JMIR Nursing

Virtualizing care from hospital to community: Mobile health, telehealth, and digital patient care
Volume 9 (2026) ISSN 2562-7600 Editor in Chief: Elizabeth Borycki, RN, PhD, FIAHIS, FACMI, FCAHS, Social Dimensions of Health Program Director, Health and Society Program Director, Office of Interdisciplinary Studies; Professor, School of Health Information Science, University of Victoria, Canada

Contents

Viewpoints

Humanistic Meaning-Making and Behavioral Governance in AI-Mediated Nursing Care (e103504)
Tino Hsieh. . . . 8

Virtual Reality–Enabled Physical AI Training for Supportive Nursing Robotics: Nurse-in-the-Loop, Site-Specific Conceptual Framework (e96848)
Seungman Kim. . . . 17

From Concept to Practice: Lessons From the Balanced Nursing Teams Decision-Support System (e92417)
Senne Vleminckx, Peter Van Bogaert, Wim De Keyser, Filip Haegdorens. . . . 29

Artificial Intelligence–Enhanced Wound Care to Improve Access, Efficacy, and Equity in Wound Care for Older Adults in Rural and Remote Regions of Canada (e85644)
Courtney Genge, Basnama Ayaz, Shannon Freeman, Heba Mohammed, Robert Fraser, Ibukun-Oluwa Abejirinde, Deirdre O'Sullivan-Drombolis, Rebecca Brookham. . . . 41

Reviews

Acceptance of Digital Technology Among Nursing Staff in Geriatric Long-Term Care: Systematic Review (e82223)
Jeton Iseni, Walter Swoboda, Daniel Houben, Roman Hilla. . . . 53

Nurses' Experiences Using AI in Clinical Practice: Systematic Review (e91238)
Ashley Scott, Qimeng Zhao, Jo-Fan Pan, Benjamin Brown, Dawn Dowding. . . . 69

Facilitating Digital Transformation in Nursing Through Nursing Development Units: Scoping Review (e89051)
Annabell Gangnus, Pascal Müller, Katrin Gayen, Uta Kirchner-Heklau, Sebastian Hofstetter, Patrick Jahn. . . . 89

Simulation-Based Training for Nursing Students to Improve Patient Safety: Systematic Review (e87898)
Azucena González-Sanz, José López-García, Elena Sutil-Rodríguez, Raúl Juárez-Vela, José Santos-Sánchez, Noelia Navas-Echazarreta, Antonio Martínez-Sabater, Consuelo Sancho-Sanchez. . . . 104

Nursing Informaticians in Spain: Scoping Review and Expert-Validated Gap Analysis (e83373)
Adrián Marco-Moyano, Patricia Verdú Rodríguez, Manuel Lillo-Crespo. . . . 127

XSL•FO
RenderX

The Shift From Efficacy to Implementation Science (2020-2026) in Nursing Practice for Digital Mental Health: Scoping Review and Bibliometric Analysis (e91498)
Shoukai Yu, Lingmei Qian, Hao Wang. . . . 154

Effectiveness of Artificial Intelligence–Based Nursing Interventions for Chronic Illness Care: Umbrella Review (e97905)
Jee Joo, Megan Liu, Youngwoo Cho, Hyungbin Cho. . . . 164

Developing a Best Practice Guideline for Clinical Practice in a Digital Health Environment: Systematic Reviews Based on the Grading of Recommendations, Assessment, Development, and Evaluation Approach (e74942)
Lauren Bailey, Lyndsay Howitt, Nafsin Nizum, Christine Buchanan, Maureen Charlebois, Jennifer Yoon, RNAO Expert Panel. . . . 183

Natural Language Processing in Nursing Practice for Clinical Reasoning and Workflow: Scoping Review (e94778)
Keisuke Iwase, Taiki Teshima, Miyae Yamakawa, Yasushi Takeya. . . . 202

Integration of Virtual and Augmented Reality Into Obstetric Nursing and Midwifery Education: Systematic Review (e90252)
Mariana Firmiano, Ana Moreira, Flaviana Vieira, Carla Sa-Couto, Alexandrina Cardoso. . . . 213

WeChat-Based Nursing Interventions in Women's Mobile Health: Systematic Review (e82679)
Jun Xiao, Rui Li, Yangyang Wu, Tong Wu. . . . 458

From Academic Integrity to Professional Formation in Generative AI–Supported Nursing Education: Integrative Review (e99284)
Peter Hartin, Paul Baretta. . . . 1036

Corrigenda and Addendas

Correction: Perceptions and Intentions of Nursing Students Regarding Digital Health: Cross-Sectional Study (e98552)
Alexandre Castonguay, Sandrine Hegg-Deloye, Guy Paré, Faustin Etindele Sosso. . . . 230

Correction: Nursing Informaticians in Spain: Scoping Review and Expert-Validated Gap Analysis (e100481)
Adrián Marco-Moyano, Patricia Verdú Rodríguez, Manuel Lillo-Crespo. . . . 231

Original Papers

Nurses' Experiences of Interprofessional Collaboration in Digitally Supported Hospital Discharge Planning: Qualitative Study (e81961)
Mera Delima, Musheer Aljaberi, Regidor Dioso. . . . 233

Explainable AI for Equitable Nurse Scheduling: Pragmatic Pre-Post Implementation Study (e94450)
Ben-Chang Shia, Szu-Ming Peng, Qui-Yang Zhang, Chiung-Yun Lo, Sheng-Ru Wang. . . . 249

The Effects of Adequate Rest on Nurse Job Satisfaction, Burnout Prevention, and Physical Health in Medical and Emergency Units at a Hospital in Western Jamaica: Qualitative Study (e84106)
Channon Smith, Keitumetse-Kabelo Murray, Natasha Croome. . . . 263

XSL•FO
RenderX

Medication-Related Errors Among Nurses by Unit Adaptation Levels: Bayesian Network–Based Exploratory Study (e87436)
Naotaka Sugimura, Katsuhiko Ogasawara. . . . 273

Implications for Virtual Nursing Role Development in Acute Nursing Care: 24-Hour Time-and-Motion Study (e87930)
Jennifer Sumner, Emily Chew, Sin Peck, Janet Lam, Pauline Chong, Mun Ching, De Ong, Jacqueline Lau, Camille Keck, Abigail Ang, Jaminah Mohamed Ali, Hui Lim, Anjali Bundele, Elya Chen, Carol Yap, Evelyn Ho, Amartya Mukhopadhyay, Alexander Yip. . . . 287

Validation and Feasibility of a Consumer-Grade Wearable Sleep Monitoring Device and an Exploratory Evaluation of Sleep Characteristics and Associated Factors Among Nurses: Observational Cohort Study (e96912)
Siew Lim, Fazila Aloweni, Phyllis Ong, Shin Ang. . . . 300

A Data Strategy for the Secondary Use of Nursing Real-World Data: Feasibility Study of the NuDaK Framework (e95546)
Renate Nantschev, Elske Ammenwerth. . . . 315

Evaluating Human-Robot Interaction With a Socially Assistive Robot in Pediatric Health Care: Mixed Methods Pilot Study (e95702)
Janika Leoste, Slavko Raki , Kristel Marmor, Kadi Krinal, Liina-Harjus Soostar, Katariina Vink, Andres Käver, Katrin Aasma, Katie Beatty, Leo Leoste. . . . 328

Social Media Use and Digital Professionalism in Nursing Practice: Cross-Sectional Questionnaire Study (e92923)
Melita Peršolja, Anton Grmšek Svetlin. . . . 338

Radar-Based Contactless Monitoring in Different Inpatient Nursing Care Settings: Mixed Methods Feasibility Study (e93624)
Madeleine Ritter-Herschbach, Sebastian Hofstetter, Dominik Behr, Melina Hager, Burkhard Mueller, Patrick Jahn. . . . 348

Changes in Nurse-Patient Communication Through Health Technologies and Nursing Practices to Recognize and Support Limited Digital Health Literacy: Qualitative Study (e82272)
Eline Dijkman, Jobke Wentzel, David Edvardsson, Carine Doggen, Constance Drossaert. . . . 360

Facilitating Teamwork With Minimal Interaction via a Mesh-Based Communication Device in a Clinical Nursing Setting: Intervention Study (e85599)
Yujiro Matsuishi, Masakazu Hirokawa, Yoshiaki Inoue, Kenji Suzuki. . . . 376

Exploring the Impact of Electronic Medical Record–Enabled Versus Paper-Based Systems on the Quality of Nursing Handover: Comparative Case-Study (e85909)
Lisa Browning, Urooj Raza Khan, Sandra Leggat, Nicholas Monypenny, James Boyd. . . . 390

Effects of a Digital Decision Support System on Diabetes Management in a Real-World Home Care Setting: Prospective Quasi-Experimental Feasibility and Implementation Study (e89687)
Katharina Lichtenegger, Lara Scholle-Winkler, Julia Kopanz, Angela Libiseller, Klaus Donsa, Felix Aberer, Thomas Pieber, Julia Mader. . . . 416

Nursing Process Data for Health Care Cost Prediction Using Machine Learning: Longitudinal Study (e93638)
María Company-Sancho, Víctor González-Chordá, María Orts-Cortés. . . . 427

Predictors of Treatment Adherence in Outpatients With Chronic Obstructive Pulmonary Disease: Cross-Sectional Study (e83132)
Hien Nguyen, Trieu Nguyen, Hoai Nguyen, Giang Luong, Doanh Vu, Thong Tran. . . . 440

XSL•FO
RenderX

The Relational Playbook Nurse Leadership Development Program Using the Whistle Systems Employee Recognition Platform: Feasibility Mixed Methods Study (e79188)
Marguerite Daus, Brigid Connelly, Drew Carter, Heather Gilmartin. . . . 450

360° Contextual Simulation Videos for Undergraduate Nursing Students: Electroencephalography-Based Quasi-Experimental Study (e84720)
Yi-Chang Chen, Chen-Jung Chen. . . . 478

The Associations of Emotional Intelligence, AI Self-Efficacy, and AI Literacy Among Nursing Undergraduates Under the NUR.S.E.S. Framework: Network Analysis (e90253)
XiaoHui Fan, Jingjing Sun, LingHui Zhang, Jinrong Wang, BinHao Dong, XiaoHong Gao, Ruihua Jin, Panpan Huai. . . . 492

Digital Health Literacy and Attitudes Toward Telehealth Use in Practice Among Nursing Students in Saudi Arabia: Cross-Sectional Study (e94722)
Muna Alharbi, Ahmed Alharthi, Abdulrahman Azab, Renad Alharthi. . . . 513

Digital Leadership Scale for Clinical Nurses: Development and Validation of an Instrument (e82101)
Ji-Hyeon Lee, Sun-Young Jung. . . . 523

Assessing Digital Literacy Among Nursing Faculty Under the Artificial Intelligence–Enhanced Technological Pedagogical Content Knowledge Framework: Cross-Sectional Survey Analysis (e88016)
Junyu Zhao, Xiaomin Hua, Rong Hu, Chu Chen, Wanxin Wu, Yaming Zhang, Yuhe Zhang, Aijing Wu, Yu Zhang. . . . 537

Informatics Competency and Technology Self-Efficacy Levels Among Undergraduate Nursing Students: Cross-Sectional Study (e85283)
Nader Alnomasy, Habib Alrashedi, Sharifah Alsayed, Petelyne Pangket, Razan Alsayed. . . . 549

Exploring the Potential for Smartphone-Based Reflection to Support Socioemotional Competence Development in Student Nurses' Clinical Placement: Mixed Methods Field Trial (e91083)
Ivo Bril-Broers, Nick Degens, Joke Fleer. . . . 561

Perceptions and Intentions of Nursing Students Regarding Digital Health: Cross-Sectional Study (e77051)
Alexandre Castonguay, Sandrine Hegg-Deloye, Guy Paré, Faustin Etindele Sosso. . . . 582

Psychometric Evaluation of the Canadian Nurse Informatics Competency Assessment Scale and the Digital-Technology Self-Efficacy Scale Among Saudi Nursing Students: Cross-Sectional Study (e88075)
Nader Alnomasy, Habib Alrashedi, Sharifah Alsayed, Petelyne Pangket, Ebtsam Abou Hashish, Razan Alsayed, Romeo Mostoles. . . . 590

Application of the Technology Acceptance Model to Predict Nursing Students' Intention to Use Informatics: Cross-Sectional Study (e85385)
Habib Alrashedi, Nader Alnomasy, Romeo Mostoles. . . . 600

Exploring Nursing Students' Experiences in a Brief Virtual Reality–Enhanced Workshop: Cross-Sectional Exploratory Study (e85780)
Niklas Carlsson, Leonie Klompstra, Mats Westas. . . . 612

The Role of Digital Health Literacy and Cultural Intelligence in Cultural Competence of Private Hospital Nurses: Cross-Sectional Study (e97334)
Emad Shdaifat, Amira Alshowkan, Bilal Alsaaideh. . . . 621

XSL•FO
RenderX

Knowledge Acquisition, Case Discussion, and Engagement in Health Professional Students Using Interactive Virtual Patient Cases Versus Written Case Studies: Randomized Controlled Trial (e98700)
Jeffrey Borckardt, Debra Henninger-Borckardt, Kelly Barth, Dusti Annan-Coultas, Lisa Langdale, Kimberly Kascak, Christopher Pelic, Michael de Arellano, Kathleen Brady. . . . 631

Train-the-Trainer Evaluation of Technology-Enhanced Learning in Preservice Midwifery Education in Nigeria: Qualitative Descriptive Study (e93596)
Isabella Garti, Nadyen Shikpup, Ophelia Tonto, Michelle Gray, Halima Abdul, Awwal Ladan. . . . 646

Preliminary Psychometric Evaluation of the Arabic Version of the NURS-DIAL Questionnaire Among Nursing Students in Saudi Arabia: Cross-Sectional Study (e106044)
Nader Alnomasy, Habib Alrashedi, Petelyne Pangket. . . . 659

Manifestations and Potential Consequences of Information Distortion in Electronic Nursing Records: Qualitative Study (e96679)
Jianan Wang, Yihong Xu, Wen Li, Zhichao Yang, Yushu Sun, Xiaoxiao Zhang, Hongying Pan. . . . 671

Nurses' Expectations of a Knowledge Management System in Nursing Practice: Qualitative Study (e78395)
Magdalena Vogt, Sebastian Müller, Glorianna Wagner-Jagfeld, Renate Ranegger, Sabin Zürcher, Janine Vetsch. . . . 683

Voice-Based Structured Nursing Documentation Using Automatic Speech Recognition and Large Language Models: Development and Evaluation Study (e88567)
Meng-Han Su, Wei-Chun Wang, Yi-Min Hsu, Shih-Yen Hou, Su-Jung Chuang, Shih-Sheng Chang. . . . 693

Barriers and Enablers for Sustaining Nurse-Led Use of Clinical Decision Support Tools for Antibiotic Stewardship: Qualitative Study (e83567)
Victoria Tiase, Aurie Tovar, Natalie Henning, Mariana Braga, Keelin McHugh, Lynn Xu, Haddy Bah, Alice Yuroff, Rachel Hess, Devin Mann, David Feldstein, Elizabeth Stevens. . . . 704

Association Between Electronic Health Record–Based Nursing Workload and Turnover: Retrospective Cohort Study (e89645)
Linlin Xia, Daphne Lew, Lindsay Tessmer, Elise Eiden, Sunny Lou, Thomas Kannampallil. . . . 715

Nurse-Led Ambient AI Scribe for Patient Safety Incident Investigation Reports (Project NARRATE): Retrospective Pre-Post Comparative Document-Quality Study (e100775)
Kai Teo, Liwen Huang, Kelly Woh, Shin Ang, Jade Ng. . . . 725

Care Partner Perspectives on Hypertension Management in Alzheimer Disease and Related Dementias in an Online Forum: Qualitative Thematic Analysis (e85818)
Jisook Ko, Kylie Meyer, Jing Wang. . . . 740

Web-Based Mindfulness Intervention for Reducing Perceived Stress in Health Care Professionals: Secondary Analysis of Moderators in a Stepped-Wedge Cluster Randomized Trial (e98624)
Gloria Guerrero-Pertiñez, Vera Carbonell Aranda, Jonathan Dawood-Histrova, Gloria Pérez-Guerrero, Adrián Pérez-Aranda, Selene Fernández-Martínez, Alicia Monreal-Bartolomé, Alberto Barceló-Soler, Javier García-Campayo, Yolanda López del Hoyo, Jesús Herrera-Imbroda, Jessica Goodman-Casanova, Jose Guzman-Parra. . . . 752

Inpatient Nurses' Anticipatory Perspectives on Assistive Robots: Qualitative Study (e91106)
Mavis Tan, Timie Chan, Charmaine Taye. . . . 766

Insights Into Factors Affecting Nurses' Knowledge of and Attitudes Toward AI and Implications for Successful AI Integration in Critical Care: Cross-Sectional Study (e85649)
Habib Alrashedi, Saad Alderaan, Nader Alnomasy, Hamdi Lamine, Khalil Saleh, Sameer Alkubati. . . . 779

Readiness Assessment for AI in Nursing Care Projects: Multimethods Study (e84148)
Kathrin Seibert, Dominik Domhoff, Janissa Altona, Sebastian Jäger, Felix Bießmann, Alessia Nowak, Rahel Gubser, Matthias Schulte-Althoff, Daniel Fürstenau, Jörg Pohle, Lea Bergmann, David Walter, Kathi Beier, Dagmar Borchers, Karin Wolf-Ostermann. . . . 789

XSL•FO
RenderX

Nurses' Evaluation of a Service Robot for Inpatient Care: Technology Acceptance Study (e86824)
Christopher Friese, Robert Klebbe, Anika Heimann-Steinert. 817

Virtual Nursing Pilot in the Inpatient Setting: Qualitative Evaluation (e76994)
Ksenia Gorbenko, Maria Bailon, Jordan Randall, James Viskochil, Clara Cualing, Patrick Healy, Madhu Mazumdar, Robert Freeman. 834

Developing a Cross-Device Platform for Robotic Systems in Nursing Care: Mixed Methods Feasibility Study (e84118)
Pascal Müller, Ruven Veit, Sebastian Hofstetter, Patrick Jahn. 846

Perception of AI Symptom Models in Oncology Nursing: Mixed Methods Evaluation Study (e82283)
Bridget Nicholson, Elizabeth Sloss, Aref Smiley, Joseph Finkelstein, Kathi Mooney. 864

Recommendations From a Health Care Provider's Perspective for Better Implementation of Telemedicine and Telenursing: Descriptive Qualitative Approach (e90610)
Rr Hariyati, Min-Huei Hsu, Hanny Handiyani, Muhammad Amiruddin, Tuti Afriani, Dara Wardhani, Satria Pranoto, Andi Wildani, Arieny Rizafni, Citra Shabrina. 876

Factors Influencing Nursing Internship Students' Readiness to Use AI: Cross-Sectional Study Using Neural Network Analysis (e92533)
Sameer Alkubati, Wesam T Almagharbeh, Talal A Alqalah, Basma Salameh, Hamdan Albaqawi, Awatif M Alrasheeday, Abdulhafith Alharbi, Layla Alshammari, Bushra Alshammari, Eddieson Pasay-an, Anwar Abdulkareem, Shimmaa Mohammed Elsayed. 891

Process Evaluation of a Digital Smartphone Intervention (URApp) to Support Adherence to Bladder Training in Community-Based Specialist Pediatric Continence Care in England: Mixed Methods Study (e95418)
Lucy Beasant, Jenny Ingram, Davina Richardson, Joanna Clothier, Carol Joinson. 903

AI Use, Perceptions, and Perceived Impact Among Nursing Students: Cross-Sectional Study (e97832)
Ilda Taka, Elona Hasalla, Albana Sula, Blerina Bahiti, Rajmonda Oboni, Blerina Bani. 916

Using Social Media Listening to Understand the Pressure Injury Experience: A Qualitative Descriptive Study (e76682)
Jessica Song, Isabel Tsang, Parthena Mouratidis, Manahil Saeed, Alma Ornelas L, Alaul Islam, Sharon Gabison. 927

Health Care Professionals' Perspectives on the Use of a Wearable Device for Early Detection and Continuous Vital Signs Monitoring of Acute Respiratory Infections in Nursing Homes: Qualitative Study (e84436)
Laura Grootegoed, Margriet Pol, Cees Hertogh, Lisa Kolodziej, Menno de Jong, Janke Schinkel, Laura van Buul. 944

Training Programs for Dementia Care Staff and Aggression in Nursing Homes: Secondary Analysis of a Cross-Sectional Online Survey (e94077)
Reed Bratches, Tamara Backhouse, Shaoqing Ge, Megan Hebdon, Dalton Riegel, Noah Stein, Katie Trainum, Rita Jablonski. 955

A One-Year Study Using Digital Biomarkers From Sensing Technologies to Assess Changes in Physical Activity Levels and Sleep Quality in Nursing Home Residents With Dementia: Observational Study (e95194)
Lydia Boyle, Monica Patrascu, Bettina Husebo, Kristoffer Haugarvoll, Ole Steihaug, Brice Marty. 965

Nurse-Led Digital Intervention for Sodium Restriction in Chronic Kidney Disease: Mixed Methods Implementation Study (e94330)
Yu Shi, Hongmei Peng, Youying Zhang, Jingyu Zhang, Yang Li, Shi Pu, Xiangchun Tang, Jinghong Zhao. 980

Patients' Perceptions of the Role of Nursing in Substance Use Disorder Treatment Programs: Qualitative Study (e82401)
Clara Lucas Guerra, Marina Gómez de Quero Córdoba, Bruno García Tardón, Guadalupe González Naranjo, Noemí Mayoral Gonzalo. 997

XSL•FO
RenderX

Multicomponent Supportive-Educative mHealth Intervention for Self-Care Among Individuals With Osteoporosis in Thailand: Exploratory Pretest-Posttest Pilot Study (e101025)
Natthida Pangutha, Sirinapha Jittimanee, Porramat Saksaen. 1010

AI-Based Structured Information Extraction From Synthetic Nursing Handover Transcripts: Comparative Evaluation of Large Language Models (e106133)
Aaron Conway, Adriana Hada, Jessica Schluter, Hui Xu, Dan Lowden, Tim Miller, Ken Donald, Andrew Teodorczuk. 1022

Thirty-Day Readmission Risk Stratification at Hospital Discharge: Single-Center Temporal Validation and Nursing Implementation Readiness Study (e101335)
Hui-Wen Po, Tzu-Yu Huang, Ying-Chien Chu, Hui-Chen Tsai, Chung-Yu Chen, Chien-Hsin Yang. 1049

The Double-Edged Sword Effect of AI Application Among Clinical Nurses Based on the Job Demands-Resources Model: Qualitative Study (e100916)
Xiaoyan Zhang, Jiaxin Fang, Sihan Chen, Jiayin Luo. 1060

Performance of Large Language Models in the Japanese Public Health Nurse National Examination: Comparative Cross-Sectional Study (e82842)
Yutaro Takahashi, Ryota Kumakura, Rie Okamoto, Shizuko Omote. 1073

Theoretical Exploration of Error Thresholds for Clinical AI Decision Support in Nursing: Exploratory Simulation Study Grounded in Human-AI Reliance Data (e99590)
Hiroyuki Tajima. 1081

Community Health Nurses' Knowledge and Perceptions of AI in Canada: National Cross-Sectional Survey (e78560)
Mary Henderson Betkus, Davina Banner, Leanne Currie, Piper Jackson, Shannon Freeman. 1095

Development of Telenursing Guidelines to Improve the Quality of Services in Diabetic Wound Care in a Hospital in Thailand: Case Study (e74228)
Chonlada Darayon, Paralee Opasanant, Siriporn Sangsrijan, Waraporn Pattaramungkhunkat, Panpimol Sukwong. 1114

Retraction

Retraction: Comparative Effectiveness of Health Communication Strategies in Nursing: A Mixed Methods Study of Internet, mHealth, and Social Media Versus Traditional Methods (e89640)
JMIR Publications Editorial Office. 1126

XSL•FO
RenderX

Humanistic Meaning-Making and Behavioral Governance in AI-Mediated Nursing Care

Tino Hsieh[1,2], MSN

[1]College of Nursing, National Yang Ming Chiao Tung University, No. 155, Sec. 2, Linong St, Beitou Dist., Taipei, Taiwan, Taiwan
[2]Department of Nursing, National Taiwan University Hospital, Taipei, Taiwan

Corresponding Author:
Tino Hsieh, MSN
College of Nursing, National Yang Ming Chiao Tung University, No. 155, Sec. 2, Linong St, Beitou Dist., Taipei, Taiwan, Taiwan

Abstract

AI has supported nursing and health care for years through decision support, prediction, monitoring, and workflow tools; what is changing is the breadth and proximity of AI-generated information to everyday care. Existing nursing AI literature has advanced AI literacy, safety, ethics, and governance, yet less attention has focused on how AI-mediated care reshapes contextual nursing judgment, patient recognition, professional behavior, and relational accountability at the point of care. Nursing digital competence may therefore increasingly need to extend beyond AI literacy toward humanistic meaning-making and behavioral governance. We offer MESH as a provisional integrative nursing lens with 4 domains: meaning-making, ethical presence, system and relational workflow governance, and human-in-the-relationship. Standardized nursing data are considered one enabling bridge through which nursing-sensitive information may remain visible within AI-supported decision environments. The MESH-action cycle—recognize, interpret, communicate, govern, and evaluate—is proposed as a 5-step practice-oriented pathway for applying these domains. MESH is not a validated competency model; its purpose is to support inquiry, education, implementation design, and empirical testing of whether AI augments or obscures nursing judgment and patient recognition.

(JMIR Nursing 2026;9:e103504) doi:10.2196/103504

KEYWORDS

artificial intelligence; AI; nursing; digital health; nursing informatics; humanistic care; clinical judgment; behavioral governance; AI-mediated care; nursing leadership; patient recognition

Introduction: AI and the Expanding Interpretive Space of Care

AI is not new to nursing or health care. Rule-based clinical decision support, automated classification, predictive models, and monitoring systems have influenced clinical work for years [1-5]. What is changing is the scope and proximity of machine-generated information to everyday care. Predictive and other AI-based technologies are increasingly used across clinical and operational functions [1,6], while generative and conversational systems can synthesize narratives, generate text, and participate more directly in clinician-patient communication [7]. AI therefore does more than add another tool to clinical workflows; it can participate in the informational and interpretive processes through which clinicians understand patients and determine care priorities.

These capabilities may improve safety, efficiency, access, and information synthesis. At the same time, as AI-enabled clinical information systems summarize, classify, and prioritize clinical information, the representation of the patient can become progressively abstracted. This raises a deeper nursing question: can care become more efficiently processed while the patient feels less personally recognized [8-10]? The concern is not limited to technical error, algorithmic bias, or inadequate literacy. It also concerns how machine-generated representations influence nursing attention, interpretation, communication, and accountability.

AI-mediated nursing care refers here to clinical situations in which AI-generated classifications, predictions, summaries, recommendations, or communications materially shape how patient information is represented, interpreted, communicated, prioritized, or acted upon within nursing workflows. It should therefore not be understood only as the use of AI tools in nursing practice. This viewpoint examines how nursing judgment, patient recognition, nursing-data visibility, professional behavior, and workflow governance interact under such conditions and proposes meaning-making, ethical presence, system and relational workflow governance, and human-in-the-relationship (MESH) as an integrative nursing lens for that interaction.

Why AI Literacy Is Necessary but Insufficient

Recent nursing AI scholarship has advanced essential conversations across several complementary areas. AI literacy

XSL•FO
RenderX

and competency work has clarified knowledge, verification, responsible use, and the capabilities nurses may need when working with AI [6,11-13]. Nursing leadership and governance scholarship has addressed implementation, sociotechnical safety, accountability, and organizational readiness [1,11,14-16]. Nursing ethics literature has emphasized dignity, respect, caring relationships, and professional accountability [9,17-19].

These contributions are necessary, but they do not fully resolve the point-of-care problem addressed here. Existing frameworks already address substantial elements of literacy, governance, communication, interpretation, and safety [1,6,11-16]. The remaining question is how these elements interact when AI becomes part of the way patients are represented, summarized, classified, and acted upon. A generated handoff note may be coherent while omitting uncertainty. A conversational AI system may provide accurate information while failing to recognize fear, shame, or confusion [7]. A staffing algorithm may optimize numbers while overlooking novice nurse support, skill mix, emotional labor, and team fatigue.

The critical question therefore shifts from "Can nurses use AI?" to "What must nursing preserve and govern when AI mediates how patients become known?" Tool use does not by itself constitute competence in AI-mediated care. Nursing competence also requires the ability to contextualize machine-generated information, recognize discordance between an output and the embodied patient, communicate uncertainty, and govern how AI changes professional behavior and clinical workflows.

Why MESH Now? From Decision Support to AI-Mediated Care

Urgency does not arise because AI has suddenly appeared in nursing. Rather, the nature and density of mediation are expanding. Earlier clinical decision support systems often delivered bounded alerts, rules, or predictions within specific tasks; contemporary generative and conversational AI systems can also summarize narratives, draft documentation, and produce patient-facing explanations, potentially functioning as an additional informational agent within clinician-patient interactions [7]. These developments increase the number of points at which machine-generated representations can shape attention, communication, and responsibility.

The need for MESH therefore rests not on technological novelty alone but on the increasing density of AI mediation across the clinical encounter. The same human factors problems observed with earlier automation—anchoring, overreliance, misuse, disuse, and automation bias—remain relevant [4,5]. What changes is the scale at which these effects may intersect with generated language, patient communication, documentation, and workflow governance.

Conceptual Positioning Among Existing Frameworks

In developing this viewpoint, we considered representative—not exhaustive—work in nursing AI literacy and competence, nursing AI governance, human factors, ethics, and implementation. This was a conceptual positioning exercise rather than a systematic literature search, and we do not claim to have excluded alternative frameworks through formal eligibility criteria.

Existing approaches already address substantial parts of the problem. The N.U.R.S.E.S. (navigate AI basics, utilize AI strategically, recognize AI pitfalls, skills support, ethics in action, and shape the future) guide emphasizes foundational AI literacy and responsible use [12]. NAIL-G (nursing artificial intelligence literacy and governance) integrates nursing AI literacy with sociotechnical safety and governance [11]. A modified Delphi study by Groeneveld et al [13] identified 10 competency domains for nurses working with AI-driven technologies, including interpretation of AI output, integration into clinical practice, communication, implementation, and evaluation. Human factors scholarship addresses automation bias, trust, use, misuse, and overreliance [4,5], while nursing ethics and relational scholarship emphasize dignity, respect, accountability, and the clinician-patient relationship [7-10,17,18].

MESH does not claim that these elements are absent from existing frameworks. Its proposed contribution is to bring them together around a nursing-specific question that becomes especially salient when AI mediates the point of care: how can nurses preserve contextual meaning-making and patient recognition while simultaneously governing behavioral drift, data visibility, workflow, and relational accountability? In this sense, MESH is an integrative lens for the interaction among domains that are often addressed separately, not a replacement for the frameworks from which those domains can be informed.

Standardized Nursing Data as a Bridge Between Humanistic Judgment and AI-Supported Decision-Making

Humanistic nursing judgment should not be positioned outside the data infrastructure of AI. Nurses continuously generate clinically meaningful information through assessments of symptoms, function, behavior, coping, care dependency, patient responses, and changes that may not yet be fully represented by conventional biomedical indicators. The challenge is that much of this knowledge may remain fragmented across narrative documentation or become insufficiently represented in the structured data used for digital decision support. As AI systems increasingly depend on structured clinical information, what becomes computationally visible may influence what becomes clinically salient.

Standardized nursing diagnoses, structured nursing assessments, Nursing Minimum Data Sets, and nursing-sensitive outcomes offer possible mechanisms for making nursing knowledge more visible within digital information infrastructures. In the protocol cited here, a Nursing Minimum Data Set is used to structure nursing diagnoses, nursing actions, and dimensions of care complexity [20]. Such standardization should not be interpreted as reducing nursing judgment to codes or as evidence that structured data automatically produce better or more humanistic AI. Rather, structured nursing data may provide one

XSL•FO
RenderX

informational bridge through which nursing-sensitive observations can become available for analysis while still requiring contextual interpretation.

Emerging evidence illustrates this potential. Cesare et al [21] used a machine learning approach to rank individual nursing diagnoses according to their predictive relevance for intensive care unit transfer. This finding should be interpreted cautiously: it does not establish that standardized nursing documentation automatically improves AI performance, nor does it imply that all dimensions of nursing judgment can be computationally encoded. It does show that structured nursing diagnoses may contain clinically meaningful information that is relevant to deterioration risk.

This distinction matters for humanistic nursing. Without adequate representation of nursing-sensitive information, AI systems may reproduce a predominantly biomedical view of the patient that underrepresents function, vulnerability, patient response, psychosocial concerns, care dependency, and other dimensions routinely recognized by nurses. Standardized nursing data are therefore treated in MESH as an enabling informational bridge—particularly for meaning-making and system and relational workflow governance—not as an additional framework domain. The relevant question is not whether nursing judgment should be human or data driven but how nursing knowledge can remain clinically meaningful, contextually interpreted, and professionally governed as portions of that knowledge become structured and computationally available.

When Machine-Generated Interpretation and the Patient Do Not Match

The following hypothetical emergency department vignette is clinically grounded and is intended to illustrate a class of machine-mediated interpretation errors rather than describe an identifiable patient.

An automated cardiac monitoring system displays a tachycardic rate. At the bedside, however, the nurse palpates a substantially slower pulse. Rather than accepting the displayed rate as definitive, the nurse reassesses the patient, reviews the waveform, obtains a more detailed medication history, and escalates the discordance for specialist review. Hospital-based continuous electrocardiogram monitoring systems can generate false-positive arrhythmia alarms, underscoring the need to interpret monitor output in conjunction with the patient and waveform rather than in isolation [22]. In this hypothetical scenario, the displayed tachycardic rate is recognized as a monitoring artifact rather than the patient's true pulse rate, while the actual bradycardia is clinically consistent with exposure to a glaucoma medication capable of systemic beta-blocking effects [23].

The clinical value of nursing judgment in this scenario lies not in rejecting technology but in recognizing that the machine-generated representation and the embodied patient do not correspond. Uncritical acceptance of the displayed rate could have prompted inappropriate tachyarrhythmia-directed evaluation or treatment. The monitor supplied a machine-generated representation; the nurse recognized the discrepancy and reconstructed its clinical meaning.

Conventional cardiac monitoring is not equivalent to contemporary generative AI. The vignette is used as an analogue of a broader machine-patient discordance problem rather than as evidence that all automated monitoring is AI. Across clinical technologies, machine-generated classifications can acquire cognitive prominence once integrated into clinical workflows. Human oversight is therefore necessary but not sufficient. Nursing requires accountable human presence: examining the patient, interpreting the clinical context, listening to the patient's story, communicating uncertainty, and taking responsibility when a technology-generated representation does not adequately reflect the person receiving care.

AI-Mediated Behavioral Drift

AI implementation can change more than information flow. As with earlier forms of automation, AI-enabled systems may alter attention, trust, questioning, communication, documentation, escalation, and responsibility [4,5]. These changes may occur gradually after implementation and may not be captured by conventional metrics such as adoption, usability, accuracy, or time saved.

We use the term "AI-mediated behavioral drift" provisionally to describe gradual, normalized postimplementation shifts across multiple professional behaviors—including attention, trust, communication, documentation, escalation, and accountability—rather than a single decision error or a synonym for automation bias. Automation bias concerns overreliance on automated output in judgment or decision-making [4]; misuse and disuse describe inappropriate patterns of automation use [5]. Behavioral drift, as proposed here, concerns how multiple AI-influenced behaviors may become normalized within clinical workflows over time. Examples may include looking first at a risk score rather than at the patient's embodied presentation, reproducing AI-generated summaries without sufficient contextual revision, or making disagreement with AI outputs difficult to document.

Such drift is not inevitable, but it warrants governance attention. In this viewpoint, behavioral governance does not mean disciplinary monitoring of individual nurses. It refers to governance of AI- and workflow-induced changes in professional attention, trust, communication, documentation, escalation, workflow, and accountability through AI system design, feedback loops, escalation structures, team learning, and nonpunitive review. Humanistic meaning-making requires calibrated trust: nurses need to judge when an output is clinically plausible, contextually incomplete, ethically concerning, or operationally useful [3-5]. Recognition of AI-clinical discordance—the mismatch between machine-generated output and bedside assessment—is therefore proposed as a candidate observable safety behavior rather than an informal act of professional intuition.

XSL•FO
RenderX

The MESH Framework

To move this discussion from abstract concern to practice-oriented inquiry, we offer MESH as a provisional conceptual framework for humanistic meaning-making and behavioral governance in AI-mediated nursing care. MESH denotes 4 domains—meaning-making, ethical presence, system and relational workflow governance, and human-in-the-relationship—and should not be confused with Medical Subject Headings (MeSH). In this domain, "system" refers to the sociotechnical health care environment and unit-level workflow in which AI is implemented, rather than only to an electronic health record system. MESH is offered in conversation with existing nursing AI literacy, competence, leadership, and governance frameworks [11-14]. It is not presented as a validated theory, complete competency model, or exhaustive taxonomy.

Meaning-making refers to the nursing capacity to interpret AI outputs through patient narratives, embodied assessment, family concerns, clinical trajectory, structured nursing data, and available resources. Ethical presence concerns the moral obligations that remain active during AI-mediated care, including dignity, autonomy, equity, advocacy, vulnerability, uncertainty, and moral accountability [8-10,17,18]. System and relational workflow governance concerns how AI system design, health care leadership, and unit-level workflow shape documentation, escalation, responsibility, invisible labor, and relational burden. Human-in-the-relationship concerns accountable human relational presence: the nurse remains responsible for listening, contextualizing, explaining uncertainty, responding to patient and family questions, supporting shared understanding, and ensuring that automated communication does not substitute for therapeutic interaction.

MESH is not intended to replace existing AI literacy, competency, ethics, leadership, or governance approaches [6,11-15,17-19], nor does it claim novelty for its component concepts. Its proposed contribution is their point-of-care integration around AI-mediated patient representation, nursing meaning-making, behavioral adaptation, nursing-data visibility, workflow governance, and relational accountability, together with translation of these concerns into provisional targets for education, implementation, audit, and empirical study.

What MESH Adds to Existing Frameworks

MESH does not aim to replace established nursing theories, AI competency frameworks, implementation science, human factors models, or approaches to AI literacy. Each contributes important concepts for understanding professional judgment, technology adoption, workflow, safety, ethics, and person-centered care [4,5,11-14]. The distinctive purpose of MESH is narrower: to organize these concerns around the conditions created when machine-generated information becomes an active intermediary in nursing care.

In AI-mediated practice, interpretation, relational recognition, behavioral adaptation, data representation, and workflow governance do not occur independently. An AI-generated recommendation may simultaneously redirect a nurse's attention, influence how a patient is interpreted, alter communication, modify escalation behavior, and redistribute accountability within workflow. Frameworks designed primarily to explain technology adoption, professional competency, ethical principles, or health care technology implementation may illuminate individual dimensions of this process without making their interaction at the point of nursing care explicit.

MESH therefore functions as an integrative nursing lens rather than a competing grand theory. Its contribution lies in connecting 4 concurrent questions: What does the AI output mean in this patient's context? What moral concerns and dimensions of the person remain visible? How is the technology shaping professional behavior, data visibility, and workflow? And who remains relationally present and accountable when nurses, patients, families, and AI-generated information interact? Its proposed contribution lies in this point-of-care integration and in translating these questions into observable targets for education, implementation, audit, and empirical research.

Operationalizing MESH in Clinical Practice

The practical value of MESH depends on whether its domains can be observed, taught, audited, and empirically examined. Table 1 therefore links each domain to a clinical question, observable indicators, implementation examples, and potential research directions. The indicators are intentionally provisional; they are not presented as validated measures. Rather, they illustrate how concepts such as patient recognition and behavioral drift can be converted into testable implementation constructs.

XSL•FO
RenderX

Table . Operationalizing the MESH (meaning-making, ethical presence, system and relational workflow governance, and human-in-the-relationship) framework in AI-mediated nursing care.

MESH domain	Key clinical question	Observable indicators	Example implementation	Future empirical direction
Meaning-making	Does the AI output improve or distort the interpretation of this patient's situation?	Nurse compares output with bedside assessment, patient narrative, trajectory, family context, and structured nursing information; AI-clinical discordance is recognized and investigated as a candidate safety behavior.	A cardiac monitor displays a tachycardic rate that conflicts with the palpated pulse; the nurse reassesses, reviews the waveform and medication history, and escalates the discrepancy rather than treating the machine-generated display as definitive.	Candidate indicators: frequency and resolution of AI-clinical discordance, override patterns, missed contextual cues, and escalation consequences.
Ethical presence	Does AI-mediated care preserve moral obligations related to dignity, autonomy, equity, advocacy, vulnerability, and uncertainty?	Patient preferences, vulnerability, goals, concerns, and ethically relevant uncertainty remain visible in documentation and decision-making.	AI-generated discharge information is reviewed and adapted to the patient's comprehension, concerns, family or community support, and ability to self-manage.	Candidate indicators: omission of preference-sensitive information, perceived respect, trust, and individualized care measures.
System and relational workflow governance	How does AI alter attention, escalation, accountability, workload, documentation, and relational work?	Defined escalation and override pathways, audit trail, alert and documentation burden monitoring, frontline feedback, and nonpunitive review of recurrent AI-influenced workflow problems.	A unit reviews recurrent overrides, alert fatigue, delayed responses, documentation revisions, and examples of behavioral drift at regular governance meetings.	Candidate indicators: alert fatigue, automation bias, behavioral drift, escalation performance, workload, and documentation burden.
Human-in-the-relationship	Does AI augment rather than substitute accountable human relational presence?	Nurse listens, contextualizes, explains uncertainty, responds to questions, supports shared understanding, and remains accountable for communication.	AI prepares preliminary education or summarizes information; the nurse verifies it, explains uncertainty, and contextualizes it through conversation and teach-back.	Candidate indicators: patient comprehension, shared understanding, teach-back performance, trust, and perceived relational quality.

The MESH-Action Cycle

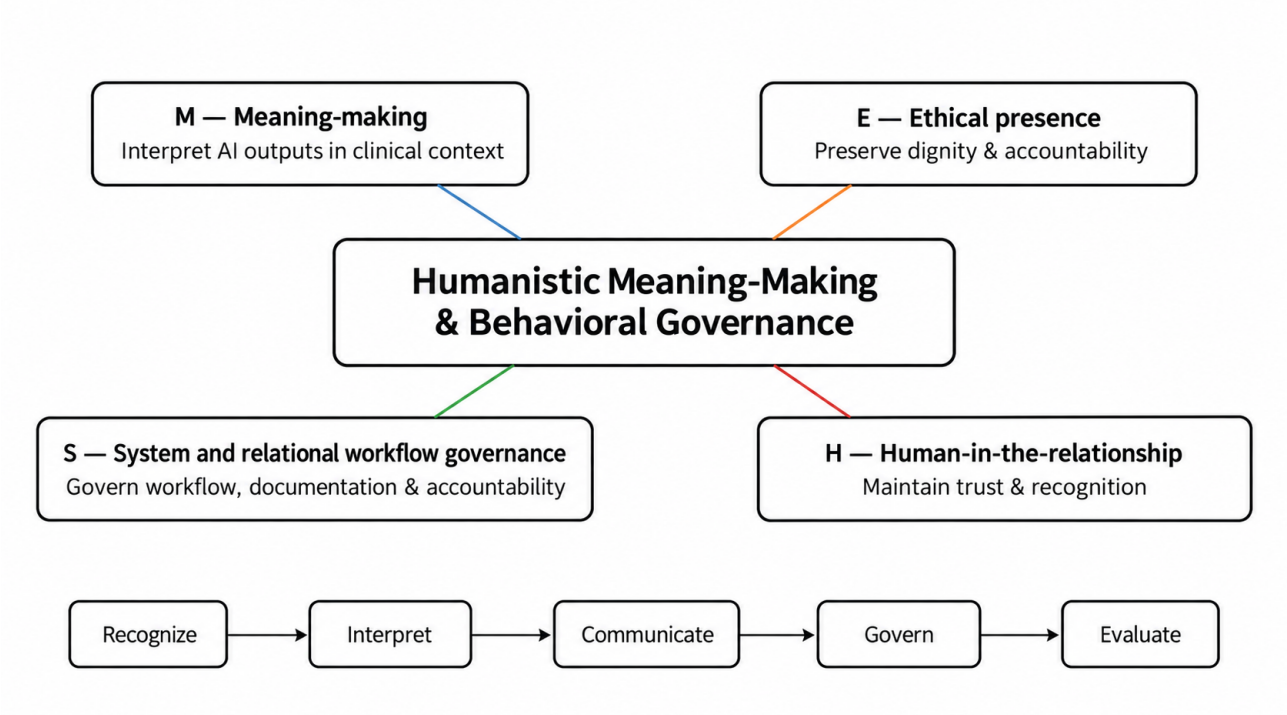
A conceptual lens that identifies a problem but does not support application is of limited value for nursing leadership. We therefore pair the 4 MESH domains with a 5-step MESH-action cycle: recognize, interpret, communicate, govern, and evaluate. MESH identifies what should be examined and preserved; the action cycle is a practice-oriented application pathway for beginning that work, not a second framework or theory. Each step draws on established bodies of knowledge. Figure 1 illustrates the relationship between the four MESH domains and the MESH-action cycle. Table 2 summarizes the five steps of the MESH-action cycle, including the corresponding leadership actions, intended purposes, and supporting knowledge bases.

XSL•FO
RenderX

Table . MESH (meaning-making, ethical presence, system and relational workflow governance, and human-in-the-relationship)-action cycle for nurse leaders.

Action	Nurse leaders should...	Purpose	Starting knowledge base
Recognize	Identify how AI reshapes attention, trust, communication, documentation, data visibility, and accountability	Detect behavioral drift and representational gaps early	Human factors, usability, and safety science [4,5]
Interpret	Train nurses to compare AI outputs with bedside assessment, patient narratives, family concerns, structured nursing data, and clinical trajectory	Prevent automation bias, premature closure, and unexamined discordance	Clinical reasoning, automation bias, calibrated trust, and AI competency work [3-5,13]
Communicate	Prepare nurses to explain AI-supported care and uncertainty in language that preserves trust and understanding	Maintain patient and family recognition and participation	Therapeutic and relational communication, respect, and shared understanding [7-10,13]
Govern	Establish escalation, override, documentation, and feedback pathways when AI and nursing judgment diverge	Protect accountability, safety, and professional agency	Sociotechnical safety, implementation science, and nursing leadership and governance [1,11,14,15]
Evaluate	Use nursing-sensitive, behavioral, relational, and humanistic indicators in addition to accuracy, adoption, and efficiency metrics	Ensure AI-supported care remains clinically meaningful and humane	Nursing-sensitive outcomes and implementation evaluation [1,13,15,16]; human factors and automation research [4,5]

Figure 1. The MESH (meaning-making, ethical presence, system and relational workflow governance, and human-in-the-relationship) framework and MESH-action cycle for AI-mediated nursing care. The 4 MESH domains identify what nursing should examine and preserve in AI-mediated care, whereas the 5-step MESH-action cycle provides a practice-oriented pathway for applying these domains in leadership, education, implementation, and evaluation; it is not a second framework or theory.



"Recognize" can draw on human factors and safety science to identify AI-induced changes in attention, trust, communication, data visibility, workflow, and accountability before they become normalized [4,5]. "Interpret" builds on clinical reasoning, automation bias research, calibrated trust, and AI competency work to reconcile machine output with the embodied patient, narrative, nursing data, trajectory, and context [3-5,13]. "Communicate" focuses on AI-supported information, uncertainty, therapeutic communication, and shared understanding [7-10,13]. "Govern" draws on sociotechnical safety, implementation, and nursing leadership scholarship to establish escalation, override, accountability, audit trails, and feedback systems [1,11,14,15]. "Evaluate" extends technical performance assessment to candidate nursing-sensitive,

XSL•FO
RenderX

behavioral, relational, workload, and patient-experience consequences [1,4,5,13,15,16].

Potential indicators include AI-nursing judgment discordance, documented rationales for overriding AI recommendations, alert burden, AI-generated documentation revision rates, patient understanding of AI-supported care, family communication concerns, staff confidence, missed care, perceived moral distress, and nursing-sensitive outcomes. These indicators do not replace technical evaluation; they broaden evaluation to the human and professional consequences of implementation.

AI as an Enabler and Risk Amplifier of Humanistic Care

AI should not be framed only as a threat to humanistic nursing. Reviews of AI-based technologies in nursing describe applications in decision support, prediction, education, and workflow-related functions [1,6,16]. Generative and conversational systems may additionally support information synthesis and communication [7]. Whether these functions actually reduce cognitive or administrative burden or create more time for relational care is an empirical question rather than an assumed benefit. Under appropriate conditions, however, AI may support rather than diminish humanistic care.

At the same time, the humanistic consequences of AI are not inherent properties of the technology; they are produced through its design, data, interpretation, implementation, and governance. The same AI-enabled clinical information system that surfaces a clinically meaningful pattern can also promote automation bias; the same generated summary that reduces documentation effort can narrow a clinical narrative; and the same decision support tool that standardizes care can shift responsibility or normalize behavioral drift [3-5,7].

The relevant distinction is therefore not humanistic nursing vs AI but humanistically governed AI vs poorly governed AI-mediated care. AI outputs should function as inputs into nursing meaning-making, not as substitutes for it. The objective of governance is to preserve the benefits of computational support while maintaining contextual judgment, patient recognition, and professional accountability.

Professional Advocacy and Leadership Implications

Nursing should not be positioned merely as the workforce that adapts to AI-enabled technologies after they are designed, purchased, and implemented [1,14,15,17,18]. Nurses and nurse leaders are governance actors because AI-enabled clinical information systems can alter clinical attention, documentation, escalation, workload, patient communication, and the visibility of nursing knowledge.

At the unit level, MESH may guide organizations to examine more than technical infrastructure and staff training. Candidate governance questions include where AI-generated information enters clinical workflows, what nursing-sensitive information is included or omitted, who remains accountable when AI and bedside assessment diverge, whether escalation and override pathways are available, how alert and documentation burden change, and whether frontline feedback identifies recurrent mismatches between AI-enabled clinical information system outputs and clinical reality. These questions are intended for organizational learning rather than individual surveillance or blame.

For nursing education, MESH may extend AI literacy toward AI-mediated clinical reasoning [4-6,12]. Simulation could expose learners to plausible but incomplete outputs, conflicting bedside findings, and situations in which AI-generated communication requires contextual revision. Training can cultivate calibrated trust rather than blind compliance or defensive rejection.

For research, validation should proceed in stages rather than assume that MESH improves outcomes. Initial qualitative work and stakeholder consultation could refine domain boundaries and language. Delphi studies could then assess content relevance and clarity among frontline nurses, nurse leaders, informaticians, patients, and ethicists. Subsequent work could develop candidate indicators or audit items and test them in simulation or observational studies, including recognition of AI-clinical discordance and override behavior. Only after conceptual and operational refinement should implementation studies examine feasibility, acceptability, workflow effects, patient experience, and associations with nursing-sensitive outcomes.

Limitations and Future Validation

The MESH framework is proposed as a provisional conceptual framework rather than an empirically validated model. Its purpose is to organize emerging concerns in AI-mediated nursing care and generate practice-oriented questions for education, leadership, implementation, and research. We do not claim that the 4 domains are exhaustive, universally applicable, psychometrically established, or unique in their component concepts.

The framework's proposed contribution is integrative rather than exclusive. The conceptual positioning in this viewpoint was based on a nonsystematic comparison of representative frameworks and was not intended to establish that MESH uniquely incorporates concepts absent from all previous models. Future work should test whether its domains are distinct, useful, measurable, and applicable across clinical contexts. The proposed operational indicators require further empirical development and validation, while the relationship between standardized nursing data and AI-supported decision-making warrants investigation beyond its early applications in predictive modeling. MESH should therefore be understood as a starting language for inquiry and governance, not as a final competency standard.

Conclusions

AI will increasingly shape nursing practice, but the central issue is not whether nursing becomes more or less technological. It is whether nursing judgment, nursing-sensitive information, patient recognition, and professional accountability remain visible and actionable within technology-mediated care. Existing

XSL•FO
RenderX

AI literacy and governance frameworks provide an essential foundation [6,11,12,14,17-19]; an important next task is to examine what happens when AI enters not only the workflow but also the process through which patients become known [7-10].

To move from general readiness to practice-level evaluation, nursing needs a structured way to examine how AI-mediated care affects interpretation, patient recognition, workflow, and relationships at the point of care. This viewpoint offers MESH as a provisional integrative nursing lens and the MESH-action cycle as a practice-oriented application pathway for leadership, education, implementation, and research. The central claim is not that nursing should resist AI or that MESH introduces unprecedented component concepts. Its proposed value lies in integrating nursing meaning-making, ethical presence, nursing-data visibility, AI-induced behavioral and workflow change, and accountable human presence at the point where AI mediates how the patient is represented and understood.

When AI helps the health care system detect patterns in patient data, nursing must ensure that the patient remains recognized as a person whose experiences, vulnerabilities, preferences, and context matter.

Generative AI Use

During manuscript preparation and revision, the author used ChatGPT (GPT-5.6 Sol; OpenAI, 2026) to assist with language editing, organization and refinement of arguments, drafting and refinement of responses to peer-review comments, consistency checking, and preparation of the schematic layout for Figure 1 based on the author-developed MESH concepts. The author critically reviewed and revised all AI-assisted output, independently verified factual claims and cited sources, and takes full responsibility for the final content of the manuscript. Generative AI was not used as an author or as an independent source of evidence.

Funding

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

Data sharing is not applicable to this viewpoint because no datasets were generated or analyzed.

Authors' Contributions

TH conceptualized the manuscript, developed the MESH framework and MESH-action cycle, drafted and revised the manuscript, and approved the final version for submission.

Conflicts of Interest

None declared.

References

1. von Gerich H, Moen H, Block LJ, et al. Artificial intelligence -based technologies in nursing: a scoping literature review of the evidence. Int J Nurs Stud 2022 Mar;127:104153. [doi: 10.1016/j.ijnurstu.2021.104153] [Medline: 35092870]
2. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019 Jan;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
3. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. N Engl J Med 2018 Mar 15;378(11):981-983. [doi: 10.1056/NEJMp1714229] [Medline: 29539284]
4. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc 2012;19(1):121-127. [doi: 10.1136/amiajnl-2011-000089] [Medline: 21685142]
5. Parasuraman R, Riley V. Humans and automation: use, misuse, disuse, abuse. Hum Factors 1997 Jun;39(2):230-253. [doi: 10.1518/001872097778543886]
6. El Arab RA, Al Moosa OA, Abuadas FH, Somerville J. The role of AI in nursing education and practice: umbrella review. J Med Internet Res 2025 Apr 4;27:e69881. [doi: 10.2196/69881] [Medline: 40072926]
7. Campos HO, Wolfe D, Luan H, Sim I. Generative AI as third agent: large language models and the transformation of the clinician-patient relationship. J Particip Med 2025 Aug 11;17:e68146. [doi: 10.2196/68146] [Medline: 40826831]
8. Sim I, Cassel C. The ethics of relational AI—expanding and implementing the Belmont principles. N Engl J Med 2024 Jul 18;391(3):193-196. [doi: 10.1056/NEJMp2314771] [Medline: 39007542]
9. Paladino MS. Artificial intelligence in nursing care: a reflection from the ethics of care. Nurs Ethics 2025 Dec;32(8):2477-2489. [doi: 10.1177/09697330251366619] [Medline: 40796224]
10. Beach MC, Duggan PS, Cassel CK, Geller G. What does “respect” mean? Exploring the moral obligation of health professionals to respect patients. J Gen Intern Med 2007 May;22(5):692-695. [doi: 10.1007/s11606-006-0054-7] [Medline: 17443381]

XSL•FO
RenderX

11. Ibrahim FM, Shahrour G, Abuijlan IA, Abdallah AM, Alsuwayed A, Alanazi AR. NAIL-G: a conceptual best-fit framework for nursing AI literacy and governance anchored in socio-technical safety science. BMC Nurs 2026 Apr 11;25(1):370. [doi: 10.1186/s12912-026-04555-6] [Medline: 41965650]
12. Hoelscher SH, Pugh A. N.U.R.S.E.S. embracing artificial intelligence: a guide to artificial intelligence literacy for the nursing profession. Nurs Outlook 2025;73(4):102466. [doi: 10.1016/j.outlook.2025.102466] [Medline: 40561583]
13. Groeneveld SW, van Os-Medendorp H, van Gemert-Pijnen JE, et al. Essential competencies of nurses working with AI-driven lifestyle monitoring in long-term care: a modified Delphi study. Nurse Educ Today 2025 Jun;149:106659. [doi: 10.1016/j.nedt.2025.106659] [Medline: 40056483]
14. Burford JS, Booth RG, McIntyre A. Exploring the intersection of nursing leadership and artificial intelligence: scoping review. JMIR Nurs 2025 Nov 14;8:e80085. [doi: 10.2196/80085] [Medline: 41237291]
15. Ronquillo CE, Peltonen LM, Pruinelli L, et al. Artificial intelligence in nursing: priorities and opportunities from an international invitational think-tank of the Nursing and Artificial Intelligence Leadership Collaborative. J Adv Nurs 2021 Sep;77(9):3707-3717. [doi: 10.1111/jan.14855] [Medline: 34003504]
16. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on the domains of nursing: scoping review. JMIR Nurs 2020 Dec 17;3(1):e23939. [doi: 10.2196/23939] [Medline: 34406963]
17. Stokes L. ANA position statement: the ethical use of artificial intelligence in nursing practice. Online J Issues Nurs 2025;30(2). [doi: 10.3912/OJIN.Vol30No02PoSCol03]
18. 2025 code of ethics for nurses provision 7.5: considerations related to ethics, technology, and policy. American Nurses Association. 2025. URL: https://codeofethics.ana.org/provision-7-5 [accessed 2026-08-22]
19. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell 2019;1:389-399. [doi: 10.1038/s42256-019-0088-2]
20. Cesare M, D'Agostino F, Cocchieri A. Exploring the association between complexity of care, medical complexity, and length of stay in the paediatric setting using a Nursing Minimum Data Set: a study protocol. Nurs Rep 2024 Oct 10;14(4):2923-2934. [doi: 10.3390/nursrep14040213] [Medline: 39449450]
21. Cesare M, Nurchis MC, Damiani G, Cocchieri A, Nursing and Public Health Group. Ranking nursing diagnoses by predictive relevance for intensive care unit transfer risk in adult and pediatric patients: a machine learning approach with random forest. Healthcare (Basel) 2025;13(11):1339. [doi: 10.3390/healthcare13111339] [Medline: 40508953]
22. Cosentino N, Zhang X, Farrar EJ, et al. Performance comparison of 6 in-hospital patient monitoring systems in the detection and alarm of ventricular cardiac arrhythmias. Cardiovasc Digit Health J 2024;5(2):70-77. [doi: 10.1016/j.cvdhj.2024.02.002] [Medline: 38765622]
23. Pratt NL, Ramsay EN, Kalisch Ellett LM, Nguyen TA, Roughead EE. Association between ophthalmic timolol and hospitalisation for bradycardia. J Ophthalmol 2015;2015:567387. [doi: 10.1155/2015/567387] [Medline: 25874117]

Abbreviations

MeSH: Medical Subject Headings
MESH: meaning-making, ethical presence, system and relational workflow governance, and human-in-the-relationship
N.U.R.S.E.S: navigate AI basics, utilize AI strategically, recognize AI pitfalls, skills support, ethics in action, and shape the future
NAIL-G: nursing artificial intelligence literacy and governance

Edited by E Borycki; submitted 03.Jun.2026; peer-reviewed by E Barey, M Cesare; revised version received 07.Sep.2026; accepted 21.Oct.2026; published 06.Nov.2026.

Please cite as:
Hsieh T
Humanistic Meaning-Making and Behavioral Governance in AI-Mediated Nursing Care
JMIR Nursing 2026;9:e103504
URL: https://nursing.jmir.org/2026/1/e103504
doi:10.2196/103504

© Tino Hsieh. Originally published in JMIR Nursing (https://nursing.jmir.org), 6.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Virtual Reality–Enabled Physical AI Training for Supportive Nursing Robotics: Nurse-in-the-Loop, Site-Specific Conceptual Framework

Seungman Kim, PhD, MEd

School of Nursing, Texas Tech University Health Sciences Center, 3601 4th Street, Lubbock, TX, United States

Corresponding Author:
Seungman Kim, PhD, MEd
School of Nursing, Texas Tech University Health Sciences Center, 3601 4th Street, Lubbock, TX, United States

Abstract

Interest in physical AI and robotics in health care is increasing, but the nursing literature shows that the evidence base remains early, nurse-centered applications are underdeveloped, and real-world experiential evidence is limited. Nurses are more likely to accept robots that reduce physically demanding and repetitive work while preserving the interpersonal and judgment-intensive core of nursing practice. This conceptual paper proposes a nursing-centered framework in which virtual reality functions not merely as a simulator but as a scaffolded training infrastructure for supportive physical AI systems, enabling nurse augmentation, site-specific adaptation through digital twins, and staged simulation-to-real transfer. The framework was developed through a conceptually integrative and implementation-aware synthesis drawing on nursing robotics, AI in nursing, immersive simulation, digital twins, human-in-the-loop learning, and physical AI development literature. It is organized around 5 linked elements and supported by a 4-layer technical architecture that outlines functional requirements and implementation pathways. Four key propositions ground the framework: (1) nursing robot training should focus on competency formation, not on decontextualized data accumulation; (2) training should proceed through progressive fidelity and staged autonomy; (3) digital twins should function as operational bridges for local ward adaptation; and (4) simulation-to-real transfer should be governed by explicit nursing-relevant validation criteria and retained human accountability. The proposed nurse-in-the-loop, site-specific virtual reality framework offers a nursing-centered complement to general-purpose physical AI pipelines by making workflow fit, role boundaries, local adaptation, and governed transfer explicit design requirements.

(JMIR Nursing 2026;9:e96848) doi:10.2196/96848

KEYWORDS

virtual reality; physical AI; nursing robotics; digital twin; simulation-based learning; nurse augmentation; human-in-the-loop learning; simulation-to-real transfer; nursing informatics

Introduction

AI in nursing has advanced rapidly in decision support, workflow optimization, and documentation, yet embodied AI for nursing remains underdeveloped as a clinically grounded field. Recent umbrella and scoping reviews show unresolved ethical, social, and implementation questions, while robotics reviews suggest that the evidence base is early and that explicitly nurse-centered robotic support remains limited [1,2]. Shaw and Chen [3] identify physical AI as a nascent but consequential frontier for nursing, emphasizing that robotics applications must be evaluated through nursing-specific lenses rather than imported wholesale from industrial or general-purpose contexts. Qualitative synthesis further confirms that nurses perceive AI adoption as acceptable only when it preserves professional autonomy and relational care [4].

Nursing work is not simply a set of motor actions. It is dynamic, interruption-prone, socially dense, ethically constrained, and deeply local. A single nurse on a medical-surgical unit may manage 6 or more patients simultaneously, responding to unpredictable deteriorations, coordinating with multiple disciplines, and making real-time prioritization decisions that cannot be reduced to algorithmic rules. Whether a robot can function appropriately in this context depends not only on navigation or manipulation capabilities but also on understanding role boundaries, fitting into existing workflows without creating new burdens, supporting appropriate escalation when uncertainty arises, and preserving the human clinical judgment that remains the irreducible core of safe patient care [5,6]. These requirements are not peripheral constraints to be addressed after technical development; they are the design specifications that should drive development from the outset.

Industry-scale physical AI platforms such as Cosmos (NVIDIA) and Isaac Sim (NVIDIA) provide powerful enabling technologies—world foundation models, synthetic data generation, physically based simulation—that future nursing robotics will leverage. These platforms have demonstrated remarkable progress in warehouse logistics, manufacturing, and autonomous driving, where environments are relatively structured and task specifications can be defined independently of individual human relationships. However, these platforms

XSL•FO
RenderX

are optimized for generalizable autonomy at scale, whereas nursing requires an additional nurse-centered orchestration layer that addresses workflow fit, role clarity, safety governance, and local adaptation as first-class design requirements. The gap is not primarily technological; it is translational. The enabling components exist, but the orchestration layer that connects them to nursing's specific clinical, ethical, and organizational realities does not.

This paper proposes a nurse-in-the-loop virtual reality (VR) framework with site-specific digital twin alignment and governed simulation-to-real transfer. The framework contributes both a conceptual structure and a practical technical architecture, the latter detailed in Multimedia Appendix 1 to support the assessment of conceptual plausibility and future empirical implementation.

The framework was developed through a conceptually integrative synthesis rather than a systematic review, and the procedure is made explicit here for transparency. Source literature was identified purposively across 6 domains—nursing robotics, AI in nursing, immersive simulation, digital twins, human-in-the-loop learning, and physical AI development—prioritizing peer-reviewed work, recent reviews and authoritative reports (including systematic reviews, umbrella reviews, and a National Academies report), and sources judged relevant to the translational question of how nursing expertise can govern robot learning. This selection was purposive and interpretive rather than exhaustive or reproducible in the sense of a systematic search protocol. Synthesis proceeded thematically. Recurring requirements and constraints were extracted across these domains and grouped into the dimensions that structure the framework: task decomposition, expert-in-the-loop learning, fidelity progression, local adaptation, and governed transfer. Conceptual abstraction then moved from specific technologies and empirical findings to architecture-level principles, distilling the recurring patterns into the 4 propositions that ground the framework and into the 5 linked elements and 4-layer architecture that operationalize them. The framework is therefore a structured, literature-grounded hypothesis whose components are intended to be tested empirically rather than a set of validated findings.

Why Nursing Requires a Distinct Physical AI Training Paradigm

Reviews consistently show nurses prefer robots that reduce physically burdensome repetitive work while preserving interpersonal and judgment-intensive care [7]. A systematic review of collaborative robots found that most applications are patient-centered rather than nurse-centered, and an umbrella review identified persistent concerns around patient safety, liability, and ethics [1,2]. This pattern reveals a fundamental misalignment in how nursing robotics has been conceived: robots are designed for tasks defined by engineers or physicians, then offered to nurses as end users, rather than developed from the outset around nurse-identified needs and nurse-defined boundaries. The consequence is a literature rich in technical demonstrations but thin on evidence of sustained clinical adoption, workflow integration, or measurable workload relief for nursing staff.

The relevant design question is therefore not "How can we maximize robotic autonomy?" but "Which parts of nursing work are appropriate for bounded robotic support, under which workflow and safety constraints, and with what forms of nurse oversight?" This reframing has several implications. It shifts the unit of analysis from the robot's capabilities to the nurse's workflow. It requires that task selection be informed by practicing nurses rather than inferred from external observation. It demands that autonomy levels be calibrated to clinical risk rather than maximized for technical elegance, and it positions evaluation around nursing-relevant outcomes, including workload relief, trust, care quality, and escalation appropriateness, rather than task success rates in isolation. Collectively, these shifts define a translational agenda that positions nurses as architects of robotic scope rather than recipients of autonomous systems.

A Nurse-in-the-Loop, Site-Specific Conceptual Framework

Framework Structure and Training Progression

The framework organizes VR-based physical AI training into 5 linked elements, summarized in Table 1. Figure 1 summarizes the proposed architecture and its feedback loops. In this framework, progressive fidelity refers to the structured increase of environmental and behavioral complexity across training tiers; staged autonomy refers to calibrating the robot's permitted actions to clinical risk; and governed transfer refers to treating simulation-to-real progression as a gated clinical-governance decision rather than a purely technical step. The 5 elements proceed sequentially but include feedback loops: nurse-in-the-loop corrections at each fidelity stage feedback to refine the task library and role-boundary map, while digital twin synchronization provides continuous environmental grounding across all training stages. The governed transfer element enforces stage-gate progression, ensuring that advancement from simulation to physical deployment is conditioned on explicit nursing-relevant validation criteria rather than technical metrics alone.

XSL•FO
RenderX

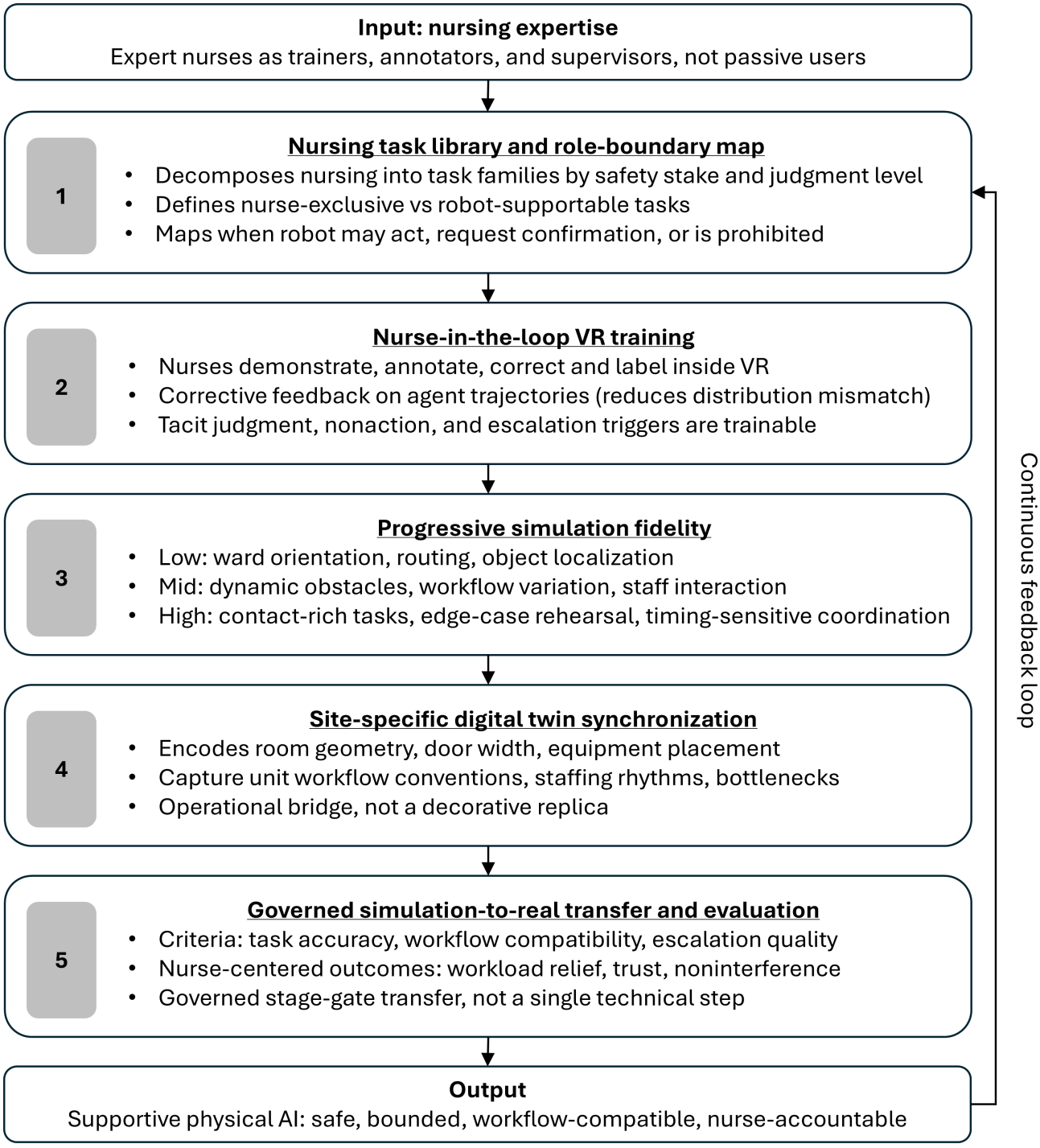
Table . The 5 linked elements of the nurse-in-the-loop, site-specific framework, with the purpose and core mechanism of each.

Element	Purpose	Core mechanism and output
Nursing task library and role-boundary map	Define which nursing tasks are appropriate for bounded robotic support and at what level of autonomy	Decomposes nursing into task families by safety stakes, paired with a 4-zone boundary map: independent execution, nurse-confirmed execution, gated expansion, and permanently prohibited action
Nurse-in-the-loop VR[a] training	Capture expert nursing judgment as the primary training signal rather than treating nurses as end users	Inside VR, nurses demonstrate tasks, annotate context, correct trajectories, label unsafe actions with explanations, and set escalation thresholds, yielding multimodal input richer than conventional labeled datasets
Progressive simulation fidelity	Build robot competence incrementally and avoid premature exposure to clinical complexity	Structured low-fidelity, midfidelity, and high-fidelity tiers, with promotion to a higher tier contingent on demonstrated performance at the current tier
Site-specific digital twin synchronization	Adapt generic robot capabilities to the physical and operational realities of one clinical unit	A continuously synchronized operational model constrained to variables that materially affect task performance: room geometry, equipment, circulation pathways, staffing density, and workflow conventions
Governed simulation-to-real transfer and evaluation	Condition physical deployment on nursing-relevant validation rather than on technical metrics alone	Staged validation gates, each with explicit sign-off and mandatory regression on degradation, evaluated against workload relief, nurse trust, and care quality

[a]VR: virtual reality.

XSL•FO
RenderX

Figure 1. Nurse-in-the-loop, site-specific conceptual framework for virtual reality (VR)–based physical AI training in nursing.



Nursing Task Library and Role-Boundary Map

Rather than treating "nursing" as a single training domain, the framework decomposes nursing into task families with different safety stakes, data needs, and permissible autonomy levels. This decomposition is essential because the spectrum of nursing work ranges from logistical tasks with minimal patient contact (eg, supply transport between a storage room and a nurses' station) to high-acuity interventions requiring moment-to-moment clinical judgment (eg, responding to hemodynamic instability). Treating these as a single training domain would either constrain the robot to trivial functions or expose patients to premature autonomy in high-stakes situations.

Near-term supportable tasks include supply transport, room-to-room delivery, retrieval of noncontrolled materials, and bounded monitoring-support workflows such as periodic vital sign equipment checks. The framework's clinical scope is delimited accordingly: it is confined to supportive tasks—noncontact logistical and environmental tasks in its initial scope, with possible extension to direct physical assistance such as patient repositioning support only under the staged validation gates—and it permanently excludes surgical or otherwise invasive intervention, medication administration, and independent clinical judgment, including autonomous decisions based on the interpretation of physiological signals, all of which remain with the nurse and clinical team. The task library is

XSL•FO
RenderX

paired with a role-boundary map that defines 4 zones of permissible action: independent execution (low-risk, logistical tasks with no patient contact), nurse-confirmed execution (tasks requiring verification before action, such as delivering items to a specific patient room), gated expansion (tasks involving direct physical assistance, such as patient repositioning support, which fall outside the autonomy permitted at the framework's initial stage and may be introduced only progressively under the staged validation gates described in layer 4), and permanently prohibited action (medication administration and independent clinical decision-making, which remain reserved to the nurse at all stages). This graduated boundary structure ensures that autonomy is not a binary on or off switch but a spectrum aligned with clinical risk, in which direct-contact assistance is neither categorically excluded nor permitted at the outset but is gated behind staged, nurse-governed validation.

Nurse-in-the-Loop VR Training

Nurses are expert contributors to policy formation, not passive end users. This distinction is not merely rhetorical; it determines the entire data collection and learning paradigm. In VR, nurses serve multiple roles simultaneously: they demonstrate tasks by performing them in the simulated environment; annotate contextual factors that affect task execution (eg, "this corridor is typically congested at shift change"); correct robot trajectories when the agent's behavior deviates from acceptable practice, label actions as unsafe or inappropriate with explanations; and define the escalation thresholds that determine when the robot should pause and request human guidance. This multimodal expert input is substantially richer than the labeled datasets typically used in supervised robot learning, because it captures not only what to do but also why, when, and under what conditions actions are appropriate or inappropriate. The approach is consistent with interactive imitation learning, which demonstrates that corrective feedback on agent trajectories in the agent's own state distribution reduces compounding errors more effectively than offline demonstration alone [8]. It also aligns with human-in-the-loop optimization research showing that iterative human participation in the learning loop produces applied robotic systems that better match real-world task requirements [9].

Progressive Simulation Fidelity

Fidelity is matched to learning objectives rather than maximized indiscriminately. A common assumption in simulation-based training is that higher fidelity is always better; however, this is not supported by the evidence. Excessively complex environments can impede early-stage learning by introducing irrelevant variation before foundational competencies are established. This staged approach draws on the scaffolding principle in learning theory, in which task complexity is incrementally increased as competence develops [10]. Applied here, the robot's training environment progresses through structured tiers of environmental and behavioral complexity, with promotion to a higher-fidelity tier contingent on demonstrated performance at the current tier. This prevents the robot from encountering clinical complexity it is not yet equipped to handle. Meta-analyses confirm that structured VR environments improve knowledge acquisition, skill development, and problem-solving in nursing education [11-14], and the same principle applies to robot training: structured progression through complexity produces more robust and generalizable policies than exposure to full complexity from the outset. The 3 fidelity tiers are defined as follows: low fidelity: ward orientation, routing, object localization, and collision avoidance; mid fidelity: dynamic obstacles, workflow variations, staff interactions, and monitoring support; high fidelity: contact-rich assistance, edge-case rehearsal, and timing-sensitive coordination.

Site-Specific Digital Twin Synchronization

In this framework, the digital twin is defined as a continuously synchronized operational model of a specific nursing unit, constrained to the variables that materially affect robotic task performance, rather than a comprehensive, photorealistic, or static replica. It is an operational bridge, not a decorative replica. The distinction matters because health care digital twin implementations vary enormously in scope and ambition, from comprehensive hospital-wide models to narrowly scoped process simulations. A recent National Academies report identifies foundational research gaps in digital twin validation, interoperability, and domain-specific adaptation, cautioning that the term is frequently applied to systems that lack the bidirectional synchronization and predictive capability that define a true twin [15]. Health care–specific reviews note similar definitional variability and largely partial implementations [16-18]. The pragmatic approach adopted here constrains the twin to the minimum set of variables that materially affect robotic task performance in a specific nursing unit: room geometry and doorway clearances, equipment locations and dimensions, circulation pathways and bottlenecks, staffing density patterns, and unit-specific workflow conventions such as supply routes and shift-change traffic patterns. This constraint is deliberate. An overly ambitious twin risks becoming an engineering project in itself, diverting resources from the core research questions. A pragmatically scoped twin, by contrast, serves its intended function—bridging the gap between a generic simulation and the particular physical and operational realities of the deployment site—without requiring comprehensive hospital informatization as a prerequisite.

The level of fidelity a twin requires is determined by what the task depends on. Supportive tasks, such as navigation, transport, retrieval, and environmental monitoring, depend on spatial geometry, circulation patterns, equipment placement, and workflow timing, all of which the pragmatic twin represents; they do not depend on patient-specific anatomy or tissue behavior. Surgical tasks, by contrast, depend on patient-specific anatomical structure, tissue deformation, and submillimeter spatial accuracy, which is why surgical applications require high-fidelity, patient-specific twins. The pragmatic twin is therefore not an oversimplified version of a surgical twin but a different object matched to a different task class: for supportive nursing tasks it is sufficient, whereas for surgery it would be inadequate. Framing twin fidelity as task-determined, rather than as a single scale of ambition, clarifies why the scoping adopted here is appropriate rather than simplistic.

XSL•FO
RenderX

Governed Simulation-to-Real Transfer and Evaluation

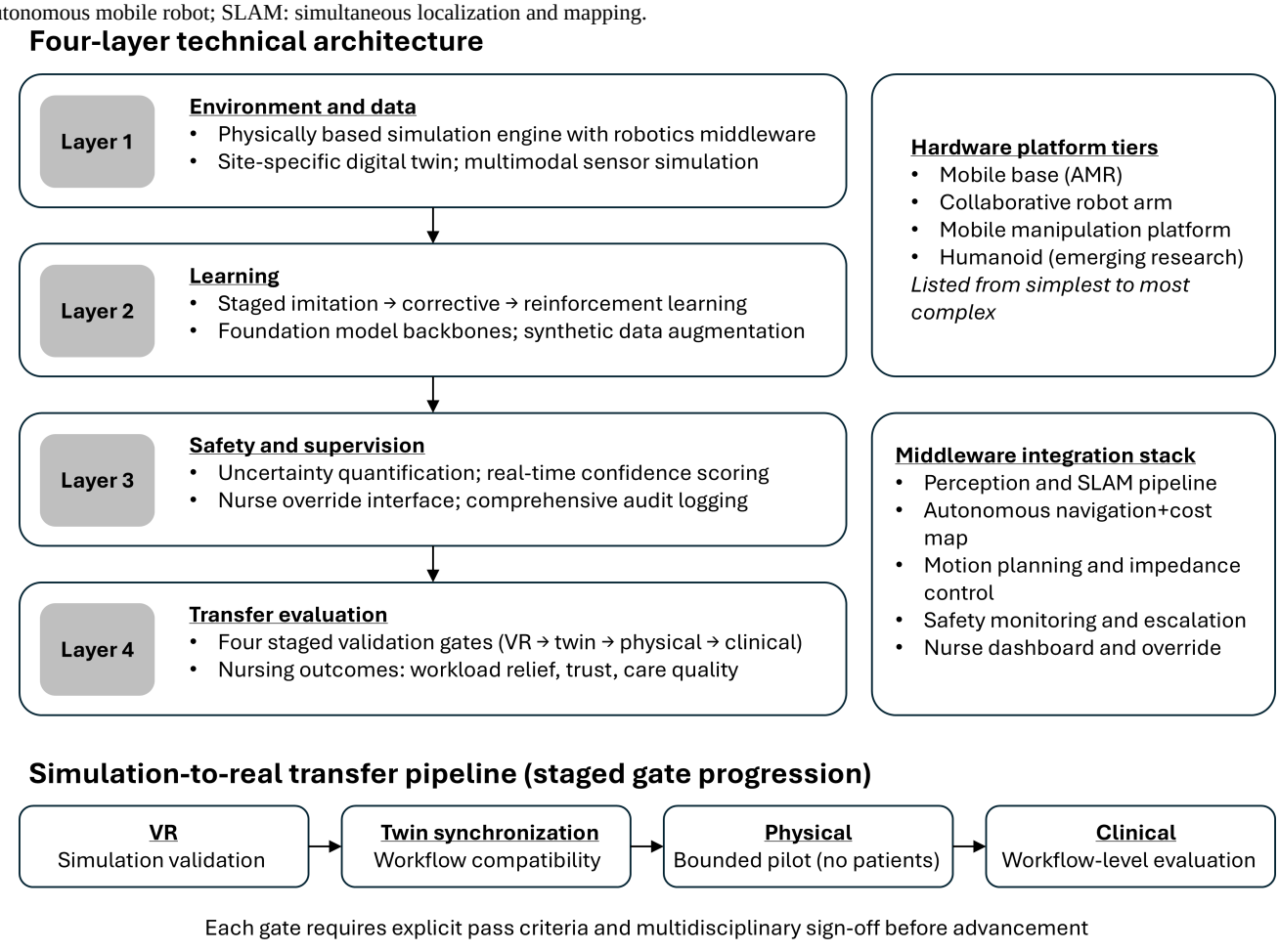
Transfer criteria extend beyond task success to include workflow compatibility, interruption recovery, escalation quality, and noninterference with bedside care. In conventional robotics, simulation-to-real transfer is treated primarily as a technical challenge: closing the gap between simulated and physical dynamics through domain randomization [19], dynamics randomization [20], and system identification. These techniques remain essential and are incorporated into the framework's technical architecture (Multimedia Appendix 1). However, in a nursing context, the transfer gap is not only physical but also social, organizational, and ethical. A robot that navigates a physical ward successfully may still fail if it disrupts nurse workflow, creates alarm fatigue through excessive escalation, or undermines patient trust through inappropriate proximity or timing. Therefore, simulation-to-real transfer is treated here as a governance event rather than a final technical step. Transfer proceeds through staged validation gates, each requiring explicit sign-off before progression. This stage-gate structure ensures that technical performance is necessary but not sufficient for clinical deployment; nursing-relevant outcomes, including measurable workload relief, nurse trust as assessed by validated human-robot interaction scales, and care quality indicators, are the ultimate evaluation standard.

Conceptual Technical Architecture

Four-Layer Technical Architecture

The framework is grounded in a 4-layer technical architecture (Figure 2), with illustrative implementation options. The conceptual reference architecture, including evaluation logic and architectural requirements for each layer, is provided in Multimedia Appendix 1. Throughout this section, the contribution lies in the layered functional structure of each layer rather than in any specific product or method; all named technologies are current illustrative instantiations that satisfy a layer's functional requirements and may be substituted by implementing teams as the tooling landscape evolves.

Figure 2. Four-layer technical architecture and simulation-to-real transfer pipeline for virtual reality (VR)–based physical AI training in nursing. AMR: autonomous mobile robot; SLAM: simultaneous localization and mapping.



Layer 1: Environment and Data

The VR environment requires graphics processing unit–accelerated, physically based simulation with direct interoperability with robotics middleware, specifically Robot Operating System 2 (ROS 2); platforms such as Isaac Sim or Unity with a ROS 2 bridge are illustrative options that meet this requirement. For example, Isaac Sim provides native integration with a physics engine supporting both rigid and deformable body dynamics—useful for simulating clinical equipment interactions—together with built-in synthetic data generation. These are cited as representative current tools rather than as required choices; the functional requirement, not the specific platform, is what the framework specifies. Site-specific digital twins are constructed via light detection and ranging scanning

XSL•FO
RenderX

or photogrammetry of target clinical units, with the resulting 3D environments synchronized with operational workflow data drawn from structured observation and, where available, real-time locating system feeds. Multimodal sensor channels are simulated within the environment to match the sensor configuration of the target physical robot platform. For the noncontact logistical and environmental tasks that constitute the framework's initial scope, the active sensing channels are red, green, blue, and depth cameras, light detection and ranging, and inertial measurement units. Force and haptic channels are additionally represented in simulation to preserve fidelity with contact-capable platforms, including collaborative-arm and emerging humanoid configurations, which increasingly incorporate fingertip tactile and joint-level force-torque sensing as standard equipment. Because simulation-to-real transfer requires the simulated sensor suite to mirror the deployed platform, these contact channels are not engaged as policy inputs for noncontact tasks; they become active only if and when the framework is extended to contact-rich assistance under the staged validation gates (layer 4). Expert demonstrations and corrective feedback are captured as structured task-state logs, including joint positions, end-effector poses, gaze direction, verbal annotations, and contextual state labels.

Layer 2: Learning

The learning layer operates in 3 sequential phases, each defined by its functional role rather than by a specific algorithm. First, foundation policy formation (eg, behavioral cloning from expert nurse demonstrations) provides a policy that captures the basic structure of each task. Second, interactive correction, using methods such as dataset aggregation (DAgger) or intervention-weighted regression, refines this policy by incorporating nurse feedback on the robot's own attempted trajectories, weighting corrections by intervention timing and confidence to address the distribution mismatch problem inherent in offline imitation learning [21]. Third, policy optimization (eg, reinforcement-learning fine-tuning with algorithms such as proximal policy optimization or soft actor-critic under nursing-specific reward shaping) optimizes the policy for task completion, safety margins, workflow timing, escalation quality, and trajectory smoothness. Pretrained robot foundation-model backbones, such as OpenVLA [22] or Octo [23], may be leveraged to reduce data requirements and improve generalization across task families. The named methods are representative instances of each functional role and may be substituted as the field advances. Synthetic data augmentation using domain randomization and world foundation models expands coverage of rare but safety-critical clinical edge cases, such as unexpected patient falls, equipment malfunctions, or crowded corridor scenarios, without requiring real-world exposure to these events. Importantly, reinforcement-learning fine-tuning is not proposed as a means of solving contact-rich reality-gap problems in simulation; rather, it is applied incrementally only after lower-risk policies have passed the staged validation gates and only within the constraints of supervised real-world evaluation.

Layer 3: Safety and Supervision

The safety layer crosscuts all other layers and operates at multiple timescales. At the real-time level, uncertainty quantification via Monte Carlo dropout or conformal prediction generates continuous confidence scores for the robot's actions and perceptions. Hard stop rules trigger immediately when force thresholds are exceeded, when confidence drops below task-specific thresholds (τ), or when the robot detects an unrecognized situation outside its training distribution. A nurse override API, implemented as a ROS 2 action server, allows real-time human intervention with a target response latency of less than 500 milliseconds. At the session level, all decisions, confidence scores, escalation events, and nurse override instances are logged in a structured format for post hoc audit. These logs serve dual purposes: they provide the audit trail necessary for institutional safety review, and they generate the training signal for policy refinement in subsequent learning iterations. The safety layer embodies the principle that the robot should fail conservatively: when uncertain, it should stop and escalate rather than proceed with a potentially harmful action.

Layer 4: Transfer Evaluation

Transfer proceeds through 4 stage gates, each with explicit pass or fail criteria and required sign-off before advancement. Gate 1 validates VR against task accuracy, collision rate, and escalation quality benchmarks. Gate 2 validates in the digital-twin–synchronized environment for workflow compatibility, route efficiency, and timing alignment with shift-level task patterns. Gate 3 involves bounded physical pilot deployment in a controlled clinical area without patients, evaluating physical task performance, nurse override latency, and usability. Gate 4 extends to workflow-level evaluation in a real clinical unit under supervision, integrating nursing-relevant outcomes, including measurable workload relief (NASA–Task Load Index [NASA-TLX]), nurse trust (validated human-robot interaction trust scales), and care quality indicators. No gate may be bypassed, and regression to an earlier gate is required if performance degrades at any stage. A formal multidisciplinary review, including nursing leadership, biomedical engineering, and safety officers, is required before advancing from Gate 3 to Gate 4. Within each gate, advancement is conditioned on predefined quantitative criteria of 2 kinds: absolute safety constraints that must not be violated, such as avoidance of collisions and unsafe proximity, and relative performance criteria benchmarked against acceptable task execution, such as task-completion accuracy and escalation appropriateness. Because the appropriate numeric thresholds depend on the task family, clinical unit, and risk profile, they are established and calibrated empirically during validation rather than fixed as universal constants at the conceptual stage; Multimedia Appendix 1 specifies the gate environments, evaluation-criteria categories, and nursing-relevant indicators against which these thresholds are set.

The conceptual reference architecture, including illustrative platform comparisons, sensor requirements, middleware integration logic, and a staged validation structure, is provided in Multimedia Appendix 1.

XSL•FO
RenderX

Ethical, Professional, and Practice Implications

Governance should be built into the training pipeline from the outset, not retrofitted after technical development is complete. This means nurses should define the task library, escalation logic, and performance criteria, rather than only provide end-stage user feedback on systems designed without their input. The ethical stakes are significant: a robot operating in a clinical environment has the potential to affect patient safety, nurse workload, and the therapeutic relationship between nurse and patient. The framework addresses these stakes through several mechanisms. First, it preserves nurse authority over all ethically consequential decisions by design; the role-boundary map explicitly prohibits robotic action in domains requiring clinical judgment. Second, robotic support is bounded to clearly delegable task families where the consequences of error are manageable and reversible. Third, VR carries a governance function alongside its training function: it provides a predeployment space where action boundaries, error modes, and workflow consequences can be examined without risk to patients [5,6,24,25]. Fourth, the staged transfer gates embed ethical review into the development pipeline, ensuring that the decision to deploy a robot in a clinical unit is made by a multidisciplinary team, rather than by the technical development team alone. These design choices reflect a broader commitment to the principle that AI in nursing should augment professional practice rather than displace professional judgment [5,6].

Beyond these design commitments, 3 areas require explicit treatment because they condition whether such a system can be responsibly developed and deployed. The first concerns regulatory classification and legal responsibility. A learned policy that supports nursing tasks may, depending on its function and claims, fall within medical-device-software regulation, such as the US Food and Drug Administration framework for Software as a Medical Device [26,27], and, in other jurisdictions, within risk-based regimes such as the European Union Artificial Intelligence Act [28], both of which impose validation, documentation, and postdeployment monitoring obligations that a nursing robotics program must anticipate from the outset. Closely related is the allocation of liability when an autonomous system contributes to an adverse event: responsibility may be distributed among the nurse who supervises the system, the institution that deploys it, and the developer or manufacturer that builds it, and this distribution is not yet settled in law or policy. The framework does not resolve this question, but its retained human accountability, explicit role boundaries, and multidisciplinary sign-off at each transfer gate are designed to keep responsibility traceable rather than diffuse, providing an auditable record of who authorized what at each stage. The second and third areas concern data governance and patient consent. Nurse demonstrations, corrective feedback, and deployment logs generate behavioral and environmental data, and in a clinical setting this data may incidentally capture patient-related information; such data must be handled under institutional governance consistent with health-privacy law, including, in the United States, HIPAA-consistent deidentification and the exclusion of patient-identifiable information from training corpora. A distinct and less frequently acknowledged question is the ownership of and consent for the nurses' tacit expertise that the framework deliberately elicits and encodes, because the nurse-in-the-loop paradigm treats expert judgment as a primary training signal, the professionals whose knowledge is captured have a legitimate interest in how that knowledge is used, attributed, and governed. Finally, patients have an interest in robotic support operating in their care environment even when tasks are noncontact; institution-level notice-and-consent processes should make patients aware of robotic support and provide a means to decline it, preserving patient autonomy as a first-class consideration rather than an afterthought. Treating these regulatory, data-governance, and consent questions as design requirements, rather than as compliance steps addressed after development, is consistent with the framework's commitment to building governance into the pipeline from the outset.

Discussion

This paper repositions VR from a visualization tool to a translational training infrastructure connecting nursing expertise with robotic execution. The claim is not that nursing requires an entirely separate technological universe from big-tech physical AI; rather, nursing needs an additional nurse-centered orchestration layer that specifies workflow fit, role clarity, safety governance, and local adaptation as nonnegotiable adoption criteria. This complements the call by Shaw and Chen [3] for nursing-specific engagement with physical AI by providing a concrete architectural pathway from concept to implementation.

The 4-layer technical architecture grounds the framework in concrete implementation considerations. By specifying graphics processing unit–accelerated simulation, robotics middleware integration, staged imitation-to-reinforcement learning, foundation model backbones, uncertainty quantification, and staged transfer gates, the framework provides an implementation-aware roadmap that connects conceptual nursing robotics discourse with concrete implementation pathways. Each architectural layer addresses a specific translational challenge: layer 1 bridges the gap between clinical environments and simulation; layer 2 structures the learning process around nurse expertise rather than autonomous exploration; layer 3 embeds safety governance into real-time operation; and layer 4 enforces validation against nursing-relevant outcomes before any clinical contact occurs. This layered structure is intended to make the framework's plausibility conceptually arguable and incrementally verifiable through independently testable stages, rather than being assumed.

The strategic contribution is clarity of scope and sequence. Rather than proposing a generalized nursing humanoid—an aspiration that outpaces both current technology and clinical readiness—the framework defines a tractable developmental pathway: build a prototype VR environment for bounded supportive nursing tasks, test nurse-in-the-loop data collection and correction workflows, synchronize with one real clinical unit, and evaluate transfer against nursing-relevant outcomes. This incremental approach is more clinically defensible because it allows each stage to be independently validated and because

XSL•FO
RenderX

it confines early-stage risk to simulation rather than exposing patients or staff to insufficiently tested systems. It also aligns with how nursing practice itself manages complexity: through structured protocols, graduated responsibility, and continuous reassessment.

Several features distinguish this framework from existing approaches to health care robotics. Most current health care robots are developed outside nursing and offered to nurses as finished products; this framework places nurses at the center of the development process from task selection through policy training to transfer evaluation. Most simulation-to-real pipelines in robotics treat transfer as a technical challenge to be solved through domain randomization; this framework treats transfer as a governance event requiring nursing-relevant validation. Further, most health care digital twin implementations aspire to comprehensive hospital models; this framework adopts a pragmatically scoped twin constrained to variables that materially affect robotic task performance. These distinctions are not arbitrary; they follow from the recognition that nursing's requirements are not merely a subset of general robotics requirements but include qualitatively different demands around workflow integration, role clarity, and preserved professional authority.

This contrast is clearest when supportive autonomy is set against surgical autonomy. Surgical autonomy centers on contact-rich, submillimeter, and often irreversible manipulation of patient anatomy, which requires patient-specific high-fidelity modeling and fine force-level control. Supportive autonomy, as framed here, centers instead on logistical and environmental tasks whose errors are comparatively bounded and recoverable; where direct physical assistance is eventually introduced, it enters only under staged, nurse-governed validation gates rather than as the operating premise. This difference in kind is precisely why the framework's architectural choices fit supportive autonomy and would be insufficient for surgery: the pragmatically scoped digital twin, the workflow-centered and role-centered validation, and the gated transfer logic are calibrated to bounded support rather than to anatomical precision. What carries across domains is not the technical stack but the governance: expert-in-the-loop policy formation, staged validation, and retained human accountability are not specific to nursing and could inform surgical contexts, even though the fidelity and precision they must achieve differ fundamentally.

The framework also raises questions that warrant further investigation. How many expert nurse demonstrations are sufficient to train a robust foundation policy for a given task family? What is the minimum viable scope of a digital twin that preserves transfer validity? How should escalation thresholds be calibrated to balance safety against workflow disruption from excessive escalation? And how do nurses' trust and acceptance of robotic support evolve as they participate in the training process vs receiving a pretrained system? These are empirical questions that the framework is designed to enable rather than preemptively answer.

Limitations

Five limitations warrant explicit acknowledgment. First, this paper is conceptual and does not empirically test the framework; the proposed architecture is a structured hypothesis for future empirical validation. Second, the simulation-to-real gap remains incompletely solved. Sensor noise and clinical unpredictability will create transfer challenges that simulation cannot fully anticipate, and physical contact dynamics in particular—deformable-tissue interaction and subtle force modulation—fall outside the framework's initial noncontact scope and would require dedicated modeling and validation before any contact task could be approached. Accordingly, the transition from behavioral cloning to reinforcement learning is positioned not as a simulation step but as fine-tuning in the validated real environment: behavioral cloning establishes a safe foundation policy from expert demonstrations, while reinforcement-learning refinement is applied incrementally under the staged transfer gates rather than to anticipate contact dynamics in simulation. Multimedia Appendix 1 characterizes the principal dimensions of this transfer gap. Third, the digital twin literature remains technically immature in health care; implementers must calibrate expectations and adopt the pragmatic constraint recommended here. Fourth, hardware platform selection (autonomous mobile robot, cobot arm, and humanoid) will substantially shape which components are feasible and how the twin must be constructed. Fifth, although the framework places nurses at the center of robotic policy formation as a matter of design, this paper was itself developed through literature synthesis rather than through direct co-design with nurses, patients, or administrators. These are distinct levels of nurse centrality: positioning nurses as the authority who defines robotic scope is a property the framework specifies, whereas codeveloping that specification with practicing nurses is an empirical step the framework explicitly requires rather than one this conceptual paper performs. Accordingly, participatory co-design is not omitted but deferred as a core element of the next stage: future empirical work should begin with a participatory design phase in which nurses from the target clinical unit codevelop the task library, role-boundary map, and escalation logic before any VR environment is constructed. This sequencing—defining what nurses should govern at the conceptual stage and operationalizing that governance through participation next—resolves the apparent tension that the absence of co-design at this stage might otherwise suggest.

Future research should pursue several priorities. First, prototype VR environments should be constructed for specific bounded nursing task families, beginning with logistical tasks (supply transport and equipment retrieval) before advancing to more complex support tasks involving dynamic human interaction. Second, digital twin construction feasibility should be examined in real clinical units, including assessment of the time, cost, and institutional cooperation required to achieve adequate twin fidelity. Third, empirical studies should evaluate technical performance alongside nurse trust, workflow compatibility, and escalation quality, using mixed methods designs that combine quantitative performance metrics with qualitative analysis of nurses' experiences. Fourth, longitudinal studies should track

XSL•FO
RenderX

how nurse trust and acceptance evolve over repeated interactions with the training and deployment pipeline. Finally, cross-site generalization should be investigated: how much of the training pipeline transfers when the framework is applied to a different clinical unit with different geometry, workflow patterns, and nursing culture?

Conclusions

Physical AI in nursing should be framed as supportive robotics that reduce repetitive and physically burdensome work while preserving nurses' authority over judgment, prioritization, and relational care. This framing is not a concession to technological limitations; it is a principled design commitment grounded in what nursing is and what robots can meaningfully contribute to nursing practice in the near to medium term.

The nurse-in-the-loop, site-specific VR framework proposed in this paper addresses the current gap between conceptual promise and clinical feasibility through several integrated mechanisms. The task library and role-boundary map ensure that the robotic scope is defined by clinical judgment rather than technical capability. Nurse-in-the-loop VR training ensures that the robot's learned policies reflect expert nursing knowledge, contextual awareness, and safety priorities. Progressive simulation fidelity structures the learning process to build competence incrementally, avoiding premature exposure to clinical complexity. Site-specific digital twin synchronization bridges the gap between generic simulation and the particular physical and organizational realities of the deployment environment. Governed simulation-to-real transfer ensures that advancement from simulation to physical deployment is conditioned on explicit nursing-relevant validation criteria, with human accountability retained at every stage.

Taken together, these elements constitute a translational infrastructure rather than a single technological intervention. VR is not merely a visual training aid in this framework; it is the environment where nursing expertise is elicited, structured, annotated, stress-tested, and safely transferred to physical systems. The digital twin is not a visualization tool; it is the mechanism through which generic robot capabilities are adapted to local clinical realities. Further, the nurse is not an end user to be consulted after development; the nurse is the domain expert whose knowledge, judgment, and workflow requirements drive the entire pipeline.

This framework contributes to nursing informatics by proposing a concrete mechanism through which nursing knowledge can be formally structured and transmitted to physical AI systems, extending the discipline's engagement with AI beyond decision support and documentation into the domain of embodied, physically acting systems. It contributes to robotics by articulating a domain-specific orchestration layer that connects general-purpose physical AI capabilities with the particular requirements of a complex, safety-critical, and human-centered work environment. It also contributes to patient safety by ensuring that the path from simulation to clinical deployment is governed by explicit criteria, staged validation, and retained human authority.

The path forward requires sustained collaboration between nursing scientists, robotics engineers, clinical administrators, and the nurses who ultimately work alongside these systems. The framework proposed here is intended as a structured starting point for that collaboration—not a finished system, but a principled architecture within which empirical questions can be posed, tested, and iteratively refined.

Acknowledgments

The author used Claude (Anthropic) to assist with language editing and improve the clarity of the manuscript text and ChatGPT (OpenAI) to generate the table of contents image. The author reviewed and edited all AI-assisted output and takes full responsibility for the accuracy, originality, and integrity of the content, including all references and citations.

Funding

The author declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Technical reference architecture for the proposed framework, including virtual reality environment construction, digital twin synchronization, the learning pipeline, staged simulation-to-real validation, and hardware and robotics-middleware integration.
[DOCX File, 372 KB - nursing_v9i1e96848_app1.docx]

References

1. Babalola GT, Gaston JM, Trombetta J, Tulk Jesso S. A systematic review of collaborative robots for nurses: where are we now, and where is the evidence? Front Robot AI 2024;11:1398140. [doi: 10.3389/frobt.2024.1398140] [Medline: 38899066]
2. Adeyemo A, Coffey A, Kingston L. Utilisation of robots in nursing practice: an umbrella review. BMC Nurs 2025 Mar 4;24(1):247. [doi: 10.1186/s12912-025-02842-2] [Medline: 40038679]

XSL•FO
RenderX

3. Shaw RJ, Chen B. Physical artificial intelligence in nursing: robotics. Nurs Outlook 2025;73(5):102495. [doi: 10.1016/j.outlook.2025.102495] [Medline: 40675084]
4. Joo JY, Liu MF, Ho MH. Nurses' perceptions of artificial intelligence adoption in healthcare: a qualitative systematic review. Nurse Educ Pract 2025 Oct;88:104542. [doi: 10.1016/j.nepr.2025.104542] [Medline: 40925173]
5. El Arab RA, Al Moosa OA, Abuadas FH, Somerville J. The role of AI in nursing education and practice: umbrella review. J Med Internet Res 2025 Apr 4;27:e69881. [doi: 10.2196/69881] [Medline: 40072926]
6. Al Khatib I, Ndiaye M. Examining the role of AI in changing the role of nurses in patient care: systematic review. JMIR Nurs 2025 Feb 19;8:e63335. [doi: 10.2196/63335] [Medline: 39970436]
7. Trainum K, Liu J, Hauser E, Xie B. Nursing staff's perspectives of care robots for assisted living facilities: systematic literature review. JMIR Aging 2024 Sep 16;7:e58629. [doi: 10.2196/58629] [Medline: 39283666]
8. Celemin C, Pérez-Dattari R, Chisari E, et al. Interactive imitation learning in robotics: a survey. Found Trends Robot 2022 Nov 22;10(1-2):1-197. [doi: 10.1561/2300000072]
9. Slade P, Atkeson C, Donelan JM, et al. On human-in-the-loop optimization of human-robot interaction. Nature 2024 Sep;633(8031):779-788. [doi: 10.1038/s41586-024-07697-2] [Medline: 39322732]
10. Wood D, Bruner JS, Ross G. The role of tutoring in problem solving. J Child Psychol Psychiatry 1976 Apr;17(2):89-100. [doi: 10.1111/j.1469-7610.1976.tb00381.x] [Medline: 932126]
11. Kyaw BM, Saxena N, Posadzki P, et al. Virtual reality for health professions education: systematic review and meta-analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959. [doi: 10.2196/12959] [Medline: 30668519]
12. Liu K, Zhang W, Li W, Wang T, Zheng Y. Effectiveness of virtual reality in nursing education: a systematic review and meta-analysis. BMC Med Educ 2023 Sep 28;23(1):710. [doi: 10.1186/s12909-023-04662-x] [Medline: 37770884]
13. Cho MK, Kim MY. Enhancing nursing competency through virtual reality simulation among nursing students: a systematic review and meta-analysis. Front Med (Lausanne) 2024;11:1351300. [doi: 10.3389/fmed.2024.1351300] [Medline: 38774395]
14. Hsieh JY, Lin PC, Sun WN, Lin TR, Kuo CC, Hsu HT. Effectiveness of immersive virtual reality in nursing education for nursing students and nursing staffs: a systematic review and meta-analysis. Nurse Educ Today 2025 Aug;151:106725. [doi: 10.1016/j.nedt.2025.106725] [Medline: 40215711]
15. National Academies of Sciences, Engineering, and Medicine. Foundational research gaps and future directions for digital twins. : National Academies Press; 2024 URL: https://www.ncbi.nlm.nih.gov/books/NBK605515/pdf/Bookshelf_NBK605515.pdf [accessed 2026-09-03]
16. Drummond D, Gonsard A. Definitions and characteristics of patient digital twins being developed for clinical use: scoping review. J Med Internet Res 2024 Nov 13;26:e58504. [doi: 10.2196/58504] [Medline: 39536311]
17. Riahi V, Diouf I, Khanna S, Boyle J, Hassanzadeh H. Digital twins for clinical and operational decision-making: scoping review. J Med Internet Res 2025 Jan 8;27:e55015. [doi: 10.2196/55015] [Medline: 39778199]
18. Shen MD, Chen SB, Ding XD. The effectiveness of digital twins in promoting precision health across the entire population: a systematic review. NPJ Digit Med 2024 Jun 3;7(1):145. [doi: 10.1038/s41746-024-01146-0] [Medline: 38831093]
19. Tobin J, Fong R, Ray A, Schneider J, Zaremba W, Abbeel P. Domain randomization for transferring deep neural networks from simulation to the real world. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2017:23-30. [doi: 10.1109/IROS.2017.8202133]
20. Peng XB, Andrychowicz M, Zaremba W, Abbeel P. Sim-to-real transfer of robotic control with dynamics randomization. 2018 IEEE International Conference on Robotics and Automation (ICRA) 2018:3803-3810. [doi: 10.1109/ICRA.2018.8460528]
21. Ross S, Gordon GJ, Bagnell JA. A reduction of imitation learning and structured prediction to no-regret online learning. Proc 14th Int Conf Artif Intell Stat 2011:627-635 [FREE Full text]
22. Kim MJ, Pertsch K, Karamcheti S, et al. OpenVLA: an open-source vision-language-action model. Proc 8th Conf Robot Learn 2025;270:2679-2713 [FREE Full text]
23. Octo Model Team, Ghosh D, Walke H, et al. Octo: an open-source generalist robot policy. Presented at: Robotics: Science and Systems 2024; Jul 15-19, 2024 URL: https://www.roboticsproceedings.org/rss20/p090.pdf [accessed 2026-09-08]
24. Ball Dunlap PA, Michalowski M. Advancing AI data ethics in nursing: future directions for nursing practice, research, and education. JMIR Nurs 2024 Oct 25;7:e62678. [doi: 10.2196/62678] [Medline: 39453630]
25. Ventura-Silva J, Martins MM, Trindade LDL, et al. Artificial intelligence in the organization of nursing care: a scoping review. Nurs Rep 2024 Oct 2;14(4):2733-2745. [doi: 10.3390/nursrep14040202] [Medline: 39449439]
26. Software as a Medical Device (SaMD). US Food and Drug Administration. URL: https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd [accessed 2026-09-08]
27. Marketing submission recommendations for a predetermined change control plan for artificial intelligence-enabled device software functions: guidance for industry and food and drug administration staff. US Food and Drug Administration. 2024 Dec. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence [accessed 2026-09-08]
28. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU)

XSL•FO
RenderX

2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance). : European Union; 2024 URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng [accessed 2026-09-10]

Abbreviations

NASA-TLX: NASA–Task Load Index
ROS 2: Robot Operating System 2
VR: virtual reality

Edited by E Borycki; submitted 01.Apr.2026; peer-reviewed by A Faria, X Xiao; revised version received 24.Jun.2026; accepted 28.Jul.2026; published 05.Oct.2026.

Please cite as:
Kim S
Virtual Reality–Enabled Physical AI Training for Supportive Nursing Robotics: Nurse-in-the-Loop, Site-Specific Conceptual Framework
JMIR Nursing 2026;9:e96848
URL: https://nursing.jmir.org/2026/1/e96848
doi:10.2196/96848

© Seungman Kim. Originally published in JMIR Nursing (https://nursing.jmir.org), 5.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

From Concept to Practice: Lessons From the Balanced Nursing Teams Decision-Support System

Senne Vleminckx[1], MSc; Peter Van Bogaert[1], MSc, Prof Dr; Wim De Keyser[2], MSc, PhD; Filip Haegdorens[1], MSc, Prof Dr

[1]Workforce management, Health systems, and Outcome Research in Care – Centre for Research and Innovation in Care, Faculty of Medicine and Health Sciences, University of Antwerp, Universiteitsplein 1, Wilrijk, Antwerp, Belgium

[2]Applied Computer Science, Sustainable Industry Research Team, Karel de Grote Hogeschool, Antwerp, Flanders, Belgium

Corresponding Author:
Senne Vleminckx, MSc
Workforce management, Health systems, and Outcome Research in Care – Centre for Research and Innovation in Care, Faculty of Medicine and Health Sciences, University of Antwerp, Universiteitsplein 1, Wilrijk, Antwerp, Belgium

Abstract

The global nursing workforce crisis demands a shift from reactive staffing to strategic workforce optimization through data-driven decision-support systems. This viewpoint paper reflects on the development and attempted implementation of the balanced nursing teams system, a decision-support tool integrating approximately 250 data points—of which roughly 150 are extracted from existing organizational systems (human resources, scheduling, electronic health records, quality registries) through flexible import mechanisms, and the remainder collected through a built-in 360-degree staff survey with automated analysis—across 10 domains to evaluate nursing team balance between capacity, performance, and outcomes. Following crowdfunding by 18 Belgian health care organizations, balanced nursing teams were implemented across 8 diverse settings (home health care, general hospitals, academic centers) between 2019 and 2023. Using the Human-Organization-Technology fit framework, we analyze why evidence-informed, organization-endorsed digital innovations struggle to achieve adoption. Our analysis reveals 3 interdependent barrier categories: technological fragmentation (vendor lock-in, legacy systems, prohibitive integration costs), organizational siloing (Chief Nursing Officers [CNOs] lacking budgetary authority, nursing framed as peripheral to strategic priorities), and managerial hesitance (fear of punitive data use, cognitive overload from staffing crises). These barriers were worsened by the substantial data-integration burden that the system's breadth imposed on organizations with limited digital maturity. Critically, only one site (ie, a nurse-led home health care organization where leadership held both strategic authority and resource control) achieved sustained implementation. This contrast demonstrates that workforce optimization through data depends not on software maturity alone, but on achieving simultaneous fit across human, organizational, and technological domains. We argue that the persistent marginalization of nursing leadership within hospital governance structures represents the fundamental barrier to digital transformation in nursing workforce management. The urgency paradox is striking: while nursing represents health care organizations' highest operational cost and most direct patient interface, workforce optimization tools are consistently deprioritized in favor of regulatory compliance systems and billing infrastructure. Bridging this gap requires systemic investment in nursing leadership authority, data interoperability standards, and recognition that data-driven workforce decisions are strategic imperatives rather than operational luxuries.

(JMIR Nursing 2026;9:e92417) doi:10.2196/92417

KEYWORDS

nursing workforce management; clinical decision-support systems; implementation science; sociotechnical barriers; nursing leadership

Introduction

Overview

The global nursing workforce crisis extends beyond simple shortages. Persistent mismatches between staffing levels, skill mix, and patient demand have been linked to missed care, adverse outcomes, and failure to rescue, ultimately compromising both patient and staff outcomes [1-4]. While policymakers have focused primarily on increasing the number of nurses entering the workforce, evidence indicates that expanding supply alone does not resolve the underlying staffing challenges [5,6]. The retention of experienced nurses depends on the work environment quality and team performance [7], yet high turnover continues to impose substantial financial and operational burdens on health care organizations [8].

Current staffing models, such as mandatory nurse-to-patient ratios, fail to capture the complexity of nursing team dynamics, skill mix optimization, and broader contextual factors that

XSL•FO
RenderX

determine whether care teams can effectively meet patient demands [9]. Addressing this challenge requires a shift from ratio-based workforce planning towards data-informed team design that integrates human, organizational, and contextual dimensions of nursing care [9].

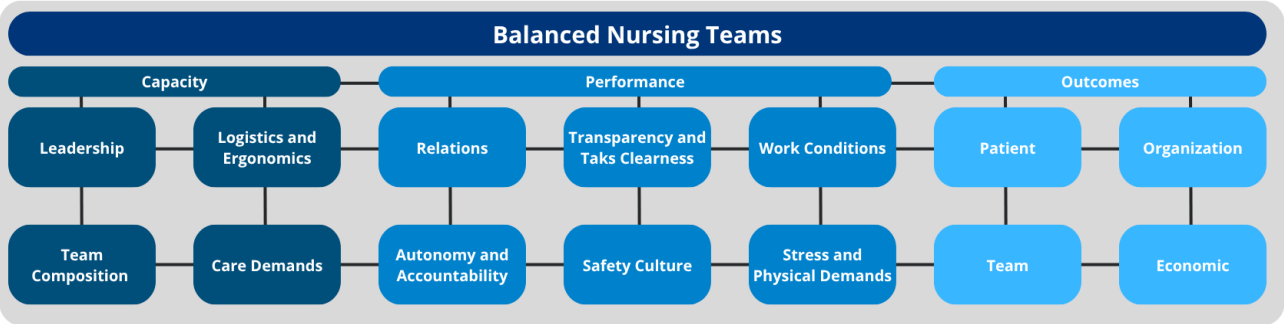
This viewpoint paper offers a critical reflection on the sociotechnical challenges of implementing digital workforce innovation in nursing. Drawing from our experience developing and testing the balanced nursing teams (BNuT) decision-support system across 8 Belgian health care organizations, we examine why evidence-informed, practitioner-endorsed digital tools struggle to achieve real-world adoption. Rather than presenting a traditional implementation study, we provide an analytical perspective on the systemic barriers that constrain digital transformation in nursing workforce management, offering lessons for researchers, health care organizations, and policymakers pursuing similar innovations.

A Nursing Teams Composition and Staffing Decision-Support System: BNuT

In response to these challenges, the BNuT initiative was launched in 2018 through a collaboration between a member organization of Belgian health care organizations and 2 research groups. The project originated from a clear practical demand: develop an instrument that enables evidence-based, proactive, and data-driven decisions regarding nursing team composition. Following proof-of-concept development, 18 health care institutions collectively crowdfunded the project, reflecting broad recognition of this critical gap in nursing workforce management.

BNuT was conceived as a computerized decision-support instrument designed to optimize team composition and team functioning in ways that existing tools cannot. While most workforce software estimates patient demand or calculates staffing levels based on mandatory ratios [10], these systems rely heavily on manual governance and contextual adjustment. In contrast, BNuT integrates data across multiple domains—including relational dynamics, autonomy, care demand, leadership, logistics, and safety—into a unified team-level model of workforce balance (Figure 1).

Figure 1. Conceptual model of the balanced nursing teams (BNuT) system. The figure depicts the 10 subcomponents supporting the BNuT decision-support system, organized into 2 overarching categories: capacity domains (team composition, care demands, staffing, and skill mix) and performance domains (leadership, logistics, safety culture, workload distribution, team dynamics, and professional development). These domains collectively inform the assessment of team balance, ranging from negative imbalance through optimal balance to positive imbalance.



The system estimates team balance on a continuum ranging from negative imbalance (eg, excessive workload and high risk of burnout) through optimal balance to positive imbalance (eg, over-capacitated teams and high risk of boreout). This continuum builds on earlier conceptualization of a balanced care team as one in which capacity, including care demands and team composition (4 domains) and operational performance processes (6 domains), is strategically aligned to optimize outcomes for both staff and patients, as well as organizational and economic outcomes [9]. Sustained imbalance between the 10 domains undermines both care quality and workforce sustainability. BNuT aims to operationalize this principle by assessing the alignment between nursing workforce organization and patient care needs, translating complex multidimensional data into actionable, evidence-based visual indicators for team and managerial decision-making.

The development of BNuT followed a staged approach designed to progressively establish validity and relevance. The underlying model was empirically grounded in a systematic scoping review that identified 35 factors across 9 domains of balanced team functioning [9]. The indicator set was subsequently refined through structured codevelopment with nursing managers from participating institutions, who assessed each variable's clinical importance, data availability, and interpretive complexity. The application itself was co-designed with end users through user-story workshops that shaped functionality, interface design, and workflow integration. This iterative process established content validity, ensured practical relevance, and confirmed usability. However, criterion validation against patient and staff outcomes, which was originally planned via the operational deployment in participating hospitals, could not be completed. This validation gap was caused by the sociotechnical barriers described in this paper, as they prevented the successful implementation necessary for outcome evaluation. BNuT should therefore be understood as a theoretically and empirically informed decision-support prototype. Its outcome validation remains contingent on achieving the organizational conditions for sustained use.

Aim and Analytical Framework

This paper provides a reflective analysis of the sociotechnical challenges encountered during the development and testing of BNuT. We move beyond technical specifications to examine

XSL•FO
RenderX

why digital innovations supported by both evidence and organizational demand struggle to achieve real-world adoption. Our analysis uses the Human-Organization-Technology fit (HOT-fit) framework [11], a model specifically developed for evaluating health information systems. This framework maps the 3 principal barrier categories observed in our work and highlights how interactions among human, organizational, and technological factors jointly shape implementation outcomes. Each dimension is examined systematically, followed by an integrative reflection demonstrating how compounding barriers can obstruct implementation even when individual system components function as designed. This viewpoint does not present outcome validation of the BNuT system; rather, it offers a structured reflective analysis of implementation dynamics, informed by 8 years of engagement with participating organizations and stakeholders across the Belgian health care field.

The insights presented in this viewpoint are grounded in 8 years of sustained engagement with the BNuT project (2018 - 2026), encompassing the full cycle from conceptualization and codevelopment to attempted implementation across 18 health care institutions. During this period, the author team engaged in extensive, iterative dialogues with stakeholders across all levels of the health care system (eg, frontline nurses, nursing managers, CNOs, data managers, data protection officers, IT departments, and external development partners) through codevelopment workshops, pilot testing sessions, steering committee meetings, and formal presentations to prospective adopters and investors. Each of these interactions generated feedback that progressively shaped the team's understanding of the barriers and enablers discussed in this paper. The reflections synthesized here emerged iteratively from this long-term, multi-stakeholder engagement. This approach prioritizes ecological breadth and depth of immersion over the rigor of structured qualitative research. The HOT-fit framework [11] was applied retrospectively to organize and analyze these accumulated insights, providing an analytical foundation.

The BNuT System and Implementation Context

BNuT Concept and Architecture

Health care organizations face a fundamental data paradox in nursing workforce management. While data seems abundant, the information remains siloed across systems, leaving leaders without the integrated insight required for strategic workforce decisions. Existing information systems were primarily designed for billing and regulatory compliance rather than operational or strategic decision-making. As a result, critical data concerning team dynamics, workload distribution, competency mix, and patient demand remain fragmented across incompatible platforms, such as electronic health records (EHRs), scheduling tools, human resources (HR) databases, and quality registries. This fragmentation and the lack of comprehensive data warehouses or business intelligence tools leave nurse managers operating largely reactively, addressing daily staffing shortages without insight into long-term workforce patterns or their impact on safety, quality, and retention. Moreover, staff lack a data-informed understanding of their team capacity, performance, and outcomes, and must rely largely on perceptions. This data paradox has become increasingly consequential as contemporary health care systems face mounting pressures: cost containment imperatives, heightened quality and safety demands, and growing clinical hyperspecialization that requires more sophisticated approaches to team composition and skill-mix optimization.

BNuT aims to address these challenges through 2 interconnected objectives that shift workforce management from reactive to strategic. The first involves comprehensive data collection and processing to enable both cross-sectional analyses (capturing current team states) and longitudinal tracking (monitoring team evolution). This supports standardized decision-making capable of distinguishing between expected variance, acceptable fluctuations, and disruptive imbalances. The second objective focuses on integrating multiple data streams into meaningful composite indicators at the clinical microsystem level, encompassing workload intensity, competency distribution, leadership effectiveness, safety climate, and related domains to estimate the balance between team capacity, performance and care demands, and outcomes.

To achieve this, the system combines a 360-degree evidence-based survey, capturing staff perceptions on statements within each capacity and performance domain, and data drawn from existing organizational systems such as scheduling software, EHRs, and quality metrics. All data are synthesized using established evidence-based standards in nursing workforce research, ensuring that the resulting indicators reflect empirically supported relationships between the domains. Crucially, the unit of analysis is the clinical microsystem, acknowledging that patient and team outcomes emerge from collective team processes rather than individual performance metrics. The system supports dynamic management through continuous monitoring and benchmarking functions, allowing leaders to compare performance across teams, identify high-performing configurations, and support continuous improvement. Team involvement in discussing these data is paramount to support teams and leaders in understanding team strengths and weaknesses, to identify improvement priorities, and to assess the extent of readiness for innovation [12]. BNuT thus provides managers and leadership with both a helicopter view for strategic oversight and the ability to zoom in on specific teams or metrics when deeper investigation is required, thereby supporting decision-making at operational, tactical, and strategic levels.

Implementation Setting and Challenges

BNuT was implemented across 8 health care organizations representing diverse contexts, including home health care, general hospitals, and academic centers. Although the degree of adoption varied, each site revealed distinct yet interrelated sociotechnical barriers that influenced feasibility and uptake:

1. At the human level, hesitance and abstention emerged as key barriers. While these responses may originate as reactions to perceived risks, the resulting behavioral patterns—reluctance to engage, delayed adoption, or outright avoidance—function as implementation barriers. Nursing leaders and managers expressed concern that

XSL•FO
RenderX

BNuT-generated insights might be misused in their performance evaluations. There was also apprehension about receiving feedback without having the resources or organizational support to act on it, creating a sense of accountability without agency. The operational pressures of managing day-to-day staffing crises left little cognitive bandwidth for strategic workforce optimization. This environment discouraged engagement with reflective data tools, despite general support for the underlying vision.

2. At the organizational level, siloed priorities and hierarchical governance structures undermined implementation. Financial executives, facing severe budget constraints (particularly relevant given that approximately 40% of Belgian hospitals operate at a financial loss [13,14], perceived workforce optimization tools as an additional cost rather than an investment. Chief technology officers, while conceptually supportive, found it difficult to justify allocating scarce IT resources to early-stage, nurse-led projects lacking immediate financial return. This division between nursing leadership (operational authority) and executive governance (strategic control) created a structural misalignment that limited implementation feasibility.
3. At the technological level, integration of existing systems proved prohibitively complex and expensive due to vendor lock-in strategies, legacy system fragmentation, and deliberate data walling between platforms. Several hospitals had outsourced IT operations, transforming what could have been internal configuration tasks into costly external service requests requiring formal budget approval. IT departments, already constrained by workforce shortages and competing project portfolios, often struggled to prioritize nurse-led digital innovations.

Out of the 8 health care organizations, one site achieved substantial success offering a contrasting perspective. Operating as a nurse-led home health organization, its leadership culture embedded the vision of data-driven workforce optimization from the outset. Alignment between operational need, executive vision, and organizational authority enabled integration of BNuT into management routines. This contrast underscores the critical role of organizational alignment and leadership ownership in realizing the potential of sociotechnical innovations such as BNuT.

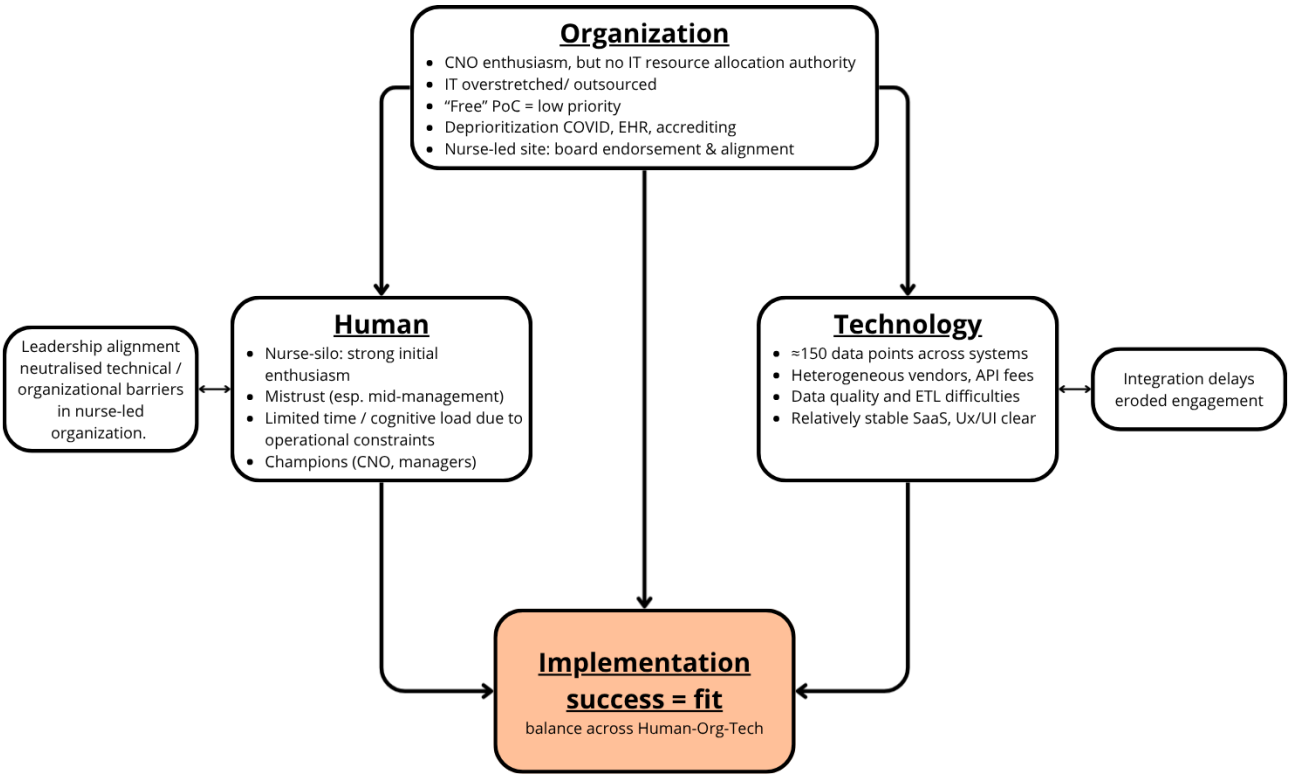
The 3 Dimensions of Barriers

Identifying the HOT-Fit

The implementation experiences of the BNuT project were examined through the HOT-fit framework [11]. The framework provides a sociotechnical lens for analyzing how system adoption and sustainability depend on the alignment, or fit, between 3 interacting domains: human, organizational, and technological. Our reflections demonstrate that BNuT's implementation challenges were not isolated to any single dimension. Rather, success depended on achieving an equilibrium across domains, and an imbalance in one frequently undermined progress in the others (Figure 2). This section reflects on lessons learned within each domain and concludes by considering the interdomain dynamics that shaped implementation outcomes.

XSL•FO
RenderX

Figure 2. HOT-fit framework applied to the BNuT implementation. The diagram illustrates the 3 interacting domains of the HOT-fit model and their relationships in the context of the BNuT decision-support system. Arrows represent interdependencies: successful implementation (fit) occurs only when all domains are simultaneously aligned. Callouts highlight two key observations: (1) strong fit within the organizational domain—where nursing leadership held both strategic authority and operational responsibility—enabled the nurse-led organization to overcome technical barriers, and (2) data-integration delays eroded user engagement. API: application programming interface; BNuT: balanced nursing teams; CNO: chief nursing officer; EHR: electronic health record; ETL: extract-transform-load; HOT: Human-Organization-Technology; PoC: proof of concept; SaaS: software-as-a-service; UX/UI: user experience/user interface.



Human: Trust, Empowerment, and Cognitive Capacity

Across the participating organizations, nursing leaders were consistently enthusiastic about BNuT's conceptual promise. The idea of a data-driven instrument capable of translating complex workforce information into actionable team-level insights resonated strongly with their daily struggles. Notably, the project itself originated from nursing leaders' explicit requests for better decision support tools, making BNuT a response to practitioner demand rather than a top-down research initiative. Early skepticism during prototype demonstrations largely dissipated once realistic use cases were presented. The urgency of developing such a system was widely acknowledged, and many leaders expressed relief that a project finally addressed the persistent gap between workforce research and managerial decision support.

> *The added value that this tool will be able to give us in the future was clear to everyone in the meeting. It will enable us to make an in-depth analysis based on objective data about the functioning of a team and also indicate on which elements action is required to remedy this.* [Chief Nursing Officer, participating hospital]

> *BNuT provides a most powerful lens for every nurse manager. Gain personal leadership and team insights to transform your department with data-driven decisions for better care and satisfied staff.* [Nurse Manager involved in BNuT pilot testing]

However, this enthusiasm coexisted with overconfidence regarding implementation feasibility. Many nursing leaders underestimated the organizational and technical barriers involved and overestimated their own authority to overcome them. Hospitals with advanced data infrastructures were paradoxically more hesitant, perceiving BNuT as "yet another system" rather than as a complement to existing business-intelligence tools. These organizations typically requested an Application Programming Interface (API) to integrate BNuT's algorithms into their own dashboards rather than adopting the full platform. Conversely, smaller organizations, particularly those without internal analytics capacity, perceived BNuT as uniquely valuable but lacked the resources to integrate it.

Hesitance surrounding data use proved a recurrent issue, though its nature differed markedly across organizational levels. Among frontline nurses, concerns were minimal once the survey's anonymity was assured; they generally perceived BNuT as a supportive and developmental tool rather than a threat. In contrast, hesitance and some mistrust were concentrated among junior and mid-level managers, who feared that aggregated results, particularly those reflecting leadership climate or organizational support, could be interpreted punitively by higher management. In one large university hospital, even the CNO voiced concern that unfavorable scores on "organizational leadership" might be politically exploited by other executives. These perceptions reveal the asymmetrical distribution of vulnerability within hierarchical hospital structures: while staff

XSL•FO
RenderX

trusted the system's safeguards, middle management occupied a position where performance data could plausibly affect professional standing.

Demonstrations and early pilots substantially increased perceived usefulness. Once users saw how disparate data streams converged into interpretable insights, engagement improved markedly. A frequently cited example involved incident reporting: one team appeared to have a disproportionately high number of reported safety incidents. Initially interpreted as poor performance, the BNuT analysis revealed that this team simultaneously scored highest on safety culture, indicating a healthy incidence-reporting climate. Such cases illustrated how integrated data interpretation could challenge misperceptions and reframe organizational learning.

Despite these successes, cognitive and temporal constraints limited sustained engagement. Nursing managers facing chronic staffing shortages and operational crises had little time to interact with a strategic-level dashboard. The tool's intuitive design mitigated digital-literacy barriers, yet the absence of formal onboarding and the overall cognitive load of crisis management restricted deeper adoption. Champions, typically CNOs or middle managers, emerged in nearly all sites, but their influence was constrained by limited institutional authority. Over time, enthusiasm at the human level could not compensate for systemic resource constraints and organizational inertia.

However, it must also be acknowledged that the system's own scope likely contributed to the cognitive and governance burden experienced by potential adopters. BNuT encompasses 10 domains and integrates approximately 250 data points, of which roughly 150 are to be extracted from organizational systems. While the user-facing interface was specifically designed to simplify interpretation by reducing this complexity to composite indicators and visual summaries, the underlying data requirements imposed demands on organizational processes that were substantial, particularly for institutions already operating under severe resource constraints. For organizations with limited digital maturity, the gap between the system's data needs and the institution's capacity to deliver that data may have been too large, regardless of the tool's analytical value. In retrospect, a more modular deployment strategy might have lowered the threshold for initial adoption: beginning with a reduced core indicator set drawn from readily available data and progressively activating additional domains as organizational data capacity matured would have allowed trust, familiarity, and demonstrable value to develop incrementally before the full model was introduced.

Organizational: Leadership Alignment and Structural Inertia

The organizational context proved decisive. In all hospitals, the CNOs were strong advocates for BNuT, yet they lacked budgetary control. IT departments, already overstretched and managing numerous competing projects, had to allocate technical resources, while financial decisions ultimately rested with chief financial officers. Because the project was offered free of charge during the proof-of-concept phase, it competed poorly for attention against commercial initiatives with contractual obligations. In hindsight, requesting a modest financial buy-in might have enhanced commitment and prioritization by signaling ownership.

Both HR and IT departments often viewed BNuT as redundant. HR staff perceived it as encroaching on their existing tools, while IT departments regarded it as a low-priority research experiment. Chronic shortages of IT personnel further reduced their willingness to support time-consuming integrations. In several hospitals, IT services had been partially outsourced, transforming even minor configuration tasks into expensive external contracts. In some cases, hospitals lacked in-house capacity to access their own databases, as software vendors maintained exclusive control and charged high fees for basic queries. These arrangements effectively held data "hostage" and created a very high dependency on vendors for data extraction.

Organizational fit, the degree to which nursing leadership held authority corresponding with operational responsibility, emerged as the critical differentiator. In hierarchical organizations where the nursing department functioned as a silo, CNOs struggled to gain board-level approval. Projects framed as nurse-led were perceived as peripheral to strategic objectives. By contrast, the one nurse-led organization integrated BNuT into its formal strategic plan, supported by executive leadership and dedicated IT allocation. This fit between operational need, strategic vision, and resource authority was essential. It enabled the project to progress from conceptualization to sustained use, highlighting the necessity of structural empowerment for nursing leadership within governance systems.

> *We use BNuT as an integrated data and communication system in our organization to better understand our practices and outcomes in home health and to lead and manage our organization proactively.* [CEO Home Health Organization]

Competing institutional priorities further hindered progress. Post-COVID recovery, ongoing EHR upgrades, and accreditation audits absorbed most organizational attention. Because BNuT did not have immediate regulatory or financial implications, it was consistently deprioritized. Organizational culture also played a role: hospitals with participatory cultures tended to show greater openness to collaboration, whereas highly hierarchical institutions rarely progressed beyond exploratory discussions. Overall, BNuT's organizational challenges reflected the entrenched marginalization of nursing leadership within hospital governance, a systemic issue that extends beyond this specific innovation.

Technological: Interoperability, Data Quality, and Usability

The technological domain presented the most concrete and persistent barriers. BNuT requires integration of about 250 data points, where approximately 150 data points can be drawn from existing systems: HRs, scheduling, EHRs, quality registries, and even locally maintained spreadsheets. Each organization used a unique constellation of software vendors and data formats, often without shared standards. Even when multiple hospitals used the same vendor, integration had to be redeveloped from scratch due to licensing and cost restrictions.

XSL•FO
RenderX

Vendors frequently imposed additional charges for API access or custom data exports, and in several cases claimed that bespoke queries were required for every new request.

Manual preprocessing was initially necessary to conform to the BNuT data structure. Anticipating this challenge, the system was designed with a flexible, tiered integration architecture: organizations could upload data manually per data point through the web interface, push standardized files through an automated importer in an agreed-upon format, or connect directly to a PostgreSQL database structure into which raw data could be pushed with automated ETL processing. To further reduce this burden, the project contracted an external ETL specialist who configured direct connections to source systems where feasible. Despite this flexibility, the approach proved only partially effective due to heterogeneous database architectures, limited vendor documentation, and the cost of obtaining data exports from third-party software providers. Data quality issues were widespread: implausible values were common, and missingness varied widely between datasets. BNuT's algorithms were designed to tolerate partial missingness, except for a few core variables essential for standardization, but the variability still limited analytic robustness. The persistence of these barriers, even given the system's deliberately flexible architecture, underscores that the bottleneck was structural and rooted in vendor lock-in, fragmented data governance, and limited hospital IT capacity.

Despite these integration difficulties, the user-facing technology was well received. A professional user experience/user interface designer ensured that the interface was intuitive and visually clear, with a simplified radar chart summarizing 10 domains of team performance. Users could benchmark teams or explore subdomains with minimal training. The system was web-based (Software as a Service) and accessible through standard browsers, providing sufficient performance and stability for managerial use.

Maintenance and updates were provided by the research team in collaboration with a contracted developer. Hospitals were able to independently distribute surveys and upload local data, and feedback loops for technical improvement remained informal. Overall, the technical obstacles were concentrated in data extraction and integration rather than in software usability or performance.

A critical question arising from the BNuT experience is whether a model integrating approximately 250 data points was over-engineered. The breadth of the model reflects our conceptual ambition of capturing team balance across all domains identified as relevant in the empirical literature [9], an ambition that was codeveloped and endorsed by the participating organizations themselves. However, the practical consequence was a data-governance burden that exceeded the current capacity of most hospital sites, particularly those reliant on legacy systems, outsourced IT functions, or fragmented data architectures. Even the flexible integration mechanisms provided could not overcome the structural barriers imposed by vendor lock-in and the cost of data extraction from proprietary systems. This tension between comprehensiveness and feasibility represents a design challenge that is not unique to BNuT but characterizes many health information systems that attempt to bridge the gap between research-level measurement and operational reality [15]. Future iterations of such systems should consider adopting a modular architecture where a core set of indicators (operationalizable with minimal integration effort and immediately actionable for nursing managers) serves as the entry point, while the full model is reserved for organizations with sufficient digital infrastructure and data-governance capacity.

Interdomain Fit: A Prerequisite for Sustainability

Analysis across HOT-fit domains highlights that implementation success depended on achieving simultaneous fit across all 3 dimensions. The nurse-led organization illustrated how a strong fit within the organizational domain, where nursing leadership held both strategic authority and resource control, could compensate for barriers in other areas. Once executive endorsement was secured, IT resources were allocated, data-sharing agreements expedited, and BNuT was embedded in strategic planning. The same technical hurdles that halted progress elsewhere became manageable. Conversely, where one domain lacked fit, most commonly technology, momentum in the others dissipated.

In several hospitals, delays in data integration eroded enthusiasm among nursing leaders who initially championed the tool. Similarly, where organizational structures prevented nursing leadership from influencing IT or financial decisions, human-level motivation could not translate into implementation capacity. Positive feedback loops only emerged when success was visible and endorsed from the top: early demonstrable value encouraged further engagement and resource allocation, but such cycles remained rare.

Overall, these patterns confirm the HOT-fit framework's central proposition: the effectiveness of health information systems depends not merely on the functionality of individual components but on the equilibrium between them. Imbalance in any single domain propagates across the system, undermining overall fit. In BNuT's case, technological constraints, particularly data integration and interoperability, acted as the principal bottleneck, yet organizational and cultural dynamics determined whether these constraints could be overcome. Strong leadership endorsement, clear strategic positioning of nursing within governance structures, and proactive attention to data-trust and communication dynamics proved essential conditions for sustaining progress. The nurse-led success case offers a critical insight: when nursing leadership holds both strategic authority and operational responsibility, implementation becomes feasible. In this organization, the same technical challenges that stalled progress elsewhere were systematically addressed because nursing leadership could directly allocate resources, prioritize IT projects, and embed BNuT within strategic planning. This demonstrates that the barriers encountered elsewhere were not primarily technical but organizational. Conversely, when nursing remains positioned as a cost-center with limited influence over budgets and technology decisions, even technically sound innovations may struggle to take root. Achieving fit across human, organizational, and technological systems is therefore not only a condition for

XSL•FO
RenderX

digital success but also a reflection of the broader power balance between clinical professions and institutional structures within health care.

Discussion

Principal Findings

The main finding of this reflective analysis is that the implementation of BNuT succeeded only where simultaneous alignment across human, organizational, and technological domains was achieved, specifically in a nurse-led home health organization where leadership held both strategic vision and budgetary authority. Across the 7 hospital sites, adoption was prevented by converging barriers: data interoperability failures driven by vendor lock-in and fragmented IT architectures, restricted nursing leadership authority within hierarchical governance structures, and the substantial data-governance demands imposed by the system's own breadth. These findings are not unique to the Belgian context but reflect structural dynamics widely documented in health informatics.

More specifically, the BNuT implementation illustrates how organizational complexity constrains decision-support adoption in nursing practice. Across sites, barriers arose primarily from misalignment between human, organizational, and technological domains, although BNuT's own data requirements may also have exceeded the digital capacity of participating institutions. This finding echoes long-standing observations within implementation science that the success of health information technology depends less on technical soundness than on contextual fit within existing workflows, governance, and professional cultures [16].

The predominance of technological barriers parallels systematic reviews showing that interoperability failures remain the most persistent obstacle to decision-support integration [17,18]. As in those studies, the inability to access or standardize data across HR, scheduling, and clinical systems eroded perceived reliability and constrained analytic capability. These "barriers of insufficiency" [17] reflect structural deficits in hospital IT resourcing rather than isolated project limitations, reinforcing the need for policy-level interoperability mandates.

The vendor lock-in and data fragmentation observed in the BNuT project are not isolated to the Belgian health care context. Internationally, similar dynamics have been extensively documented. In England, the National Programme for IT (NPfIT) was ultimately dismantled in 2011, in part because of vendor lock-in and the inability to achieve interoperability across hospital trusts [19,20]. Subsequent NHS strategies have mandated open standards and fast health care interoperability resources–based interoperability, yet in practice, vendor-controlled data architectures continue to impede secondary use of clinical and workforce data [21-23]. In the United States, the 21st Century Cures Act introduced explicit information-blocking provisions to prevent vendors from restricting data access, reflecting federal recognition that proprietary data practices constitute a systemic barrier to innovation [24]. Within the European Union, the proposed European Health Data Space regulation represents an emerging attempt to establish cross-border interoperability standards that could, in principle, reduce the per-project negotiation costs that proved limiting in the BNuT experience [25]. Collectively, these international examples confirm that the barriers encountered by BNuT reflect a structural problem in health IT governance, and that no amount of tool-level design can compensate when system-level interoperability standards are absent.

At the organizational level, the imbalance between nursing leadership enthusiasm and restricted resource authority emerged as decisive. BNuT encountered resistance not from overt opposition but from structural dependency on nonclinical decision-makers, reflecting what has been described as the "negotiation of control" in health care technology implementation [26]. When nursing leaders lacked access to financial or technical levers, implementation stalled despite clinical support [27]. This pattern may reflect broader international challenges: research suggests that even within integrated health systems, resource allocation and workflow redesign remain chronic bottlenecks [28], while CNO authority limitations and IT resource allocation challenges appear to be documented across health care systems globally [17,27]. The single nurse-led organization in our sample, where strategic and budgetary authority were unified under nursing leadership, achieved full operational use, empirically supporting the HOT-fit premise that fit across all domains is a prerequisite for sustainability.

The human dimension was characterized by enthusiasm to hesitation, abstention to mistrust. Nurses did not resist technology per se, but distrusted implementations perceived as top-down or evaluative [29]. The anonymity of BNuT's survey component mitigated some of these concerns, yet lingering fears of data misuse illustrate fragile power dynamics [26]. Time pressure and cognitive overload further limited engagement, with research showing that human-resource scarcity continues to undermine even well-designed interventions [17].

Taken together, our findings reinforce the view of health care organizations as complex adaptive systems in which technical, organizational, and human subsystems are tightly coupled [16]. Localized improvements in one domain cannot compensate for deficits in another. The nurse-led success case exemplifies this interdependence: strong executive sponsorship overcame technical and organizational inertia, producing a functioning sociotechnical equilibrium. In contrast, sites lacking such fit experienced cascading failure: data integration delays reduced managerial trust, which in turn decreased organizational priority. Implementation success thus required "all lights green," confirming HOT-fit's central premise that sustainable adoption depends on balanced interaction across domains.

The conceptual contribution of BNuT lies in extending decision support beyond patient-level staffing ratios towards team-level optimization, incorporating leadership, workload, and safety culture. This holistic approach addresses gaps identified in recent reviews of nursing decision-support systems, which report limited theoretical grounding and minimal postimplementation evaluation [30]. Yet the difficulties experienced in this project also underscore the persistent divide between research prototypes and operational systems. Bridging this divide will

XSL•FO
RenderX

require not only improved technical integration but also structural empowerment of nursing leadership to act as equal partners in digital transformation.

The lessons from BNuT have concrete implications for the design of future nursing decision-support systems. First, scalability demands a shift from monolithic, all-or-nothing deployment towards modular architectures that allow organizations to adopt core functionality with minimal integration effort and progressively expand as digital maturity grows. An API-first, cloud-native design would reduce per-site configuration costs and enable integration with diverse IT ecosystems without requiring bespoke development for each vendor. Second, adaptability requires that indicator sets are configurable to organizational context: rather than imposing a fixed 250-variable model, future systems should allow institutions to select domains aligned with their strategic priorities and data availability, with the full model serving as an aspirational benchmark. Third, integrating predictive analytics, for example, flagging units where scheduled team composition falls below stability thresholds, or forecasting turnover risk based on employment pattern trends, would shift such tools from retrospective reporting towards prospective decision support, increasing their value for both unit-level workforce management and organization-wide strategic planning. Fourth, interoperability should be embedded by design rather than retrofitted: adopting open data standards and European Health Data Space–compatible data models from the outset would reduce dependency on vendor-specific export mechanisms. For clinical data, standards such as HL7 fast health care interoperability resources provide a foundation; however, workforce-specific data (including turnover, employment patterns, and team composition metrics) are not yet covered by existing interoperability standards, highlighting the need for standardized data models for nursing workforce analytics as a prerequisite for tools like BNuT.

Taken together, the findings demonstrate that sustainable digital innovation in health care depends on achieving fit across human, organizational, and technological systems. Where one domain fails, whether through technical fragmentation, limited leadership authority, or mistrust, momentum across the others may diminish. Conversely, fit across all domains enables small successes to compound into sustained adoption.

Furthermore, they reveal a troubling urgency paradox. Given mounting workforce shortages, cost pressures, and quality demands, strategic optimization of nursing team deployment should arguably be among health care organizations' highest priorities. The nursing workforce represents both the highest operational cost and the most direct interface between an organization and its patients. Yet across the sites studied, workforce optimization tools were consistently deprioritized in favor of regulatory compliance systems, billing infrastructure, or projects with more immediate financial returns. This misalignment between stated workforce priorities and actual resource allocation warrants attention from nursing leadership and policymakers alike. Translating these insights into practice yields several actionable lessons for the key stakeholders shaping the future of the nursing workforce.

Several limitations of this viewpoint must be acknowledged. First, the authors were the developers of BNuT, which creates an inherent risk of presenting the tool more favorably than an independent analysis might. We have tried to apply a self-critical lens, particularly in examining whether the system's own complexity contributed to adoption failure, but readers should interpret the analysis with this in mind. Second, the reflections presented here are grounded in 8 years of sustained stakeholder engagement rather than in formal qualitative data collection with recorded interviews, systematic coding, and member checking. This gives us considerable ecological breadth and depth of immersion, but lacks the methodological safeguards of structured qualitative research, and the insights may be subject to recall bias and selective interpretation. Third, all implementation attempts occurred within the Belgian health care system, with its specific configuration of hospital governance, vendor markets, and regulatory frameworks. Although we have drawn parallels with international experiences, direct generalizability to other jurisdictions should not be assumed. Fourth, criterion validation of BNuT against patient and staff outcomes was planned but could not be completed because the sociotechnical barriers described in this paper prevented sustained implementation. The system's effectiveness, therefore, remains empirically undemonstrated, and the tool should be understood as a theoretically informed prototype rather than a validated intervention. Fifth, the contrast between the single successful site and the 7 hospital sites should be interpreted with caution: the successful organization differed structurally in governance, scale, and patient population, and the comparison cannot isolate which specific factors were sufficient or necessary for adoption. Nonetheless, we believe that systematically documenting implementation failure through a structured analytical framework serves an important function in the health informatics literature, where publication bias towards successful cases limits collective learning.

Lessons Learned

The synthesis of our findings, summarized in Figure 2, yields actionable lessons for the diverse stakeholders involved in digital nursing innovation. While grounded in the BNuT experience, these insights reflect broader structural issues relevant to sociotechnical implementation in health care.

Researchers

For researchers developing digital decision-support tools, the BNuT case offers a sobering lesson. BNuT was developed bottom-up from explicit practitioner demand: nursing leaders requested better decision-support tools, and 18 health care organizations collectively crowdfunded its development. The project incorporated end user co-design throughout, with nursing managers and CNOs involved in iterative development cycles. Implementation science principles guided the approach, including stakeholder mapping, staged rollout, and continuous evaluation. Yet despite these methodological strengths, BNuT still encountered fundamental barriers that prevented adoption in most sites. This suggests that the obstacles to data-driven decision support in nursing may be more systemic than methodological: even well-designed, practitioner-endorsed innovations may fail when organizational structures do not grant

XSL•FO
RenderX

nursing leadership the authority to implement them. Researchers should therefore embed evaluation within a sociotechnical framework from the outset, document contextual mechanisms rather than merely outputs, and advocate for the structural conditions that enable adoption, not only develop technically sound tools.

Health Care Organizations

For health care organizations, our analysis underlines that the success of nursing informatics projects depends on granting nursing leadership formal authority over budget and data infrastructure. CNO enthusiasm alone cannot compensate for siloed governance or outsourced IT functions where access to the health care organization's own data is locked by software vendors. Health care organizations need to address this issue in future contract negotiations with the software vendor so that access to their proper data can be faster, more flexible, and at a much lower additional cost. Organizations that positioned nursing as an equal partner in strategic planning achieved smoother integration and sustained use. Embedding nursing analytics tools such as BNuT within long-term workforce strategies, rather than treating them as stand-alone one-fits-all pilots, fosters alignment across organizational hierarchies and improves accountability.

Policymakers

For policymakers, the BNuT experience highlights that the barriers to digital nursing innovation are systemic rather than local. Fragmented data architectures, vendor lock-in, and inconsistent interpretations of data-protection regulation create environments where interoperability remains prohibitively expensive. Concretely, this requires policy action on multiple fronts. First, mandatory interoperability standards for health IT vendors would reduce the per-project negotiation costs that limited BNuT's data integration. Second, dedicated pilot funding mechanisms for nurse-led digital health innovation could provide the financial continuity that free proof-of-concept approaches cannot sustain. The BNuT experience suggests that offering a tool free of charge may paradoxically reduce organizational commitment; phased funding models tied to implementation milestones might better incentivize sustained engagement. Third, governance reforms that mandate nursing representation on hospital IT steering committees and digital transformation boards would address the chronic underrepresentation of nursing leadership that proved decisive across our implementation sites. Addressing these "barriers of insufficiency" [17] requires structural investment in digital infrastructure and governance. Equally essential is embedding nursing leadership within national digital health strategies. Policies that enable nurse leaders to access and use workforce data are not merely equity measures but economically rational interventions that enhance quality, safety, and system performance.

Conclusions

The BNuT experience demonstrates that the challenge of digital transformation in nursing may lie less in technology than in achieving fit across sociotechnical domains. Sustainable innovation appears to depend on the concurrent readiness of human, organizational, and technological systems, none of which can substitute for another. When fit is achieved, as in the nurse-led organization, digital tools can become enablers of strategic decision-making rather than isolated projects. The persistent barriers encountered across most sites reflect structural issues that extend far beyond any single system: the undervaluation of nursing leadership, fragmented data infrastructures, and chronic underinvestment in interoperability. Equally, our experience suggests that decision-support systems of this scope must be designed with explicit attention to the data-governance capacity of target organizations, adopting modular architectures that lower the threshold for initial adoption. Addressing these systemic imbalances—in interoperability governance, nursing leadership authority, and the design of implementation-ready tools—will be essential not only for future iterations of BNuT but for the broader success of nursing informatics as a driver of resilient and learning health care organizations.

Acknowledgments

The authors gratefully acknowledge the 18 health care institutions that collectively crowdfunded the BNuT initiative and the members of the Icuro steering committee for their sustained engagement and strategic guidance throughout the project. Sincere thanks are extended to the development partners at Dashdot and Basil for their technical collaboration.

Funding

SV received a PhD grant from the Special Research Fund (Bijzonder Onderzoeksfonds, BOF) of the University of Antwerp.

Authors' Contributions

SV contributed to conceptualization, project administration, visualization, writing of the original draft, and review and editing of the manuscript. PVB and FH contributed to conceptualization, supervision, validation, and review and editing of the manuscript. WDK contributed to the review and editing of the manuscript. The authors used a large language model (Anthropic Claude) to assist in the writing process, but remain fully responsible for the content.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Galanis P, Moisoglou I, Katsiroumpa A, et al. Increased job burnout and reduced job satisfaction for nurses compared to other healthcare workers after the COVID-19 pandemic. Nurs Rep 2023 Aug 14;13(3):1090-1100. [doi: 10.3390/nursrep13030095] [Medline: 37606463]
2. Haegdorens F, Van Bogaert P, De Meester K, Monsieurs KG. The impact of nurse staffing levels and nurse's education on patient mortality in medical and surgical wards: an observational multicentre study. BMC Health Serv Res 2019 Nov 21;19(1):864. [doi: 10.1186/s12913-019-4688-7] [Medline: 31752859]
3. Aiken LH, Sloane DM, Bruyneel L, et al. Nurse staffing and education and hospital mortality in nine European countries: a retrospective observational study. Lancet 2014 May 24;383(9931):1824-1830. [doi: 10.1016/S0140-6736(13)62631-8] [Medline: 24581683]
4. Ausserhofer D, Zander B, Busse R, et al. Prevalence, patterns and predictors of nursing care left undone in European hospitals: results from the multicountry cross-sectional RN4CAST study. BMJ Qual Saf 2014 Feb;23(2):126-135. [doi: 10.1136/bmjqs-2013-002318]
5. Drennan VM, Ross F. Global nurse shortages-the facts, the impact and action for change. Br Med Bull 2019 Jun 19;130(1):25-37. [doi: 10.1093/bmb/ldz014] [Medline: 31086957]
6. Shembavnekar N, Buchan J, Bazeer N, et al. REAL centre projections. : The Health Foundation; 2022 URL: https://www.health.org.uk/sites/default/files/2022-07/REALCentreWorkforceProjections_EMBARGOED.pdf
7. Ying LY, Ramoo V, Ling LW, et al. Nursing practice environment, resilience, and intention to leave among critical care nurses. Nurs Crit Care 2021 Nov;26(6):432-440. [doi: 10.1111/nicc.12551] [Medline: 32929840]
8. Bae SH. Noneconomic and economic impacts of nurse turnover in hospitals: a systematic review. Int Nurs Rev 2022 Sep;69(3):392-404. [doi: 10.1111/inr.12769] [Medline: 35654041]
9. Vleminckx S, Van Bogaert P, De Meulenaere K, Willem L, Haegdorens F. Factors influencing the formation of balanced care teams: the organisation, performance, and perception of nursing care teams and the link with patient outcomes: a systematic scoping review. BMC Health Serv Res 2024 Sep 27;24(1):1129. [doi: 10.1186/s12913-024-11625-5] [Medline: 39334182]
10. Griffiths P, Saville C, Ball J, et al. Nursing workload, nurse staffing methodologies and tools: a systematic scoping review and discussion. Int J Nurs Stud 2020 Mar;103:103487. [doi: 10.1016/j.ijnurstu.2019.103487] [Medline: 31884330]
11. Yusof MM, Paul RJ, Stergioulas LK. Towards a Framework for Health Information Systems Evaluation. : IEEE; 2006 Presented at: Proceedings of the 39th annual Hawaii International Conference on System Sciences (HICSS'06); Jan 4-7, 2006; Kauia, HI, USA. [doi: 10.1109/HICSS.2006.491]
12. Bogaert P, Clarke S. Future steps in practice and research. In: The Organizational Context of Nursing Practice: Concepts, Evidence, and Interventions for Improvement: Springer; 2017:297-307. [doi: 10.1007/978-3-319-71042-6_14]
13. MAHA-sectoranalyse. Belfius Bank. 2024. URL: https://www.belfius.be/imagingservlet/GetDocument?src=mifid&id=EEPVRACMAHA24_NL [accessed 2025-10-17]
14. Belfius Bank. MAHA-sectoranalyse.: Belfius Bank; 2025. URL: https://www.belfius.be/imagingservlet/GetDocument?src=mifid&id=EEPMAHACOMALL2024_NL
15. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
16. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. BMJ Qual Saf 2010 Oct 1;19(Suppl 3):i68-i74. [doi: 10.1136/qshc.2010.042085]
17. Chen W, O'Bryan CM, Gorham G, et al. Barriers and enablers to implementing and using clinical decision support systems for chronic diseases: a qualitative systematic review and meta-aggregation. Implement Sci Commun 2022 Jul 28;3(1):81. [doi: 10.1186/s43058-022-00326-x] [Medline: 35902894]
18. Abell B, Naicker S, Rodwell D, et al. Identifying barriers and facilitators to successful implementation of computerized clinical decision support systems in hospitals: a NASSS framework-informed scoping review. Implement Sci 2023 Jul 26;18(1):32. [doi: 10.1186/s13012-023-01287-y] [Medline: 37495997]
19. Campion-Awwad O, Hayton A, Smith L, Vuaran M. The National Programme for IT in the NHS. A case history. : University of Cambridge; 2014 URL: https://www.alejandrobarros.com/wp-content/uploads/2019/07/Historia-Proyecto-TI-NHS.pdf
20. Wachter R. Making IT work: harnessing the power of health information technology to improve care in England. : Department of Health; 2016 URL: https://rpsg.org.uk/wp-content/uploads/2020/04/Technology-to-Improve-Care-in-England.pdf
21. Bennion MR, Spencer R, Moore RK, Kenyon R. Digital capability, open-source use, and interoperability standards within the national health service in England: survey of health care trusts. JMIR Hum Factors 2025 Aug 12;12:e66398. [doi: 10.2196/66398] [Medline: 40795322]
22. Cresswell K, Sheikh A, Krasuska M, et al. Reconceptualising the digital maturity of health systems. Lancet Digit Health 2019 Sep;1(5):e200-e201. [doi: 10.1016/S2589-7500(19)30083-4] [Medline: 33323267]
23. Department of Health and Social Care. Data saves lives: reshaping health and social care with data. HM Government. 2022. URL: https://www.gov.uk/government/publications/data-saves-lives-reshaping-health-and-social-care-with-data/data-saves-lives-reshaping-health-and-social-care-with-data [accessed 2026-03-17]

XSL•FO
RenderX

24. Adler-Milstein J, Pfeifer E. Information blocking: is it occurring and what policy strategies can address it? Milbank Q 2017 Mar;95(1):117-135. [doi: 10.1111/1468-0009.12247] [Medline: 28266065]
25. Terzis P, Santamaria Echeverria OE. Interoperability and governance in the European Health Data Space regulation. Med Law Int 2023 Dec;23(4):368-376. [doi: 10.1177/09685332231165692]
26. Liberati EG, Ruggiero F, Galuppo L, et al. What hinders the uptake of computerized decision support systems in hospitals? A qualitative study and framework for implementation. Implement Sci 2017 Sep 15;12(1):113. [doi: 10.1186/s13012-017-0644-2] [Medline: 28915822]
27. Alexander C, Tschannen D, Argetsinger D, Hakim H, Milner KA. A qualitative study on barriers and facilitators of quality improvement engagement by frontline nurses and leaders. J Nurs Manag 2022 Apr;30(3):694-701. [doi: 10.1111/jonm.13537] [Medline: 34969172]
28. Shi Y, Amill-Rosario A, Rudin RS, et al. Barriers to using clinical decision support in ambulatory care: do clinics in health systems fare better? J Am Med Inform Assoc 2021 Jul 30;28(8):1667-1675. [doi: 10.1093/jamia/ocab064] [Medline: 33895828]
29. Pareek A, Soni H, Bisen A, et al. Nursing informatics adoption in clinical settings: overcoming barriers and leveraging facilitators. Universe Int J Interdiscip Res 2024 Oct;5:97-107 [FREE Full text]
30. Abi Khalil C, Saab A, Rahme J, Bouaud J, Seroussi B. Capabilities of computerized decision support systems supporting the nursing process in hospital settings: a scoping review. BMC Nurs 2025 Jul 1;24(1):677. [doi: 10.1186/s12912-025-03272-w] [Medline: 40598308]

Abbreviations

API: application programming interface
BNuT: balanced nursing teams
CNO: Chief Nursing Officer
EHR: electronic health record
HOT-fit: Human-Organization-Technology fit
HR: human resource
NPfIT: National Programme for IT

Edited by D Dordunoo; submitted 29.Jan.2026; peer-reviewed by B Dovdon, KK Lai; revised version received 24.Mar.2026; accepted 01.Apr.2026; published 11.May.2026.

Please cite as:
Vleminckx S, Van Bogaert P, De Keyser W, Haegdorens F
From Concept to Practice: Lessons From the Balanced Nursing Teams Decision-Support System
JMIR Nursing 2026;9:e92417
URL: https://nursing.jmir.org/2026/1/e92417
doi:10.2196/92417

© Senne Vleminckx, Peter Van Bogaert, Wim De Keyser, Filip Haegdorens. Originally published in JMIR Nursing (https://nursing.jmir.org), 11.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Artificial Intelligence–Enhanced Wound Care to Improve Access, Efficacy, and Equity in Wound Care for Older Adults in Rural and Remote Regions of Canada

Courtney Genge[1], PhD; Basnama Ayaz[2,3], PhD; Shannon Freeman[4,5], PhD; Heba Tallah Mohammed[3], PhD; Robert D J Fraser[3,6], PhD; Ibukun-Oluwa Omolade Abejirinde[2,7], PhD, MD; Deirdre O'Sullivan-Drombolis[8], MCISc; Rebecca Brookham[9], PhD

[1]Aging in Place Challenge Program, National Research Council Canada, 1200 Montréal Rd, Gloucester, Ottawa, ON, Canada
[2]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[3]Swift Medical, Toronto, ON, Canada
[4]School of Nursing, University of Northern British Columbia, Prince George, BC, Canada
[5]Centre For Technology Adoption for Aging in the North, Prince George, BC, Canada
[6]Arthur Labatt School of Nursing, Western University, London, ON, Canada
[7]Institute for Better Health, Trillium Health Partners, Mississauga, ON, Canada
[8]Giishkaandago'Ikwe Health Services, Fort Frances, ON, Canada
[9]Brightshores Health System, Owen Sound, ON, Canada

Corresponding Author:
Courtney Genge, PhD
Aging in Place Challenge Program, National Research Council Canada, 1200 Montréal Rd, Gloucester, Ottawa, ON, Canada

Abstract

Wound care is an increasing global challenge, with older adults among those most affected. As populations age, the demand for effective and efficient wound care increases. Over the years, various wound assessment and care techniques have been developed, including digital wound care technology (DWCT), which uses innovative artificial intelligence (AI). Many older adults, especially those living in rural and remote areas, face significant barriers in obtaining timely and effective wound care, leading to poorer health outcomes and increased health care costs related to wound care. These challenges underscore the urgent need to implement wound care models that equitably improve access to care and enhance clinical outcomes, particularly for older adults, to promote healthy aging and age-in-place. Based on evidence from the literature and the initial implementation of a DWCT in 2 community health systems in Ontario, this viewpoint paper encourages clinicians and health care leaders to embrace and expand the implementation of an AI-driven DWCT to address inequities in access to high-quality, timely care. The experiences from these implementations indicate that the use of AI can support clinical decision-making and extend access to care for individuals in rural and remote communities in Canada. By leveraging DWCT powered by AI, health care providers can enhance the accuracy and consistency of wound assessments, improve communication, streamline care processes, and more effectively allocate resources, ultimately aiming to reduce disparities in wound care outcomes.

(JMIR Nursing 2026;9:e85644) doi:10.2196/85644

KEYWORDS

wound care; AI; equity; artificial intelligence

Introduction

Wound care represents a growing challenge worldwide, particularly impacting older adults [1]. As populations age, the demand for effective and efficient wound care increases, driven by the growing number of older adults with multiple chronic conditions that predispose them to complex wounds and longer healing times. Many older adults, especially those living in rural and remote areas, face significant barriers in obtaining timely and effective wound care, leading to poorer health outcomes and increased health care costs [2]. These challenges underscore the urgent need to implement wound care models, which equitably improve access to care and enhance clinical outcomes. Traditionally, wound assessment is conducted visually to evaluate anatomical, physiological, and mechanical domains [3] or with subjective methods such as the Pressure Ulcer scale for Healing (PUSH) and Pressure Sore Status Tool (PAST) scales; these methods are widely used, can lack accuracy and consistency, leading to variable wound evaluation reports when administered by different clinicians [4].

XSL•FO
RenderX

With advances in science and technology, a range of instruments has been developed to address the growing demand for wound care. Among these, noninvasive imaging techniques have received particular attention for their potential to support wound assessment and monitoring. A systematic review of 20 studies done in 2015 summarized available approaches, including 2D stereophotogrammetry and 3D imaging methods. The authors concluded that 2D imaging is a practical wound measurement technique, offering ease of use and reliable assessments of wound size and healing progress; however, they noted that its inability to capture wound depth limits its value for care planning [5]. Building on this, a 2022 scoping review of 156 global studies examined the development, evaluation, and implementation of digital health technologies (eg, imaging and measurement) for wound care [6]. This review highlighted growing adoption of digital tools, citing evidence of their acceptability, feasibility, effectiveness, and clinical impact. At the same time, it identified important challenges, including limited evidence in hard-to-reach populations (eg, those in remote regions), low uptake linked to noninclusive design processes, and uncertainty around cost-effectiveness. It is essential to note that this review was not specifically tailored to older adults, a population at elevated risk of wound development and delayed healing due to comorbid conditions, and who also experience disproportionate barriers to care arising from age-related, geographic, and socioeconomic factors.

This study is a forward-looking perspective that highlights the current context of wound care in Canada and explores the potential for artificial intelligence (AI) driven digital wound care technology (DWCT) to improve access, efficiency, and equity. The authorship of this study reflects the multiple perspectives shaping this work, including 2 community health systems implementing a DWCT, the evaluation team supporting that implementation, academic and government partners advancing digital health innovation, and representatives from the company developing the technology. By bringing together these complementary experiences, we aim to provide a salient commentary on the merits of AI-augmented DWCT and its potential to advance equity, access, and efficiency in wound care in Canada.

The initial implementation of a DWCT in 2 rural and/or remote community health care systems, presented here as case studies, illustrates the contextual challenges faced by health care providers and patients while also offering practical insights into the technology's application and benefits. Discussion of anticipated challenges in implementing AI technologies in wound care provides a roadmap for future research directions, including identification of potential facilitators and barriers to adoption and uptake of the DWCT in the 2 networked community health systems. The potential for AI in transforming health care delivery is vast, particularly in the field of wound care, and can play an important role in ensuring that all patients receive high-quality, equitable care across diverse geographies.

Wound Care in Canada

The Canadian population is experiencing a significant demographic shift, with older adults (65 years and older) rapidly becoming the fastest-growing demographic, and by 2051, they could represent 25% of the population [7]. Aging is associated with the accumulation of cellular and molecular damage within the body, which can manifest as frailty and/or chronic health conditions [8], which increases vulnerability to complex wounds, meaning the number of older adults needing wound care is anticipated to increase substantially in the future [9]. Although the exact prevalence of wounds requiring care in Canada is unclear, estimates suggest that up to CAD $11.1 billion (US $8.18 billion) are spent annually on wound care [9]. Approximately 1 in 500 Canadians develops nonhealing wounds [10], which significantly impacts the health care system [11]. In Ontario alone, 30% - 50% of health care delivery involves patients with wounds, predominantly affecting older adults who often heal at home [11].

In 2019, older adults constituted 17% of Canada's population but accounted for 47% of total health care costs [10]. Rural and remote areas of Canada are aging more rapidly than their urban counterparts [12]; yet, they often have fewer health resources, especially health human resources (HHR) [2]. This lack of access to timely treatment for complex wounds leads to improper treatment of up to 25% of Canadian patients [13]. Chronic wounds impose significant physical and emotional burdens, decreasing patients' quality of life and increasing health care costs [14,15]. Without regular, high-quality care, these wounds can lead to serious complications such as infections, hospitalization, amputation, and even death [1].

The complexity of wound care is exacerbated by limited opportunities for comprehensive wound care education for health care providers, who typically receive fewer than 10 hours of formal instruction [16]. During training, health care providers are taught to provide quality care and maintain optimal functioning, including the treatment of health conditions such as wounds, through supportive, preventive, therapeutic, palliative, and rehabilitative care. However, these professionals need additional training to fulfill the specific needs of care and their respective professional scope of practice. For example, the College of Ontario of Nurses specified the scope of practice for registered nurses (RNs), registered practical nurses (RPNs), and nurse practitioners (NPs), which includes 5 regulated acts, including performing a prescribed procedure below the dermis or a mucous membrane [17]. While this authorized act consists of wound assessment and wound care, which are at the entry level, nurses with advanced education, such as NPs, have an extended scope of practice and are authorized to diagnose, prescribe medications, and other treatments for clients [17]. Similarly, British Columbia has specified the pacified authorities at entry level and advanced level, which is presented in the Wound Care Canada by Freeman et al [18]. In addition, the British Columbia Provincial Nursing Skin & Wound Committee, in collaboration with the wound care clinicians from across all health authorities, has devised the wound assessment and treatment flow sheet: documentation guide (version 2, 2018) that specifies nurses' roles with regard to wound assessment and documentation for varying types of wounds [19].

Traditional methods for tracking wound healing progress, such as using a paper ruler and cotton swab for wound assessment, result in high clinical variability and a 44% error rate in infection

XSL•FO
RenderX

identification [20]. Furthermore, health care education and wound care tools have traditionally been developed with lighter skin tones in mind, leading to inaccurate assessment for patients with darker skin [21]. Additionally, rural communities in Canada are diverse, not only in terms of geographic characteristics such as population size, density, or distance from urban centers, but also in social aspects, including social representation and resource availability [18,22,23]. There is growing concern regarding the shortage of health care providers in remote and rural regions, further intensifying the issue of care access for older adults. The Canadian Institute for Health Information reports on the declining trends of health professionals in rural or remote areas, including for nurses, NPs, and pharmacists over a decade, and stagnant growth for rural family physicians [2]. Reducing health disparities, particularly in rural and remote areas with finite resources, requires innovative solutions, including consultation models and enhanced access to best practice care [24]. These dual drivers, population aging and uneven resource distribution, underscore the urgent need for effective health care solutions that:

1. Enable individuals to age in place while living with aging-related health conditions, including chronic and complex wounds.
2. Support the provision of best-practice clinical care in areas where specialized health care human resources are limited.
3. Reduce the challenges associated with human error and biases.

To meet these needs, innovative technologies such as AI offer promising solutions that can enhance efficiency, access, and quality in health care delivery.

AI-Enhanced Wound Care as a Solution

AI has the potential to transform medical practices and service delivery in health care, including enhancing productivity, improving patients' flow, experience, care, and quality of life [25] and enhancing providers' experiences and safety [25,26]. The positive impacts of AI in clinical care include enhanced accuracy in diagnostics, improved patient engagement, and increased treatment support [25]. AI has been leveraged in wound care, providing more efficient and effective solutions for acute and chronic wounds [27]. AI-driven algorithms have proven effective in measuring wound dimensions and identifying prognostic features, such as tissue composition, granulation, slough, eschar, and exudate. These elements are crucial for assessing wound burden and predicting the healing trajectory [28].

Our Implementation Experience: Project and Team

To illustrate how AI-augmented wound care can enhance equity, access, and efficiency, we draw on our collective experiences implementing a DWCT in 2 networked Canadian community health systems. The following section describes the technology itself, the project sites, and the multidisciplinary team involved in its adoption, highlighting both practical considerations and lessons learned in real-world settings.

The Digital Wound Care Technology

The DWCT under implementation, Swift Skin and Wound (Swift), is developed by Swift Medical Inc. and is deployed in more than 5200 health care facilities internationally, spanning the continuum of care. Swift Skin and Wound is a software-based, noninvasive digital wound assessment application installed on smart devices. The software captures clear imaging enhanced by HealX (Swift Medical Inc), a Health Canada and US Food and Drug Administration–registered fiducial marker for real-time calibration of wound images on smart devices, using AI-driven algorithms that auto-trace wound edges, precisely measure wound surface area, and facilitate comprehensive documentation [29]. These features offer advanced analytical and tracking insights to clinicians, empowering them with data-driven assessments to support wound care management [28,29]. This DWCT also enhances communication among the interdisciplinary care team through the centralized dashboard portal, which provides a comprehensive, real-time view to support coordinated care management. The dashboard enables the care team to review patient information, track progress, and provide recommendations, facilitating coordination of care.

Recently, Swift Skin and Wound added additional features: AutoDepth, SmartTissue, and HealingIndex. AutoDepth automatically calculates the visible depth of wounds, measures the wound, and accurately identifies and records its deepest point on a smart device. SmartTissue detects and quantifies tissue types including epithelial, granulation, slough, and eschar within the wound bed, irrespective of skin tone, using deep learning and vision architectures designed to run locally on a wide range of smart devices [30]. HealingIndex uses deep learning and machine learning algorithms to analyze a range of wound characteristics, including wound size, tissue composition, wound type, location, and wound exudates' type and amount to predict healing trajectories of wounds [28]. By using these sophisticated AI-driven algorithms, providers can achieve rapid and precise assessments of wounds and their characteristics and use predictive modeling to project healing trajectories and inform clinical decision-making for care plans (Figure 1).

XSL•FO
RenderX

Figure 1. A sample wound on the ankle area, showing a clinical presentation where objective and consistent wound assessment, supported by an artificial intelligence–driven tool for measurement and tissue classification, can enhance effective monitoring, especially in settings with limited access to specialized expertise.



Evidence from implementations, mostly in US health systems, indicates that integrating Swift's DWCT into wound care programs can improve clinical outcomes [31,32], enhance operational efficiency [29,32], reduce costs [31-33], and increase clinician satisfaction [34].

While other digital wound care technologies exist, this viewpoint focuses on Swift's DWCT because the authorship team is uniquely positioned as active implementation partners in Canadian health systems. We draw on the published literature to provide context and evidence, but we triangulate these insights with our direct, real-world experience to highlight equity, access, and efficiency considerations that emerge in practice, providing insights that would be difficult to capture from literature alone.

Project Description

This project uses an implementation research approach to evaluate the adoption of the DWCT in 2 networked Canadian health systems, aiming to identify facilitators, barriers, and lessons for broader AI-driven wound care implementation.

Implementation Site 1: Giishkaandago'Ikwe Health Services

Giishkaandago'Ikwe Health Services (formerly Fort Frances Tribal Area Health Services) is an Indigenous-led health organization in Northwestern Ontario. The accredited organization provides holistic health services to the 10 Anishinaabe Communities of Southern Treaty 3. The Home and Community Care Program offers essential services in the home and community setting, including advanced wound care through a nurse-led DWCT tool.

Giishkaandago'Ikwe initiated implementation of the DWCT in February 2022 as a quality improvement initiative. This site has particular expertise in leveraging technology to provide care across remote geographic regions. Given the high amputation rates in the region, the home and community care team implemented the DWCT to strengthen decision-making for wound management through efficient documentation and communication, with the goal of enhancing healing rates and reducing amputations.

The interdisciplinary wound care team includes RNs, RPNs, including 2 Skin Wellness associate nurse (SWAN)–trained RPNs, personal support workers (PSWs), and a physiotherapist with advanced wound care training (Masters of Clinical Science in wound healing). In total, the team includes 14 wound care clinicians, supported by a wound care champion and a home and community care coordinator. Foot care nurses conduct wound assessments and communicate findings to the home care team, while other community health nurses contribute through public health programming and wound care support.

Team members complete an 8-hour online wound care training upon hire if they do not already have advanced wound care education. In addition, the team participates in an annual, full-day, hands-on internal wound care workshop to reinforce clinical skills. Weekly SWIFT rounds provide structured opportunities for case review, collaborative problem-solving, and ongoing knowledge exchange.

Implementation of the DWCT was completed over 1 month, supported by training and mentorship from both SWIFT and the site's internal champion (the home and community care coordinator). Initial training consisted of a 1-hour virtual session, followed by ongoing support throughout the implementation period. Due to the team's small size, strong collaboration, and

XSL•FO
RenderX

existing wound care expertise, adoption of the technology was rapid and well integrated into practice. The AutoDepth feature has been actively used in patient care, and team members completed SmartTissue training in January 2025.

Implementation Site 2: Brightshores Health System (Brightshores)

Brightshores Health System is a comprehensive health care network serving the rural Gray and Bruce counties in Ontario, Canada. It operates 6 hospitals (Owen Sound, Meaford, Markdale, Southampton, Wiarton, and Lion's Head), each with 24/7 emergency departments. Brightshores services encompass medical and surgical services, including complex surgeries, total joint replacements, cancer treatments, advanced diagnostic imaging such as magnetic resonance imaging and computed tomography scans, and wound care. Across the system, over 400,000 visits per year are reported, and approximately 2200 staff members with over 250 dedicated physicians deliver patient care across the region, including wound care.

Implementation of the DWCT began in 2024, initially within the regional diabetic foot ulcer clinic. Within months, 35 clinical staff were trained as technology champions, and by October 2024, the DWCT was rolled out to multiple departments across all 6 hospitals. Initial training was performed in-person (2.5 h) with an asynchronous online portion (1.5 h) facilitated by the research team and a nurse specialized in wound, ostomy, and continence (NSWOC), with a focus on use of the AI technology, clinical data entry education, and workflow processes. Users included RNs, RPNs, clinical coordinators, NPs, nurse clinicians, and PSWs. Within 6 months, it was noted that the turnover of staff necessitated continuous training opportunities. An online training module has been developed specifically for the Brightshores sites, which will be more accessible and enable quicker onboarding for staff across multiple sites. Deployment of SmartTissue began in December 2024 and continues across sites.

Discussion

AI-Powered Digital Wound Care Could Improve Access, Enhance Efficiency, and Address Inequities

Overview

Wound care is a critical component of comprehensive home health services. The prevalence of older adults as recipients of home care is notable, with a substantial proportion necessitating wound care due to conditions such as pressure injuries, postsurgical recovery, and complications arising from venous ulcers. Older adults prefer to receive wound care in community settings [35], citing pain, transportation barriers, and risk of exposure to infection in the clinic as primary drivers of preference [36]. Patients and family caregivers frequently report feeling inadequately prepared to participate effectively and confidently in wound care [37]. The presence of nonhealing wounds and associated complications can profoundly affect quality of life, increasing risk of infections and hospitalization [38].

Access

In the Canadian context, the potential for an AI-enhanced DWCT to increase access to best practice care is particularly significant. The "Caring for Canadians" study highlights a significant gap between the supply and demand for HHR in Canada, a gap that is already evident and expected to widen over the next decade, with rural and remote communities feeling the impacts most acutely [39]. Specialist HHR, such as NSWOCs or additional training in wound care, is in high demand. However, recruiting and retaining specialized health care workers is challenging, which has led to the adoption of innovative, technology-enhanced care models to optimize resource use [2,24].

In Canada, geographic location is a primary factor determining access to care, with isolated and remote communities experiencing the highest levels of deprivation [40]. These communities are frequently located in northern regions of the provinces and territories and are often home to Indigenous groups [40]. First Nations, Métis, and Inuit people, along with those of African, Caribbean, Southeast Asian, and Latin descent, have higher prevalence and incidence of diabetes and higher rates of lower extremity amputation because of diabetic foot ulcers and peripheral arterial disease [41]. This is particularly evident in Indigenous populations, where the rate of major amputations due to diabetic wounds is up to 49 times higher than in non-Indigenous populations [42]. This may be attributable to poor access to health care, barriers to clear communication, cultural differences, and discrimination [43]. Within this context, the AI-augmented DWCT offers an opportunity to increase access to best practice care by leveraging available resources and supporting available staff to work at the frontiers of their scope of practice (Textbox 1).

XSL•FO
RenderX

Textbox 1. Case study: Giishkaandago'Ikwe Health Services.

The team at Giishkaandago'Ikwe Health Services uses digital wound care technology (DWCT) for all aspects of wound assessments. Leveraging this technology has strengthened communication internally and externally with care partners. Enhanced communication is central to improved outcomes, illustrated here through a client story.

A client was referred by the Emergency Department after seeking care for painful, reddened foot wounds that had worsened following several weeks of self-treatment. The team initiated a comprehensive wound assessment completed by a personal support worker (PSW) using the SWIFT app. PSWs are unregulated care providers who assist individuals with activities of daily living (ADLs) and instrumental activities of daily living (IADLs).

A holistic treatment plan was then created by the interdisciplinary team during SWIFT rounds. Giishkaandago'Ikwe Health Services mobilized available staff to provide care, ensuring they were supported to work to their full scope of practice. The client observed weekly progress through the app, which increased their engagement and adherence to the care plan. The team also used dashboard-generated reports to share wound progress with external health care providers.

Through this collaboration, the client's care was escalated appropriately, enabling timely access to vascular surgery and chiropody for specialized offloading. The client received these services in a tertiary center and was able to return home promptly, with the Giishkaandago'Ikwe team providing ongoing follow-up.

Despite a late start to wound care and the need for surgical intervention, the client's wounds fully healed within 5 months—an excellent outcome for a case at high risk of amputation. Importantly, the majority of the client's care occurred within their home community, minimizing disruption to their family and caregiving responsibilities. The ability to coordinate and communicate effectively across providers using DWCT was key to avoiding prolonged hospitalization and ensuring continuity of care close to home.

In the Giishkaandago'Ikwe Health Services case study, the DWCT enabled a PSW who had direct contact with the patient to complete a wound care assessment even though this typically would be conducted by nursing staff. The DWCT tool enabled the team member who was geographically co-located to effectively capture the data needed for the wound care team to create a full care plan, without requiring the patient or the specialist providers to travel. In such scenarios, AI-augmented DWCT can support an available staff member to work at the frontiers of the scope of practice while maintaining direct connections for clinical escalation and consultation when needed.

This is particularly relevant in the Canadian rural and remote context, where geography and inclement weather can act as a barrier to access. A 2020 review of barriers to equitable access identified the lack of year-round roads, extreme winter weather events, and the absence of rural and remote hospital staffing as key contributors to poor access to care in Indigenous communities [44]. That same review identified telehealth and telehealth-enabled consultation models as a mitigator of inequitable access. Furthermore, the review noted that telehealth enables community health workers and nurses to provide additional services in their local community and reduces the burden on patients to travel long distances to access care. This has important implications for equitable access, as community care providers typically have higher levels of trust and greater ability to provide culturally safe care [45].

The DWCT augments traditional "telehealth" in wound care by adding empirical data in the form of high-quality, time-lapsed photos in a "dashboard" that can be accessed by patients, caregivers, and the entire wound care team. This enables tracking of wound trajectory and gives contextual clinical data to geographically remote specialist providers such as NSWOCS, endocrinologists, and vascular surgeons. Additionally, the dashboard supports team communication as it offers a shared point of reference to facilitate collaborative care planning. In the context of an optimized scope of practice, this feature provides an additional layer of protection—ensuring specialized clinical oversight of geographically remote patients being cared for by practitioners working at the frontier of scope of practice. Furthermore, this supports staff who may not have clinical training in wound care by providing avenues for consultation and escalation to complement their clinical judgment. Ultimately, this benefits patients who receive care from a trusted provider in their community, while retaining the benefit of review and consultation from clinical experts.

The AI-augmented features of the DWCT expand access to best practice care by (1) enabling patients in remote communities to access services without necessitating travel, (2) empowering clinicians practicing in rural and remote communities to work with an optimized scope of practice, and (3) supporting communication and ensuring all members of the care team have access to relevant clinical information.

Efficiency

Optimization of clinical resources is required to ensure efficiency and maximize the number of patients who receive best practice care. The AI-enhanced DWCT tool can increase clinical efficiency by conducting rapid assessment of wounds and streamlining consultations and escalations [28]. A key component of effective wound care is assessment of wound size. Traditionally, this is done with paper rulers, and the wound size is calculated using a formula that multiplies length by width. However, this traditional method is ill-suited to irregularly shaped wounds, leading to overestimation of size in 36% - 74% of nonrectangular wounds [46-48]. Patients with darker skin tones, wounds characterized by diffuse edges, irregular shapes, necrotic tissue, or unhealthy surrounding tissue are more prone to having their wounds overestimated in size [46]. A more accurate calculation of wound size can be captured using a tracing method, where transparency film is placed over the wound and outlined, but this is reported to be time-consuming [49]. Accurate and objective measurement of wounds is essential as a change in wound size is a key predictor of wound healing trajectory [5]. Digital wound measurements powered by AI are emerging as a more efficient and accurate means to measure wounds [46,50-52] (Textbox 2).

XSL•FO
RenderX

Textbox 2. Case Study: Brightshores Health System.

Brightshores Health System is spread throughout 8600 square kilometers in Southwestern Ontario. Residents are among the oldest average age in the province, with the fastest-growing segment of the population aged between 65 and 84 years. The region has a significant population of patients with chronic disease and wounds, but Brightshores is limited to 1.4 Nurses Specialized in Wound, Ostomy, and Continence (NSWOC's) for the 6 hospital sites.

This means that daily, multiple wounds can present at any one of Brightshores hospitals, and the reality of limited resources means limited and delayed response time. Furthermore, without access to specialists and with limited tools, frontline nursing staff struggle to accurately assess wounds. The DWCT has presented an opportunity for frontline staff to take photos of wounds, allowing for accurate and detailed assessments that can be digitally accessed by the NSWOC and all providers within the circle of care. Providing frontline staff at all 6 hospital sites with a tool to accurately assess wounds allowed for a reduction in reliance upon the NSWOC.

Within the first year, over 1000 km of NSWOC travel time has been saved, allowing them to devote more time to direct patient care. Since the implementation of this DWCT, initial trends show a promising decline in amputation rates and fewer visits of outpatient wound care patients attending the emergency departments due to wound complications, compared to the previous year. Patients are reporting that their assessments performed with the DWCT are more comfortable than traditional methods and are reporting that the use of DWCT has improved their wound care experience.

In one patient's story, the patient felt discouraged about the lack of visual improvement they were able to detect with their eyes. With the use of the DWCT and historical measurements, the patient was able to monitor progressive wound closure, aligning with compliance to treatment recommendations, and was happy to experience a full wound closure within a month of using Swift.

Brightshores' continued hope is that the use of this technology will empower nurses to gain confidence and become more involved in wound assessments, enabling prioritization of wound care, and that this technology will enable patients to see progression in healing over time.

This case study highlights the benefits of tools that enable the clinical care team to optimize their time and streamline processes for escalating concerns to specialized care resources. Use of the DWCT is faster and clinically equivalent in accuracy compared to manual methods of wound measurement and capture [29,53]. Enhanced DWCT reduces the overall time required for wound measurement and documentation and has been shown to result in an average time saving of 1.01 - 2.39 minutes per wound assessment when compared to the traditional manual methods [29]. Given the known constraints on clinicians' time, burden of documentation, and anticipated workforce shortages, these time savings represent an opportunity to increase both efficiency and access.

AI tools can also help identify and escalate wounds that require review. HealingIndex is a new feature of the DWCT that uses machine learning algorithms to assess the photographs of the wounds that are submitted by clinicians. This tool analyzes the photo to identify wound characteristics associated with healing including wound size, tissue composition, and exudates, to predict the healing trajectories of the wound. The HealingIndex AI Model uses deep learning techniques to examine patient wound records, images, and characteristics, producing a score from the AI model that indicates the level of risk for delayed healing and potential deterioration [28].

While traditional care pathways rely on clinicians identifying and escalating wounds that are deteriorating or at risk of deteriorating, the HealingIndex automates these processes using a predictive algorithm. This allows the wound care team to rapidly identify and prioritize wounds that require specialist input. This identification reduces the risk of deterioration and complications that necessitate hospitalization or amputation, reducing overall health care costs associated with wound complications [1,31,38]. Furthermore, the use of this AI tool can identify wounds at risk for slow healing, which may require an adapted wound management protocol. By reducing complications and adapting the treatment of slow-healing wounds, the AI tool can reduce the overall costs associated with wound care [31]. Early identification of wound deterioration is crucial as it allows timely intervention, preventing progression and potentially leading to significant savings for the health care system through avoidance of unnecessary costs [34].

The use of the predictive algorithm may also augment clinical decision-making by identifying wounds that are healing effectively and therefore may not need daily or thrice-weekly dressing changes. Reducing the frequency of dressing changes reduces supply costs and frees patient and provider time. Reducing the frequency of dressing changes to once a week, where clinically appropriate, can free up thousands of hours of nursing time over the course of a year [54]. Overall, the AI-powered DWCT can enhance clinical efficiency by automating escalation and review processes.

Additionally, the DWCT platform provides opportunities for asynchronous collaboration. Specialists are able to review patient records, including high-resolution time-lapse photos, and provide recommendations without requiring geographic co-location or congruent timing. This enables the specialist wound care nursing team to optimize their time by reducing travel between sites, reducing administrative burden, and creating care plans that can be executed by staff located near the patient. This offers the benefits of not only efficiency but also equity, as access to best practice care and consultation is no longer limited to the fortunate few who can be reviewed by the specialist nurses in person. In rural and remote parts of Canada, which, as noted above, are struggling to recruit and retain skilled providers, these tools enable more patients to access specialist consultation and review without necessitating travel.

AI-powered DWCT can support increased efficiency by (1) reducing time spent on manual assessment and measurement of wounds; (2) automating escalation by flagging deteriorating wounds for review, which enables prioritization of at-risk patients for intervention; and (3) enabling collaboration between frontline staff and specialist wound care resources without requiring geographic colocation or time alignment.

XSL•FO
RenderX

Equity

Universality and accessibility are embedded within Canada's Health Act, and equity must be centered as a priority. Inequities in the Canadian health system exist, especially related to disparities in health outcomes for rural and remote populations, and disproportionately affect First Nations, Inuit, and Métis communities who face burdens associated with travel and costs [44]. Rural health care policies are often guided by urban care models, which can exacerbate these inequities for rural communities [24].

Many rural, remote, and Indigenous communities experience inequitable access and poor outcomes, which may be attributable in part to resource shortages. Additionally, human biases, racism, colonial legacies, and historically unbalanced relationships between service users and providers also play a key role. Nyugen et al [44] explicitly state "Colonization and historical intergenerational traumas... have plagued the survivors and later generations with physical and mental trauma," and "...many Health Care providers do not acknowledge the impact of colonization on the Indigenous community and disregard the social determinants of health as explanations for illness." This, in turn, manifests as culturally inappropriate health services, which deter people from seeking health care.

In this context, the AI-enhanced DWCT can support more equitable care by enabling community care providers with existing, trusting relationships to provide culturally safe care. This approach is aligned with mitigation strategies proposed by Indigenous health experts who highlight the importance of community partnerships and provision of culturally safe care [44,45].

Furthermore, patients with darker skin tones disproportionately experience poor wound outcomes, which may be attributable to both inequities in social determinants of health and limitations in provider ability to accurately identify clinically relevant changes in tissue in pigmented skin [55]. Johnson et al [55] note that erythema is an excellent example, as its presentation varies based on skin tone, and historically, medical education has failed to account for diverse skin tones in training and published literature [21]. Darker skin tone is associated with overestimation of wound size when using standard "width X length" measurements [46]. In comparison, the DWCT with AI-augmented AutoTrace, AutoDepth, and SmartTissue is able to provide accurate measurements and can enable color calibration to adjust for light conditions and skin pigmentation [46]. Thus, the use of the AI-enhanced DWTC can support more accurate measurements and potentially overcome some of the implicit biases that health care providers manifest.

Next Steps

While AI-enhanced tools offer immense promise to support better care efficiency and outcomes, there are several key considerations that must be addressed to ensure equitable and trustworthy implementation. One of the pressing concerns is the representation of diverse patient populations in AI systems. When datasets fail to capture diversity across factors such as skin tone, age, and other demographic variables, resulting models can produce biased or inaccurate outputs. These biases are particularly problematic in clinical contexts like wound care, where diagnostic accuracy is critical. Diverse groups, such as individuals with darker skin tones or older adults, are often underrepresented in training datasets, leading to biases in AI outcomes. This underrepresentation can result in inaccurate diagnoses or recommendations, which are harmful and counter to the goals of quality wound care. Human biases embedded in AI systems can become enshrined or even amplified if not intentionally designed with equity in mind [25]. Such biases arise when AI models are trained on datasets that do not adequately represent the diversity of the population, including those most likely to experience wounds, such as older adults [56]. Chu et al [56] emphasize the known challenges of ageism in AI, noting that models often overlook older adults due to insufficient data representation. To counteract these biases, it is essential to increase the diversity of AI training datasets and maintain robust evaluation protocols that ensure accuracy and interrater reliability across patient populations. These safeguards are crucial in mitigating the risks of bias and ensuring that AI systems are fair and inclusive.

In addition to addressing bias, overcoming skepticism and implementation barriers is crucial for the successful adoption of AI-enhanced technologies, particularly among older adults. While the stereotype of older adults as technophobic is increasingly being disproven [57], structural barriers such as the cost of technology and access to reliable internet remain significant hurdles for many older adults in rural and remote communities [58]. Despite these challenges, patients are generally open to the use of AI in health care if appropriate regulatory and safety oversights are in place [59]. Pilot studies, such as those by Wang et al [52] and Raismam et al [60], show promising results regarding patient acceptance of AI for wound care [52,60] while Mohammed et al [34] highlight positive provider acceptance [34]. These studies indicate a growing receptiveness toward an AI-augmented DWCT, but continued exploration into unique barriers, especially those pertinent to the Canadian context, is necessary.

The evaluation being conducted by this team will enhance understanding of implementing AI-enhanced DWCT in rural and remote communities in Canada. This evaluation will identify specific barriers and facilitators of adoption, ultimately leading to the creation of a framework that supports wider adoption and integration of technologies to support healthy aging. By addressing these issues, the path toward equitable, efficient, and widespread use of AI technologies in health care becomes clearer, fostering a more inclusive technological future.

Funding

Author BA is funded by a Mitacs Accelerate postdoctoral fellowship.

XSL•FO
RenderX

Conflicts of Interest

Authors HTM and RDJF are current employees of Swift Medical Inc. Other authors have no conflicts to declare.

References

1. Sen CK. Human wound and its burden: updated 2020 compendium of estimates. Adv Wound Care (New Rochelle) 2021 May;10(5):281-292. [doi: 10.1089/wound.2021.0026] [Medline: 33733885]
2. Keeping pace with changing population needs. Canadian Institute for Health Information. 2024. URL: https://www.cihi.ca/en/the-state-of-the-health-workforce-in-canada-2022/keeping-pace-with-changing-population-needs [accessed 2025-03-11]
3. Ud-Din S, Bayat A. Non-invasive objective devices for monitoring the inflammatory, proliferative and remodelling phases of cutaneous wound healing and skin scarring. Exp Dermatol 2016 Aug;25(8):579-585. [doi: 10.1111/exd.13027] [Medline: 27060469]
4. Pillen H, Miller M, Thomas J, Puckridge P, Sandison S, Spark J. Assessment of wound healing: validity, reliability and sensitivity of available instruments. Wound Pract Res 2009;17(4):208-217 [FREE Full text]
5. Tan P, Lim J, Moiemen N. Imaging techniques used for wound healing assessment: a systematic review part 1 chronic wounds. EBJ 2021;2(4):194-214. [doi: 10.3390/ebj2040015] [Medline: 34308485]
6. Shi C, Dumville JC, Juwale H, Moran C, Atkinson R. Evidence assessing the development, evaluation and implementation of digital health technologies in wound care: a rapid scoping review. J Tissue Viability 2022 Nov;31(4):567-574. [doi: 10.1016/j.jtv.2022.09.006] [Medline: 36192302]
7. Population projections for Canada (2021 to 2068), provinces and territories (2021 to 2043). Statistics Canada. 2022. URL: https://www150.statcan.gc.ca/n1/pub/91-520-x/91-520-x2022001-eng.htm [accessed 2025-03-31]
8. Ageing and health. World Health Organization. 2018. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-04-08]
9. Queen D, Botros M, Harding K. International opinion-the true cost of wounds for Canadians. Int Wound J 2024 Jan;21(1) [FREE Full text] [doi: 10.1111/iwj.14522] [Medline: 38084491]
10. National health expenditure trends, 1975 to 2019. Canadian Institute for Health Information. 2019. URL: https://www.cihi.ca/sites/default/files/documents/nhex-trends-narriative-report-2019-en.pdf [accessed 2026-03-06]
11. Four Pillars: Recommendations for Achieving a High Performing Health System. 2011. URL: https://nhh.ca/document/four-pillars-recommendations-for-achieving-a-high-performing-health-system-pdf [accessed 2024-02-10]
12. Dandy K, Bollman RD. Seniors in rural Canada. Stat Can 2008;7(8) [FREE Full text]
13. Dhalla IA, Tepper J. Improving the quality of health care in Canada. CMAJ 2018 Oct 1;190(39):E1162-E1167. [doi: 10.1503/cmaj.171045] [Medline: 30274993]
14. Gouin JP, Kiecolt-Glaser JK. The impact of psychological stress on wound healing: methods and mechanisms. Immunol Allergy Clin North Am 2011 Feb;31(1):81-93. [doi: 10.1016/j.iac.2010.09.010] [Medline: 21094925]
15. Laforet K, Allen JO, McIsaac C. Evidence-based wound care, home care perspective. The Canadian Home Care Association (CHCA). 2012. URL: https://cdnhomecare.ca/wp-content/uploads/2020/03/Wound_Management_-_Home_Care_Perspective_-_English_2012.pdf [accessed 2026-03-06]
16. Ousey K, Stephenson J, Cook L, Kinsey L, Batt S. Final year student nurses' experiences of wound care: an evaluation. Br J Community Nurs 2013 Mar;Suppl(Sup3):S7-S8. [doi: 10.12968/bjcn.2013.18.sup3.s7] [Medline: 23682497]
17. Practice standard- scope of practice. : College of Nurses- Ontario; 2025 URL: https://cno.org/assets/cno/documents/standard-and-learning/practice-standards/49041-scope-of-practice.pdf [accessed 2026-03-13]
18. Freeman S, Sargent MJ, Galarza LR, Rossnagel E. The use of AI informed technology for wound care progress tracking and data analysis: rural and northern perspectives. WCC 2025;23(1). [doi: 10.56885/124790twwpdj]
19. Wound assessment & treatment flow sheet: documentation guide. CLWK (Connecting Learners with Knowledge). URL: https://www.clwk.ca/get-resource/wound-assessment-documentation-guide-landscape-2pg/ [accessed 2025-09-25]
20. Keast DH, Bowering CK, Evans AW, Mackean GL, Burrows C, D'Souza L. Measure: a proposed assessment framework for developing best practice recommendations for wound assessment. Wound Repair Regen 2004;12(3 Suppl):S1-17. [doi: 10.1111/j.1067-1927.2004.0123S1.x] [Medline: 15230830]
21. Mukwende M, Tamony P, Turner M. Mind the Gap: A Handbook of Clinical Signs in Black and Brown Skin: St George's, University of London; 2020. [doi: 10.24376/rd.sgul.12769988.v1]
22. Laurin F, Pronovost S, Carrier M. The end of the urban-rural dichotomy? Towards a new regional typology for SME performance. J Rural Stud 2020 Dec;80:53-75. [doi: 10.1016/j.jrurstud.2020.07.009]
23. Keating N, Phillips J. A critical human ecology perspective on rural ageing. In: Keating N, editor. Rural Ageing: A Good Place to Grow Old? Ageing and the Lifecourse Series: Bristol University Press, Vol. 2008:1-10. [doi: 10.46692/9781847424037]
24. Wilson CR, Rourke J, Oandasan IF, Bosco C. Progress made on access to rural healthcare in Canada. Can J Rural Med 2020;25(1):14 [FREE Full text] [doi: 10.4103/CJRM.CJRM_84_19]
25. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ 2023 Sep 22;23(1):689. [doi: 10.1186/s12909-023-04698-z] [Medline: 37740191]

XSL•FO
RenderX

26. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021 Jul;8(2):e188-e194. [doi: 10.7861/fhj.2021-0095] [Medline: 34286183]
27. Encarnação R, Manuel T, Palheira H, Neves-Amado J, Alves P. Artificial intelligence in wound care education: protocol for a scoping review. Nurs Rep 2024 Mar 14;14(1):627-640. [doi: 10.3390/nursrep14010048] [Medline: 38535720]
28. Gupta R, Goldstone L, Eisen S, et al. Towards an AI-based objective prognostic model for quantifying wound healing. IEEE J Biomed Health Inform 2024 Feb;28(2):666-677. [doi: 10.1109/JBHI.2023.3251901] [Medline: 37028088]
29. Mohammed HT, Bartlett RL, Babb D, Fraser RDJ, Mannion D. A time motion study of manual versus artificial intelligence methods for wound assessment. PLoS ONE 2022;17(7):e0271742. [doi: 10.1371/journal.pone.0271742] [Medline: 35901189]
30. Ramachandram D, Ramirez-GarciaLuna JL, Fraser RDJ, Martínez-Jiménez MA, Arriaga-Caballero JE, Allport J. Fully automated wound tissue segmentation using deep learning on mobile devices: cohort study. JMIR Mhealth Uhealth 2022 Apr 22;10(4):e36977. [doi: 10.2196/36977] [Medline: 35451982]
31. Mohammed HT, Mannion D, Cassata A, Fraser RD. Trends in pressure injury prevalence rates and average days to healing associated with adoption of a comprehensive wound care program and technology in skilled nursing facilities in the United States. Wounds 2024 Jan;36(1):23-33. [doi: 10.25270/wnds/23089] [Medline: 38417821]
32. Mohammed HT, Corcoran K, Lavergne K, et al. Clinical, operational, and economic benefits of a digitally enabled wound care program in home health: quasi-experimental, pre-post comparative study. JMIR Nurs 2025 Apr 8;8:e71535. [doi: 10.2196/71535] [Medline: 40198913]
33. Mohammad HT, Peterson L, Cassata A, Mitchell J, Shippy C, Robert DJ. Evaluating the efficacy and financial impacts of integrating AI-powered skin and wound solution in post-acute care skilled nursing facilities over a three-year period. Swift. 2024. URL: https://swiftmedical.com/efficacy-financial-impacts-ai-powered-wound-solution-post-acute-care/ [accessed 2025-02-17]
34. Mohammed HT, Cassata A, Fraser RDJ, Mannion D. Implementing technology in practice: factors associated with clinicians' satisfaction with an AI wound assessment solution. Int J Nurs Health Care Res 2023 Jul 28. [doi: 10.29011/2688-9501.101452]
35. Rogers LC, Armstrong DG, Capotorto J, et al. Wound center without walls: the new model of providing care during the COVID-19 pandemic. Wounds 2020 Jul;32(7):178-185. [Medline: 32335520]
36. Huang Y, Hu J, Xie T, et al. Effects of home-based chronic wound care training for patients and caregivers: a systematic review. Int Wound J 2023 Nov;20(9):3802-3820. [doi: 10.1111/iwj.14219] [Medline: 37277908]
37. Moore Z, Kapp S, Loney A, et al. A tool to promote patient and informal carer involvement for shared wound care. Wounds Int 2021;12(3) [FREE Full text]
38. Sen CK. Human wound and its burden: updated 2022 compendium of estimates. Adv Wound Care (New Rochelle) 2023 Dec;12(12):657-670. [doi: 10.1089/wound.2023.0150] [Medline: 37756368]
39. Caring for Canadians: Canada's future health workforce – the Canadian health workforce education, training and distribution study. Health Canada. 2025 Jan 30. URL: https://www.canada.ca/en/health-canada/services/health-care-system/health-human-resources/workforce-education-training-distribution-study.html [accessed 2025-03-28]
40. Richardson KL, Driedger MS, Pizzi NJ, Wu J, Moghadas SM. Indigenous populations health protection: a Canadian perspective. BMC Public Health 2012 Dec 20;12(1):1098. [doi: 10.1186/1471-2458-12-1098] [Medline: 23256553]
41. Blanchette V, Kayssi A. Sex, gender, race and ethnicity matter in limb preservation in North America. LPJ 2022;3(1):30-36. [doi: 10.56885/FDBN2892]
42. Blanchette V, Patry J, Brousseau-Foley M, et al. Diabetic foot complications among Indigenous peoples in Canada: a scoping review through the PROGRESS-PLUS equity lens. Front Endocrinol (Lausanne) 2023;14:1177020. [doi: 10.3389/fendo.2023.1177020] [Medline: 37645408]
43. Bernatchez J, Mayo A, Kayssi A. The epidemiology of lower extremity amputations, strategies for amputation prevention, and the importance of patient-centered care. Semin Vasc Surg 2021 Mar;34(1):54-58. [doi: 10.1053/j.semvascsurg.2021.02.011] [Medline: 33757636]
44. Nguyen NH, Subhan FB, Williams K, Chan CB. Barriers and mitigating strategies to healthcare access in indigenous communities of Canada: a narrative review. Healthcare (Basel) 2020 Apr 26;8(2):112. [doi: 10.3390/healthcare8020112] [Medline: 32357396]
45. Peiris D, Brown A, Cass A. Addressing inequities in access to quality health care for indigenous people. CMAJ 2008 Nov 4;179(10):985-986. [doi: 10.1503/cmaj.081445] [Medline: 18981431]
46. Alonso MC, Mohammed HT, Fraser RD, Ramirez Garcia Luna JL, Mannion D. Comparison of wound surface area measurements obtained using clinically validated artificial intelligence-based technology versus manual methods and the effect of measurement method on debridement code reimbursement cost. Wounds 2023 Oct;35(10):E330-E338. [doi: 10.25270/wnds/23031] [Medline: 37956346]
47. Langemo D, Anderson J, Hanson D, Hunter S, Thompson P. Measuring wound length, width, and area: which technique? Adv Skin Wound Care 2008 Jan;21(1):42-45. [doi: 10.1097/01.ASW.0000284967.69863.2f] [Medline: 18156829]
48. Rogers LC, Bevilacqua NJ, Armstrong DG, Andros G. Digital planimetry results in more accurate wound measurements: a comparison to standard ruler measurements. J Diabetes Sci Technol 2010 Jul 1;4(4):799-802. [doi: 10.1177/193229681000400405] [Medline: 20663440]

XSL•FO
RenderX

49. Oien RF, Håkansson A, Hansen BU, Bjellerup M. Measuring the size of ulcers by planimetry: a useful method in the clinical setting. J Wound Care 2002 May;11(5):165-168. [doi: 10.12968/jowc.2002.11.5.26399] [Medline: 12055939]
50. Bowling FL, Paterson J, Ndip A. Applying 21st century imaging technology to wound healing: an Avant-Gardist approach. J Diabetes Sci Technol 2013 Sep 1;7(5):1190-1194. [doi: 10.1177/193229681300700536] [Medline: 24124945]
51. Foltynski P, Ciechanowska A, Ladyzynski P. Wound surface area measurement methods. Biocybern Biomed Eng 2021 Oct;41(4):1454-1465. [doi: 10.1016/j.bbe.2021.04.011]
52. Wang SC, Anderson JAE, Evans R, et al. Point-of-care wound visioning technology: reproducibility and accuracy of a wound measurement app. PLOS ONE 2017;12(8):e0183139. [doi: 10.1371/journal.pone.0183139] [Medline: 28817649]
53. Au Y, Beland B, Anderson JAE, Sasseville D, Wang SC. Time-saving comparison of wound measurement between the ruler method and the swift skin and wound app. J Cutan Med Surg 2019;23(2):226-228. [doi: 10.1177/1203475418800942] [Medline: 30841726]
54. Lindholm C, Searle R. Wound management for the 21st century: combining effectiveness and efficiency. Int Wound J 2016 Jul;13 Suppl 2(Suppl 2):5-15. [doi: 10.1111/iwj.12623] [Medline: 27460943]
55. Johnson J, Johnson AR, Andersen CA, Kelso MR, Oropallo AR, Serena TE. Skin pigmentation impacts the clinical diagnosis of wound infection: imaging of bacterial burden to overcome diagnostic limitations. J Racial Ethn Health Disparities 2024 Apr;11(2):1045-1055. [doi: 10.1007/s40615-023-01584-8] [Medline: 37039975]
56. Chu CH, Nyrup R, Leslie K, et al. Digital ageism: challenges and opportunities in artificial intelligence for older adults. Gerontologist 2022 Aug 12;62(7):947-955. [doi: 10.1093/geront/gnab167] [Medline: 35048111]
57. Sixsmith A, Horst BR, Simeonov D, Mihailidis A. Older people's use of digital technology during the COVID-19 pandemic. Bull Sci Technol Soc 2022 Jun;42(1-2):19-24. [doi: 10.1177/02704676221094731] [Medline: 38603230]
58. Freeman S, Pelletier C, Ward K, et al. Factors influencing participation in physical activity for persons living with dementia in rural and northern communities in Canada: a qualitative study. BMJ Open 2022 Jun 16;12(6):e060860. [doi: 10.1136/bmjopen-2022-060860] [Medline: 35710242]
59. Esmaeilzadeh P. Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. BMC Med Inform Decis Mak 2020 Jul 22;20(1):170. [doi: 10.1186/s12911-020-01191-1] [Medline: 32698869]
60. Raizmam R, Ramírez-GarciaLuna JL, Newaz T, et al. Empowering patients and caregivers to use AI and computer vision for wound monitoring: a feasibility study. JMIR Preprints. Preprint posted online on Dec 1, 2024. [doi: 10.2196/preprints.69470]

Abbreviations

AI: artificial intelligence
DWCT: digital wound care technology
HHR: health human resource
NP: nurse practitioner
NSWOC: Nurses Specialized in Wound, Ostomy, and Continence
PAST: Pressure Sore Status Tool
PSW: personal support worker
PUSH: Pressure Ulcer Scale for Healing
RN: registered nurse
RPN: registered practical nurse
SWAN: Skin Wellness associate nurse

Edited by E Borycki; submitted 17.Oct.2025; peer-reviewed by C Milne, S Mohanadas; accepted 13.Jan.2026; published 30.Mar.2026.

Please cite as:
Genge C, Ayaz B, Freeman S, Mohammed HT, Fraser RDJ, Abejirinde IOO, O'Sullivan-Drombolis D, Brookham R
Artificial Intelligence–Enhanced Wound Care to Improve Access, Efficacy, and Equity in Wound Care for Older Adults in Rural and Remote Regions of Canada
JMIR Nursing 2026;9:e85644
URL: https://nursing.jmir.org/2026/1/e85644
doi:10.2196/85644

© Courtney Genge, Basnama Ayaz, Shannon Freeman, Heba Tallah Mohammed, Robert D J Fraser, Ibukun-Oluwa Omolade Abejirinde, Deirdre O'Sullivan-Drombolis, Rebecca Brookham. Originally published in JMIR Nursing (https://nursing.jmir.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Acceptance of Digital Technology Among Nursing Staff in Geriatric Long-Term Care: Systematic Review

Jeton Iseni[1], MSc; Walter Swoboda[2], Dr Prof; Daniel Houben[3], Dr Prof; Roman Hilla[4]

[1]DIWAG, Health Management, University of Applied Sciences Neu-Ulm, Wileystr. 1, Neu-Ulm, Germany

[2]DigiHealth, Health Management, University of Applied Sciences Neu-Ulm, Neu-Ulm, Germany

[3]Faculty of Social Work, University of Applied Sciences Landshut, Landshut, Germany

[4]Business Intelligence, Administration, St. Elisabeth Pflegezentrum Senden, Senden, Germany

Corresponding Author:
Jeton Iseni, MSc
DIWAG, Health Management, University of Applied Sciences Neu-Ulm, Wileystr. 1, Neu-Ulm, Germany

Abstract

Background: Digital technologies are increasingly being introduced into the health care system and in settings such as hospitals and geriatric long-term care (LTC) facilities, offering potential benefits such as improved care quality, reduced workload, or enhanced documentation processes. However, the success of these technologies also depends on the acceptance by the primary users, that is, the nursing staff.

Objective: This review synthesizes empirical studies that have explored the acceptance of digital technologies by nursing staff in geriatric LTC settings, building upon the foundational work by Yu et al (2009). The goal is to identify influencing factors, assess the extent of existing evidence, and highlight research gaps in this care setting.

Methods: A systematic literature review was conducted following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines. The SPIDER (sample, phenomenon of interest, design, evaluation, research type) framework was used for eligibility criteria. Databases searched included PubMed, ACM Digital Library, Web of Science, and the Health Administration Database ProQuest. Studies were included if they empirically examined the acceptance of digital technologies by nursing staff in geriatric LTC settings. Two reviewers independently screened the studies, extracted data, and assessed methodological quality using the CASP (Critical Appraisal Skills Programme) checklist.

Results: A total of 3 studies met the criteria, highlighting a gap in research on this topic. The studies applied cross-sectional quantitative designs and highlighted critical determinants of technology acceptance, including perceived usefulness, ease of use, digital competence, and organizational support. The studies involved a total of 1019 participants from Germany, Australia, and the Netherlands. Barriers included lack of user involvement, lack of training, poor system design, and demographic differences in digital affinity.

Conclusions: This review shows that the acceptance of digital technologies by nursing staff in geriatric LTC settings is shaped by a constellation of individual factors, such as digital competence and perceived relevance of technology, as well as organizational factors such as access to training and involvement of staff in the implementation process. Despite these insights, the limited number of empirical studies highlights a research gap in this care setting. To ensure sustainable digital transformation in geriatric LTC, future research should prioritize rigorous and participatory approaches, using longitudinal, intervention-based, or multilevel study designs.

(JMIR Nursing 2026;9:e82223) doi:10.2196/82223

KEYWORDS

digitalization; elderly care; health information technology; geriatric nurse; long-term care; LTC; nursing; organizational innovation; systematic review; technology acceptance

Introduction

Overview

"A promising approach to understanding social dynamics lies in conceiving our society as a globalized knowledge society undergoing a comprehensive and multifaceted digital transformation" [1]. The adoption of digital technologies in health care and nursing care reflects the complex digital transformation taking place across society [2]. Digital technologies are already having an immense impact on how nursing care is delivered [3-10]. In elderly care settings, particularly in geriatric long-term care (LTC) facilities, digital technologies such as electronic health records, assistive robotic systems, telehealth apps, assistive sensory systems, information

XSL•FO
RenderX

and communication technologies, or artificial intelligence monitoring platforms [4,9,11] offer important opportunities to address current and future challenges [12-15]. These include workforce shortages, improving working conditions, or increasing the attractiveness of the nursing profession. The demographic shift associated with an aging population [16] is also one of the major challenges in this context. In Germany, the number of individuals in need of LTC rose to over 5.7 million people by December 2023 [17], with projections indicating a further increase in this number. In Germany, several programs were initiated for supporting the digital pathway [18,19]. The Bavarian State Chancellery decided in a cabinet meeting on March 19, 2024, to promote digitalization in health care and nursing. The goal is to further improve medical and nursing care for the population [20]. On the other hand, not only is the demand for LTC places increasing, but also the need for nursing staff in general is growing [21].

The real-world implementation of digital innovations in the health care system, especially elderly care, remains inconsistent and is frequently challenging [7,22-25]. One of the most significant challenges is the level of acceptance among nursing staff [26-29]. As the primary users of these technologies and new systems, nursing staff play a crucial role in determining whether such tools will be adopted and integrated into everyday work [15-19]. While research in acute and primary care has increasingly examined digital transformation through staff training, workflow redesign, and implementation frameworks, geriatric LTC remains comparatively underexplored. In acute care settings, digital competence programs and structured IT implementation strategies are often supported by institutional infrastructure [30,31]. Theoretical models such as the technology acceptance model (TAM) [32] or TAM2 [33] highlight that perceived usefulness and perceived ease of use are key predictors of user acceptance [34]. However, practical experience shows that digital transformation, especially in the field of care, often falters at the stage of user engagement, particularly when it fails to consider organizational, cultural, ethical, and educational conditions [2,5,10,35-38]. In geriatric LTC, where staff is more involved in basic care of older adults, these challenges become even more important [14,39-41]. A simple example of how digital technology in geriatric LTC could avoid high risks and time waste of the nursing staff is the occurrence of discrepancies between medication plans sent via fax by general practitioners and the actual administration records in nursing homes. Paper-based updates made during medical visits are sometimes not transferred into the official documentation, creating dangerous information gaps and avoidable risks for residents. This example illustrates how outdated communication practices and the lack of integrated digital infrastructures can compromise care quality and safety. It further highlights the importance of user-accepted digital solutions in daily nursing work and a scientifically grounded framework for implementation in LTC. Geriatric LTC facilities often face limited access to training resources and less technical and managerial support for digital adoption. Consequently, empirical evidence on how nursing staff in LTC acquire digital skills, engage in technology implementation, and perceive organizational support remains scarce. This gap underscores the need for research specifically focusing on acceptance factors, training needs, and contextual barriers unique to geriatric LTC, rather than extrapolating findings from hospital-based studies. Despite the critical role of nursing staff in implementing digital innovations, scientific evidence addressing their perspectives, needs, and acceptance in LTC contexts remains very low [25,42].

Objective

Despite considerable political interest and investments in digital transformation, the success of such efforts in the care setting hinges on a crucial factor that remains underexplored, at least in the geriatric LTC, which is the acceptance of digital technologies by nursing staff. Their perspective is not only relevant but essential to the sustainable implementation of digital solutions in care. The primary objective of this systematic review is to synthesize existing empirical research that investigates the acceptance of digital technologies among nursing staff in geriatric LTC settings, building upon the work of Yu et al [39], which was one of the first studies with focus on acceptance factors among nursing staff in LTC, published in 2009. By identifying the most relevant influencing factors, the review contributes to a better understanding of the conditions under which circumstances digital innovations can be effectively and successfully implemented in geriatric LTC environments, with particular attention to the acceptance factors of the nursing staff in this setting.

Methods

Study Design

This systematic review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines (Checklist 1) [43]. For the development of the eligibility criteria, the SPIDER (sample, phenomenon of interest, design, evaluation, research type) framework [44] was applied to ensure a structured and targeted selection of studies.

Eligibility Criteria

The eligibility criteria and methodological steps were defined a priori; however, no protocol was registered for this review. The inclusion and exclusion criteria were defined in alignment with the SPIDER components (Table 1), focusing for instance on studies involving nursing personnel in LTC (sample); their acceptance of digital technologies (phenomenon of interest); and empirical research designs with quantitative, qualitative, and mixed methods approaches (design and research type). This focus reflects the review's aim to identify scientific evidence on how acceptance shapes digital adoption among LTC nursing staff. The studies had to be peer-reviewed and published in English or German. Exclusion criteria comprised studies conducted only in hospital, outpatient, or home care environments, as well as research focusing on other professional groups without separately analyzing the nursing staff perspective. Although qualitative and mixed methods studies were eligible according to the SPIDER framework, no such studies met all inclusion criteria (ie, focus on nursing staff in geriatric LTC and explicit assessment of technology acceptance). As a result, all included studies employed cross-sectional

XSL•FO
RenderX

quantitative designs. This limitation is discussed in the *Results* and *Discussion* sections, but the inclusion parameters were retained to ensure methodological consistency and comparability across studies.

Table . Inclusion and exclusion criteria—SPIDER (sample, phenomenon of interest, design, evaluation, research type) components.

SPIDER components	Inclusion criteria	Exclusion criteria
S=Sample	• Nursing staff employed in long-term care facilities (nursing homes, elderly care)	• Studies focusing in general on non-nursing staff (eg, administrators, managers) • Studies involving participants who are not working in long-term care facilities • Studies with samples not clearly defined as nursing staff in geriatric long-term care
PI =Phenomenon of interest	• Acceptance, adoption, barriers, experiences related to digital innovations in care settings, including technologies like electronic health records, telehealth services, assistive robotics, digital documentation, sensory, ICT[a], IoT[b], AI[c]-driven decision support systems	• Studies focusing only on nondigitalized operations in long-term care • Studies exclusively addressing competencies and education without looking at technology acceptance • Studies not involving digital technologies
D=Design	• Intervention studies, observational or cross-sectional surveys, studies employing qualitative, mixed methods designs	• Nonresearch
E=Evaluation	• Outcomes related to staff attitudes, perceptions, barriers, willingness to use, fears, and facilitators to adoption, satisfaction, perceived usefulness of digital technologies in long-term care	• Studies not reporting on outcomes related to staff digital technology acceptance • Studies focusing solely on managerial or administrative evaluations without staff input. • Studies focusing only on nursing staff from hospitals or private home care settings
R=Research type	• Qualitative, quantitative, or mixed methods research focusing on the care employees regarding digital innovation adoption • Peer-reviewed journal articles published between January 1, 2010, and December 31, 2024 in English or German	• Conference papers, reviews, editorials, letters to the editor, and studies not published in peer-reviewed journals • Publications not in English or German • Studies published outside the specified date range before January 1, 2010 (except for Yu et al [39])

[a]ICT: information and communication technology.

[b]IoT: internet of things.

[c]AI: artificial intelligence.

Search Strategy

The search strategy employed an inclusive keyword combination, which was discussed and refined beforehand. Boolean operators were used to capture the intersection of acceptance, digitalization, technology, nursing, and geriatric LTC. The primary search string used was as follows: (“acceptance” AND (“digital technology” OR “digital” OR “technological” OR “artificial” OR “robotic” OR “digitalization” OR “artificial intelligence” OR “IoT” OR “robot” OR “virtual reality” OR “socially assistive robots” OR “digital tools” OR “telehealth” OR “Internet of Things” OR “EHR”)) AND (“nursing homes” OR “elderly” OR “geriatric” OR “inpatient home” OR “care facility” OR “nursing facilities” OR “nursing home” OR “aged care” OR “care home” OR “long-term care” OR “senior living center” OR “LTC”). Exact search strings for each database are documented in Multimedia Appendix 1.

The literature search was conducted across PubMed, Web of Science, ProQuest, and the ACM Digital Library. These databases were selected to ensure broad interdisciplinary coverage of nursing, health care, and technology-related research. Gray literature was not searched systematically. However, 1 relevant report identified through manual search [45] was used to provide contextual information for the discussion and was not part of the primary evidence base.

Although specialized databases, such as CINAHL, were not included due to missing license at University of Applied Sciences Neu-Ulm, the chosen databases offer considerable overlap. This limitation and the potential risk of missed studies are acknowledged in the *Discussion* section. To enhance comprehensiveness, the database search was supplemented by citation tracking and manual searches. Searches were limited to the period from January 1, 2010, to December 31, 2024. As noted previously, 1 of the included studies [39] falls outside the formal inclusion window set; however, it was retained based

XSL•FO
RenderX

on discussions among all internal reviewers involved and due to the fact that this study represents the first known empirical study with the focus on the acceptance of digital technology among nursing staff in LTC settings. The identification process is illustrated in the PRISMA 2020 flow diagram, in the *Results* section.

The systematic search was conducted on April 14, 2025, following an initial exploratory search for an overview of the existing literature on October 25, 2024 (Multimedia Appendix 2). The primary researcher (JI) led the systematic review process, including database search, screening, and data extraction. The second reviewer (RH) independently screened the publications and also evaluated them for eligibility. Any discrepancies or critical assessments concerning study relevance, methodological quality, or thematic clarity were discussed in regular virtual meetings with senior reviewers (WS) and (DH). To manage the studies, the open-source software Zotero, version 7.0.11 (64-bit) was used as the reference software.

Study Selection

Study selection was conducted in 2 phases. The first phase was the selection via title and reading the abstract. In the second phase, the full texts of potentially eligible studies were reviewed in detail. Studies that met the inclusion criteria and passed quality checks were included in the synthesis. Excluded studies and reasons for exclusion are presented in Multimedia Appendix 3.

Due to the limited number of eligible studies, a formal sensitivity analysis was not possible. However, the impact of study quality on synthesis outcomes was qualitatively assessed during reviewer discussions.

Data items from included studies were extracted with the following variables using an Excel form:

- Study identification: authors, title, year of publication, country
- Study design: methodological approach
- Participants: number, role (nursing staff, management)
- Aim of study: nature of the digital technology studied
- Key findings: outcome measures, determinants, and facilitators affecting acceptance

Quality Assessment

Methodological quality was assessed using the Critical Appraisal Skills Programme (CASP) [46], applying item-level judgments ("Yes," "Can't tell," "No"). The overall confidence rating was categorized as "low," "moderate," or "high," with no studies falling into the "high" category. Each study was independently assessed by 2 reviewers across all checklist domains, including study aims, design, recruitment strategy, data collection, analysis, and potential bias. Discrepancies between the reviewers were resolved through consensus. To further strengthen methodological rigor and confirm the reliability of the CASP-based evaluations, the AXIS (Appraisal tool for Cross-Sectional Studies) [47] checklist for cross-sectional studies was additionally applied as a supplementary framework (Multimedia Appendix 4).

Synthesis Approach

With regard to the synthesis approach, due to the heterogeneity of the included studies (technologies, outcome measures, countries), a narrative synthesis approach was applied keeping in mind the principles of thematic content analysis [48]. Data were coded inductively to identify recurring themes related to determinants and facilitators of digital technology acceptance. These themes were subsequently compared and mapped to ensure conceptual coherence across studies [49]. As this review analyzed previously published studies, no ethical approval was required.

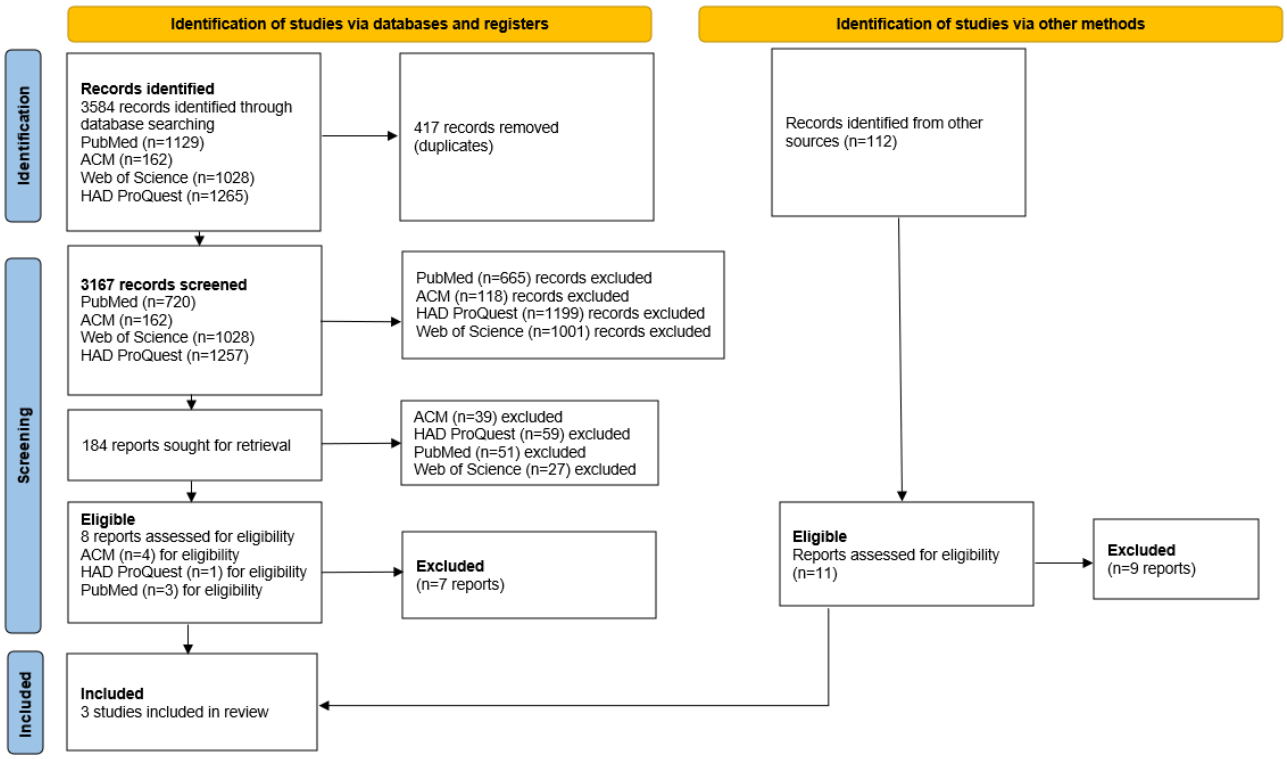
Results

Study Selection

The outcome of the literature search initially yielded 3584 records from the databases and an additional 112 studies from citation tracking and manual searching as demonstrated in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram—identification of studies.



After the removal of duplicates, the screening of the studies, and the application of eligibility criteria, 3 studies were included in the final analysis [12,26,39]. The included studies reported quantitative findings using the following measures:

- Likert-scale derived scores: these were used in all studies to assess acceptance variables (eg, attitudes, fears, perceived usefulness).
- Regression coefficients: these were reported in Barisch-Fritz et al [12] and Yu et al [39] to identify predictors of acceptance (eg, age, gender, professional group).
- Descriptive statistics: frequencies, means, and standard deviations were commonly used to present the results.

For this study, qualitative studies were eligible; however, none were identified for the final selection. Although limited in number, these studies offer initial insights into key acceptance factors and provide a basis for further investigation. These studies were conducted in Germany, the Netherlands, and Australia. Together, they involved 1019 participants, most of whom were direct care workers in nursing homes or LTC facilities. Across all studies, 867 were nursing staff, 99 were nursing home managers, and 53 were other staff members in LTC facilities (eg, clerks). The technologies under investigation ranged from electronic documentation systems to assistive robotic devices and digital communication platforms.

Risk of Bias

The risk of bias was assessed using the CASP checklist for cross-sectional studies (Table 2). Individual checklist items were evaluated qualitatively to appraise methodological rigor, and each cell in Table 2 represents the reviewer's consensus.

XSL•FO
RenderX

Table . Critical Appraisal Skills Programme (CASP) evaluation.

CASP	Barisch-Fritz et al (2023) [12]	de Veer et al (2011) [26]	Yu et al (2009) [39]
1. Did the study address a clearly focused issue?	Yes	Yes	Yes
2. Did the authors use an appropriate method to answer their question?	Yes	Yes	Yes
3. Were the subjects recruited in an acceptable way?	Yes	Yes	Cannot tell
4. Were the measures accurately measured to reduce bias?	Cannot tell	No	Yes
5. Were the data collected in a way that addressed the research issue?	Yes	Yes	Yes
6. Did the study have enough participants to minimize the play of chance?	Cannot tell	Yes	Cannot tell
7. How are the results presented and what is the main result?	Yes	Yes	Yes
8. Was the data analysis sufficiently rigorous?	Yes	Yes	Yes
9. Is there a clear statement of findings?	Yes	Yes	Yes
10. Can the results be applied to the local population?	Cannot tell	Yes	Cannot tell
11. Is the research valuable?	Cannot tell	Yes	Yes

While the CASP checklist provided a structured approach to appraising methodological quality, we also considered the AXIS critical appraisal tool, and it confirmed the initial CASP-based judgments.

Two studies [12,39] were rated as having moderate risk of bias. The study by Barisch-Fritz et al [12] addressed a clearly focused issue using validated instruments. Although the sample was good, it was not randomly selected, introducing potential self-selection bias. The study by Yu et al [39] also had a moderate risk of bias. It demonstrated strong internal validity through the use of validated TAM2-based instruments and a clearly defined research aim. However, some limitations remain; for instance, the convenience sampling reduced the strength of the recruitment process. Also, the relatively small sample size limits generalizability. These facts increase the potential for selection and sampling bias. The study by de Veer et al [26] demonstrated low risk of bias, supported by transparent reporting and robust measurement design. Nevertheless, the study lacked detailed information on how bias was addressed through measurement design.

Study Characteristics

The key findings and characteristics of the included studies are summarized in Table 3. Across the 3 studies, several patterns emerged regarding the implementation and acceptance of technology in nursing and LTC settings. In the study by de Veer et al [26], approximately half of the nursing staff had encountered new technologies within the past 3 years and generally perceived these introductions positively. However, actual use was hindered by technology-related factors, such as ease of use, patient relevance, and potential risks. Respondents emphasized the need for structured innovation strategies and organizational support. Similarly, Yu et al [39] in Australia confirmed the validity of a modified TAM2 model for LTC facilities, identifying perceived usefulness, ease of use, professional image, and computer skills as primary determinants of the intention to adopt health IT applications. The German nationwide survey by Barisch-Fritz et al [12] extended these findings, showing that acceptance and technology affinity depend on education, professional role, and sociodemographic characteristics. Lower acceptance was observed among older employees.

XSL•FO
RenderX

Table . Included studies: key findings.

Authors	Title	Year of publication	Research method	Country	Aim of study	Which technology?	Participants included	Key findings
de Veer et al [26]	Successful implementation of new technologies in nursing care: a questionnaire survey of nurse-users	2011	Questionnaire survey	The Netherlands	To gain a better understanding of determinants influencing the success of the introduction of new technologies as perceived by nursing staff	• New technologies introduced in the past three years • Electronic information systems • Distant care technology • Medical devices	• 685 nursing staff	• Half of the respondents were confronted with the introduction of new technology in the past 3 years • Half of them rated the introduction of the technology as positive • Factors impeding actual use were related to the technology itself: ie, malfunctioning, ease of use, relevance for patients, risk to patients • Nursing staff stressed the importance of an adequate innovation strategy

XSL•FO
RenderX

Authors	Title	Year of publication	Research method	Country	Aim of study	Which technology?	Participants included	Key findings
Yu et al [39]	Health IT acceptance factors in LTC[a] facilities: a cross-sectional survey	2009	Self-administered questionnaire	Australia	To examine the factors determining the acceptance of health IT applications by caregivers in LTC facilities	• Health IT applications (software, documentation)	• 134 questionnaires • Nurses (n=105) • LTC clerks (n=11) • Nursing managers (n=18)	• Approved the validity of a modified TAM2[b] in LTC facilities • Factors influencing caregivers' intention to use IT technology were perceived usefulness, perceived ease of use, image, and computer skills
Barisch-Fritz et al [12]	Are nursing home employees ready for the technical evolution? German-wide survey on the status quo of affinity for technology and technology interaction	2023	Online survey	Germany	Examine affinity for technology and technology interaction and related sociodemographic confounders, as well as detect possible requirements and boundary conditions relevant for the development and implementation of assistive technologies among nursing home employees	• Technology, assistive technologies (eg, networked systems, assistive humanoid or social robots, mobile applications)	• 200 nursing home employees • Nursing and therapy operation (n=77) • Nursing home manager (n=81) • Others in LTC (n=42)	

XSL•FO
RenderX

Authors	Title	Year of publication	Research method	Country	Aim of study	Which technology?	Participants included	Key findings
								• Positive consequences depended on education and professional group and the affinity for technology varied across age and gender • Lower acceptance with increasing age • Lower acceptance for females • Lower acceptance among nursing home managers

[a]LTC: long-term care.

[b]TAM2: technology acceptance model.

Despite differences in geographic context and methodological design, the studies share some overlapping findings regarding common factors that influence the acceptance. Perceived usefulness and perceived ease of use [26,39] consistently emerged as important determinants of acceptance. In addition, digital competence, defined as the ability to interact confidently with digital tools, was positively associated with willingness to use technology, particularly among younger staff members [12]. Organizational support, including leadership endorsement, training opportunities, and the involvement of staff in decision-making processes, also acted as a strong facilitator [26].

These cross-cutting themes are summarized in Table 4, which illustrates the main factors affecting the acceptance across studies.

XSL•FO
RenderX

Table . Factors affecting acceptance.

Authors	Title	Year of publication	Strengths	Weaknesses	Practical relevance	Factors affecting acceptance
de Veer et al [26]	Successful implementation of new technologies in nursing care: a questionnaire survey of nurse-users	2011	Strategic depth, very practical, multisectoral representativeness	• Little quantitative analysis • Mainly qualitative; not 100% LTC-specific[a]	Very high: helpful for implementation planning, LTC sector, and hospital	• Involvement of nursing staff during development and implementation affects acceptance. • Organizational support, such as leadership endorsement, communication, and available training does increase adoption. • Perceived relevance of the technology for patient care enhances likelihood of use.
Yu et al [39]	Health IT acceptance factors in LTC facilities: a cross-sectional survey	2009	Theoretically grounded, structural modeling, clear implications	• Limited representativeness, convenience sample, preimplementation data	Moderate to high: theoretical insights; limited practical transferability relevant for IT strategies in the LTC context	• Perceived usefulness is the strongest predictor of care staff intention to use digital technologies. • Digital competence correlates positively with willingness to use technology, particularly among younger staff. • Negative perceptions through IT use (image factor) reduce acceptance. Ease of use significantly influences both perceived usefulness and intention to adopt technology. • Perceived relevance of the technology for patient care enhances likelihood of use.

XSL•FO
RenderX

Authors	Title	Year of publication	Strengths	Weaknesses	Practical relevance	Factors affecting acceptance
Barisch-Fritz et al [12]	Are nursing home employees ready for the technical evolution? German-wide survey on the status quo of affinity for technology and technology interaction	2023	Good sample, valid measurement instruments, differentiated results	• Confounder control • Nonrandom sampling, response bias likely	High: directly applicable to nursing homes	• Digital competence correlates positively with willingness to use technology, particularly among younger staff. • Technology affinity varies strongly across age, gender, and professional role. • Organizational support, such as leadership endorsement, communication, and available training does increase adoption. • Ethical concerns can limit technology acceptance.

[a]LTC: long-term care.

All 3 studies contributed important evidence regarding factors influencing acceptance, organizational support, training availability, perceived usefulness, and digital competence. To account for heterogeneity across technologies and study designs, the extracted data were grouped thematically into 3 analytical levels: individual, organizational, and technological (Table 5).

This comparative thematic structure enabled a coherent synthesis across diverse contexts. Perceived usefulness, digital competence, organizational readiness, and usability emerged consistently across studies, supporting central constructs of the TAM.

XSL•FO
RenderX

Table . Thematic synthesis of factors affecting acceptance.

Level	Technology type	Emerging themes	Example evidence	Studies contributing
Individual	Electronic information and documentation systems; telecare software	Digital literacy, perceived usefulness, professional image, computer self-efficacy	Staff with higher digital competence and positive attitudes toward electronic documentation and telecare reported higher acceptance. Perceived usefulness and ease of use predicted intention to adopt these systems.	Barisch-Fritz et al (2023) [12] ; Yu et al (2009) [39]
Organizational	EHR[a] systems; digital readiness tools	Training, managerial support, workload, innovation climate	Organizational readiness, management involvement, and access to training facilitated technology use, while workload and lack of structured implementation strategies reduced uptake.	de Veer et al (2011) [26]; Barisch-Fritz et al (2023) [12]
Technological	Assistive technologies; robots; health IT software	Usability, reliability, system relevance, perceived ethical and professional implications	Usability and reliability were decisive for acceptance across all technologies, whereas assistive and robotic technologies introduced concerns regarding trust, ethics, and role identity.	de Veer et al (2011) [26]; Yu et al (2009) [39]; Barisch-Fritz et al (2023) [12]

[a]EHR: electronic health record.

Discussion

Interpretation of Findings

The synthesis of the 3 studies revealed that the acceptance of digital technologies in geriatric LTC depended on a combination of individual and organizational factors. Consistent with TAM and its extensions, usefulness and ease of use were the most robust predictors across the studies.

Beyond individual and organizational determinants, contextual factors, such as organizational culture, leadership style, and national policy frameworks, also influence digital readiness in LTC. Environments with a long-standing emphasis on innovation and participatory care culture may facilitate staff involvement in digital implementation, whereas strict data-protection orientation and reliance on paper-based processes may hinder the change. National eHealth infrastructures, such as Germany's Telematics Infrastructure and reimbursement policies, can affect incentives for adoption. Recognizing these dimensions is essential, as technological acceptance should not be understood in isolation from broader policy and organizational environments [50].

Previous reviews have also highlighted the importance of user attitudes and digital competencies for successful implementation [51-53]. Staff who feel confident in their ability to use digital tools are more willing to adopt them. This is particularly relevant given the generational differences observed in digital affinity. Younger staff members tend to have higher levels of acceptance, while older staff may require more training and support.

Organizational conditions further contribute to acceptance. Early staff involvement in the selection, testing, and implementation of new technologies, combined with training and transparent communication, fosters adoption.

A valuable complement to the peer-reviewed evidence is the BGW report "Pflege 4.0" [45], which constitutes gray literature but offers important contextual insights. Drawing on a mixed methods dataset of 576 professional caregivers in Germany—140 of whom were from geriatric LTC facilities—the report explored both actual technology use and perceived barriers to adoption. Using various 5-point Likert scales (ranging from "does not apply" to "fully applies"; from "not familiar at all" to "very familiar"), the survey identified key concerns, such as fear of job loss, data protection concerns, lack of technical skills, and low participation in implementation processes. While the professional composition of respondents was not fully specified, the findings add practical relevance by highlighting workplace-level perceptions that mirror those reported in the peer-reviewed studies.

Limitations of Evidence

The limited number (n=3) of eligible studies and their predominantly cross-sectional nature restrict the ability to draw clear conclusions, even though they identify relevant influencing factors. Additionally, the studies differ in the types of technologies investigated, outcome measures used, representation of demographic groups, and regional contexts. This heterogeneity complicates direct comparisons, and it further limits the generalizability of the findings. For instance, Yu et al [39] conducted a preimplementation survey based on TAM2 in an Australian LTC context, focusing on intention to use the technology. On the other hand, de Veer et al [26] investigated actual technology implementation across multiple health care sectors in the Netherlands, including nursing homes, but not exclusively. Barisch-Fritz et al [12] explored technology affinity in German nursing homes, but their heterogeneous sample included managers and other staff in LTC facilities with a

XSL•FO
RenderX

relatively small response rate. This fact raises concerns regarding representativeness. These limitations hinder generalizability.

Although comprehensive efforts were made to include all relevant research, the review was limited to publications in English or German, and no protocol was registered in advance. In addition, the CINAHL database was not searched due to a missing institutional license. As CINAHL is a relevant source, other studies may not have been captured.

Implications

Given the limited number of studies and their methodological heterogeneity, the implications hereby should be interpreted with caution. Nevertheless, the evidence indicates that the successful implementation of digital technologies in geriatric LTC relies on strategies that are aligned to the needs, competencies, and experiences of nursing staff. Policies should prioritize ongoing digital training programs based on the different groups of users. Furthermore, implementation efforts should involve staff from the earliest planning stages, ensuring that their expertise informs both system design and rollout. Organizational support and transparent communication regarding the objectives, benefits, and limitations of new systems are essential to build trust and reduce uncertainty among nursing staff. Ethical concerns must be addressed proactively, particularly in relation to surveillance technologies and the preservation of interpersonal care dynamics. In terms of research, there is definitely a need for more robust, rigorous, and longitudinal studies to enhance external validity and provide a more comprehensive understanding of technology acceptance among nursing staff in geriatric LTC.

Conclusion

This systematic review demonstrates that the acceptance of digital technologies by nursing staff in geriatric LTC settings is shaped by a constellation of individual and organizational factors. Three key determinants emerged consistently across all studies.

First, digital competence significantly influences willingness to adopt new technologies. Nursing staff with higher digital affinity, especially younger staff members, show greater readiness to engage with digital tools in the workplace. This highlights the need for training programs that target all age and experience groups.

Second, perceived relevance of technologies to daily care practice affects acceptance. Nursing staff are more likely to accept innovations that support main aspects of nursing home care, such as documentation efficiency, communication, or safety.

Third, organizational support, including communication, managerial encouragement, access to training, and staff participation in the implementation processes of digital technologies, plays a crucial role.

In light of the structural and demographic relevance of geriatric LTC, future research should be directed toward building a strong evidence base on technology acceptance. This review offers several testable hypotheses derived from the synthesized evidence. Future studies should empirically examine how early involvement of nursing staff in the development and implementation of digital technologies affects subsequent acceptance and sustained use. It can be hypothesized that organizational support mechanisms, including leadership endorsement, effective communication, and targeted digital training, strengthen the relationship between perceived usefulness and intention to use. Likewise, digital competence may mediate the relationship between training and technology adoption, while factors such as technology affinity, age, and professional role may moderate these effects. Furthermore, perceived relevance for patient care likely increases acceptance by reinforcing the perceived usefulness of digital tools, whereas ethical concerns or a negative professional image of IT use may inhibit adoption. Testing these mechanisms through longitudinal, intervention-based, or multilevel study designs could provide stronger causal evidence for the transformation strategies in geriatric LTC.

Acknowledgments

JI and the coauthors would like to thank Dr Felix Holl from Hochschule Neu-Ulm and also Amela (Gjoni) Brechlin, BA, from the nursing home St. Elisabeth Pflegezentrum gGmbH in Senden for the interesting discussions and the experience shared with regard to this topic.

Authors' Contributions

JI, WS, and DH participated in the design of the study. JI conducted the search. JI and RH were responsible for the screening of the results. JI, WS, and DH reviewed, evaluated critically, and edited the manuscript. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strings for the databases searched.
[PNG File, 213 KB - nursing_v9i1e82223_app1.png]

XSL•FO
RenderX

Multimedia Appendix 2
Overview search in PubMed conducted on October 25, 2025, to explore the topic prior to the systematic search.
[PNG File, 196 KB - nursing_v9i1e82223_app2.png]

Multimedia Appendix 3
Excluded publications after eligibility assessment according to the exclusion criteria.
[XLSX File, 15 KB - nursing_v9i1e82223_app3.xlsx]

Multimedia Appendix 4
AXIS (Appraisal tool for Cross-Sectional Studies) quality assessment of included studies.
[XLSX File, 12 KB - nursing_v9i1e82223_app4.xlsx]

Checklist 1
PRISMA 2020 checklist.
[PDF File, 190 KB - nursing_v9i1e82223_app5.pdf]

References

1. Houben D. Reflexive digitalisierung und kritische digitalkompetenz [Book in German]. In: Newiak D, Romppel J, Martin A, editors. Digitale Bildung Jetzt!: Springer; 2023:75-89. [doi: 10.1007/978-3-658-40845-9_5]
2. Konopik J, Blunck D. Development of an evidence-based conceptual model of the health care sector under digital transformation: integrative review. J Med Internet Res 2023 Jun 8;25:e41512. [doi: 10.2196/41512] [Medline: 37289482]
3. Rantanen T, Lehto P, Vuorinen P, Coco K. The adoption of care robots in home care—a survey on the attitudes of Finnish home care personnel. J Clin Nurs 2018 May;27(9-10):1846-1859. [doi: 10.1111/jocn.14355] [Medline: 29575204]
4. Swoboda W, Seifert N. Digitale Innovationen in Der Pflege [Book in German]: Springer; 2024. [doi: 10.1007/978-3-662-67914-2]
5. Bitterlich S, Götze U. Pflege 4.0 – digitale transformation eines pflegeunternehmens [Book in German]. In: Controlling & Innovation 2022: Springer; 2022:27-60. [doi: 10.1007/978-3-658-36484-7_2]
6. Beaney P, Hatfield R, Hughes A, Schmid M, Chambers R. Creating digitally ready nurses in general practice. Nurs Manag (Harrow) 2019 May 28;26(3):27-35. [doi: 10.7748/nm.2019.e1840] [Medline: 31468839]
7. Schlicht L, Wendsche J, Melzer M, Tschetsche L, Rösler U. Digital technologies in nursing: an umbrella review. Int J Nurs Stud 2025 Jan;161:104950. [doi: 10.1016/j.ijnurstu.2024.104950] [Medline: 39603090]
8. Huter K, Krick T, Domhoff D, Seibert K, Wolf-Ostermann K, Rothgang H. Effectiveness of digital technologies to support nursing care: results of a scoping review. J Multidiscip Healthc 2020;13:1905-1926. [doi: 10.2147/JMDH.S286193] [Medline: 33328736]
9. Krick T, Zerth J, Klawunn R. Pflegeinnovationen in der praxis: erfahrungen und empfehlungen aus dem „cluster zukunft der pflege“, einführung und orientierung [Book in German]. In: Pflegeinnovationen in Der Praxis: Springer; 2023:1-16. [doi: 10.1007/978-3-658-39302-1]
10. Hübner U, Shaw T, Thye J, et al. Technology Informatics Guiding Education Reform—TIGER. Methods Inf Med 2018 Jun;57(S 01):e30-e42. [doi: 10.3414/ME17-01-0155] [Medline: 29956297]
11. Pfannstiel MA. Technologien Und Technologiemanagement Im Gesundheitswesen: Potenziale Nutzen, Lösungen Entwickeln, Ziele Erreichen [Book in German]: Springer Gabler Wiesbaden; 2024. [doi: 10.1007/978-3-658-43860-9]
12. Barisch-Fritz B, Krafft J, Rayling S, et al. Are nursing home employees ready for the technical evolution? German-wide survey on the status quo of affinity for technology and technology interaction. Digit Health 2023;9:20552076231218812. [doi: 10.1177/20552076231218812] [Medline: 38144174]
13. Wangmo T, Lipps M, Kressig RW, Ienca M. Ethical concerns with the use of intelligent assistive technology: findings from a qualitative study with professional stakeholders. BMC Med Ethics 2019 Dec 19;20(1):98. [doi: 10.1186/s12910-019-0437-z] [Medline: 31856798]
14. Kruse CS, Mileski M, Alaytsev V, Carol E, Williams A. Adoption factors associated with electronic health record among long-term care facilities: a systematic review. BMJ Open 2015 Jan 28;5(1):e006615. [doi: 10.1136/bmjopen-2014-006615] [Medline: 25631311]
15. Ho KF, Ho CH, Chung MH. Theoretical integration of user satisfaction and technology acceptance of the nursing process information system. PLOS ONE 2019;14(6):e0217622. [doi: 10.1371/journal.pone.0217622] [Medline: 31163076]
16. Ageing and health. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-04-03]
17. Gesundheit pflege [Web page in German]. Statistisches Bundesamt. 2024. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Pflege/_inhalt.html [accessed 2025-04-21]

XSL•FO
RenderX

18. Konzertierte aktion pflege – zweiter bericht zum stand der umsetzung der vereinbarungen der arbeitsgruppen 1 bis 5 [Report in German]. : Bundesregierung; 2021 URL: https://www.publikationen-bundesregierung.de/pp-de/publikationssuche/konzertierte-aktion-pflege-2053798 [accessed 2025-04-21]
19. Digitale–Versorgung–und–Pflege–Modernisierungs–Gesetz (DVPMG) [Web page in German]. Bundesgesundheitsministerium. 2021. URL: https://www.bundesgesundheitsministerium.de/service/gesetze-und-verordnungen/guv-19-lp/dvpmg.html [accessed 2025-06-03]
20. Pressemitteilungen - bericht aus der kabinettssitzung vom 19. märz. 2024 [Web page in German]. Bayerische Staatskanzlei. URL: https://www.bayern.de/bericht-aus-der-kabinettssitzung-vom-19-maerz-2024 [accessed 2026-01-03]
21. Bis 2049 werden voraussichtlich mindestens 280 000 zusätzliche Pflegekräfte benötigt [Web page in German]. Statistisches Bundesamt. 2024. URL: https://www.destatis.de/DE/Presse/Pressemitteilungen/2024/01/PD24_033_23_12.html [accessed 2025-05-02]
22. Venkataraman A, Fatma N, Edirippulige S, Ramamohan V. Facilitators and barriers for telemedicine systems in India from multiple stakeholder perspectives and settings: a systematic review. Telemed J E Health 2024 May;30(5):1341-1356. [doi: 10.1089/tmj.2023.0297] [Medline: 38206654]
23. Müller P, Jahn P. Cocreative development of robotic interaction systems for health care: scoping review. JMIR Hum Factors 2024 Sep 12;11:e58046. [doi: 10.2196/58046] [Medline: 39264334]
24. Negro-Calduch E, Azzopardi-Muscat N, Krishnamurthy RS, Novillo-Ortiz D. Technological progress in electronic health record system optimization: systematic review of systematic literature reviews. Int J Med Inform 2021 Aug;152:104507. [doi: 10.1016/j.ijmedinf.2021.104507] [Medline: 34049051]
25. Krick T, Huter K, Domhoff D, Schmidt A, Rothgang H, Wolf-Ostermann K. Digital technology and nursing care: a scoping review on acceptance, effectiveness and efficiency studies of informal and formal care technologies. BMC Health Serv Res 2019 Jun 20;19(1):400. [doi: 10.1186/s12913-019-4238-3] [Medline: 31221133]
26. de Veer AJE, Fleuren MAH, Bekkema N, Francke AL. Successful implementation of new technologies in nursing care: a questionnaire survey of nurse-users. BMC Med Inform Decis Mak 2011 Oct 27;11:67. [doi: 10.1186/1472-6947-11-67] [Medline: 22032728]
27. Morton ME, Wiedenbeck S. EHR acceptance factors in ambulatory care: a survey of physician perceptions. Perspect Health Inf Manag 2010 Jan 1;7(Winter):1c. [Medline: 20697466]
28. Giebel GD, Speckemeier C, Abels C, et al. Problems and barriers related to the use of digital health applications: scoping review. J Med Internet Res 2023 May 12;25:e43808. [doi: 10.2196/43808] [Medline: 37171838]
29. Trainum K, Liu J, Hauser E, Xie B. Nursing staff's attitudes, needs, and preferences for care robots in assisted living facilities: a systematic literature review. 2024 Mar 11 Presented at: HRI '24; Mar 11-14, 2024. [doi: 10.1145/3610978.3640690]
30. Konttila J, Siira H, Kyngäs H, et al. Healthcare professionals' competence in digitalisation: a systematic review. J Clin Nurs 2019 Mar;28(5-6):745-761. [doi: 10.1111/jocn.14710] [Medline: 30376199]
31. Egbert N, Thye J, Hackl WO, Müller-Staub M, Ammenwerth E, Hübner U. Competencies for nursing in a digital world. Methodology, results, and use of the DACH-recommendations for nursing informatics core competency areas in Austria, Germany, and Switzerland. Inform Health Soc Care 2019;44(4):351-375. [doi: 10.1080/17538157.2018.1497635] [Medline: 30148411]
32. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
33. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
34. Trainum K, Liu J, Hauser E, Xie B. Nursing staff's perspectives of care robots for assisted living facilities: systematic literature review. JMIR Aging 2024 Sep 16;7:e58629. [doi: 10.2196/58629] [Medline: 39283666]
35. Nittas V, Daniore P, Chavez SJ, Wray TB. Author correction: challenges in implementing cultural adaptations of digital health interventions. Commun Med (Lond) 2024 Mar 20;4(1):54. [doi: 10.1038/s43856-024-00479-x] [Medline: 38509219]
36. Niwa LMS, Radovich NMF, Ciosak SI. Safe Embrace: technological innovation for elderly safety in the use of toilets. Rev Bras Enferm 2018;71(suppl 6):2833-2836. [doi: 10.1590/0034-7167-2017-0751] [Medline: 30540063]
37. Walsh K, Callan A. Perceptions, preferences, and acceptance of information and communication technologies in older-adult community care settings in Ireland: a case-study and ranked-care program analysis. Ageing Int 2011 Mar;36(1):102-122. [doi: 10.1007/s12126-010-9075-y]
38. Sun SL, Hwang HG, Dutta B, Peng MH. Exploring critical factors influencing nurses' intention to use tablet PC in patients' care using an integrated theoretical model. Libyan J Med 2019 Dec;14(1):1648963. [doi: 10.1080/19932820.2019.1648963] [Medline: 31357919]
39. Yu P, Li H, Gagnon MP. Health IT acceptance factors in long-term care facilities: a cross-sectional survey. Int J Med Inform 2009 Apr;78(4):219-229. [doi: 10.1016/j.ijmedinf.2008.07.006] [Medline: 18768345]
40. Pavlova M, Praznovszky L, Tambor M, Golinowska S, Groot W. Introduction. Challenges in long-term care in Europe. OZ 2019;17(3):113-118. [doi: 10.4467/20842627OZ.19.013.11970]

XSL•FO
RenderX

41. Yuan S, Coghlan S, Lederman R, Waycott J. Social robots in aged care: care staff experiences and perspectives on robot benefits and challenges. Proc ACM Hum-Comput Interact 2022 Nov 7;6(CSCW2):1-23. [doi: 10.1145/3555220]
42. Khalaily J. The importance of digital technology for elderly care in nursing homes. spas 2024;12(35):123-127. [doi: 10.70147/s35123127]
43. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
44. Cooke A, Smith D, Booth A. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. Qual Health Res 2012 Oct;22(10):1435-1443. [doi: 10.1177/1049732312452938] [Medline: 22829486]
45. Merda M, Schmidt K, Kähler B. Pflege 4.0 – Einsatz moderner Technologien aus der Sicht professionell Pflegender [Report in German]. : Berufsgenossenschaft für Gesundheitsdienst und Wohlfahrtspflege; 2017 URL: https://www.bgw-online.de/resource/blob/20346/e735030f6178101cf2ea9fa14e1bc063/bgw09-14-002-pflege-4-0-einsatz-moderner-technologien-data.pdf [accessed 2026-01-03]
46. CASP checklist: CASP cross-sectional studies checklist. Critical Appraisal Skills Programme. URL: https://casp-uk.net/casp-tools-checklists/cross-sectional-studies-checklist/ [accessed 2026-01-03]
47. Downes MJ, Brennan ML, Williams HC, Dean RS. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). BMJ Open 2016 Dec 8;6(12):e011458. [doi: 10.1136/bmjopen-2016-011458] [Medline: 27932337]
48. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
49. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
50. Kachimanga C, Divala TH, Ket JCF, et al. Adoption of mHealth technologies by community health workers to improve the use of maternal health services in sub-Saharan Africa: protocol for a mixed method systematic review. JMIR Res Protoc 2023 May 4;12:e44066. [doi: 10.2196/44066] [Medline: 37140981]
51. Wutz M, Hermes M, Winter V, Köberlein-Neu J. Factors influencing the acceptability, acceptance, and adoption of conversational agents in health care: integrative review. J Med Internet Res 2023 Sep 26;25:e46548. [doi: 10.2196/46548] [Medline: 37751279]
52. Jacob C, Sanchez-Vazquez A, Ivory C. Understanding clinicians' adoption of mobile health tools: a qualitative review of the most used frameworks. JMIR Mhealth Uhealth 2020 Jul 6;8(7):e18072. [doi: 10.2196/18072] [Medline: 32442132]
53. Harst L, Lantzsch H, Scheibe M. Theories predicting end-user acceptance of telemedicine use: systematic review. J Med Internet Res 2019 May 21;21(5):e13117. [doi: 10.2196/13117] [Medline: 31115340]

Abbreviations

AXIS: Appraisal tool for Cross-Sectional Studies
CASP: Critical Appraisal Skills Programme
LTC : long-term care
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SPIDER: sample, phenomenon of interest, design, evaluation, research type
TAM: technology acceptance model

Edited by E Borycki; submitted 12.Aug.2025; peer-reviewed by M Hasseler, MA Zoubi; revised version received 07.Nov.2025; accepted 28.Dec.2025; published 15.Jan.2026.

Please cite as:
Iseni J, Swoboda W, Houben D, Hilla R
Acceptance of Digital Technology Among Nursing Staff in Geriatric Long-Term Care: Systematic Review
JMIR Nursing 2026;9:e82223
URL: https://nursing.jmir.org/2026/1/e82223
doi:10.2196/82223

© Jeton Iseni, Walter Swoboda, Daniel Houben, Roman Hilla. Originally published in JMIR Nursing (https://nursing.jmir.org), 15.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Nurses' Experiences Using AI in Clinical Practice: Systematic Review

Ashley J S Scott[1,2], MSc; Qimeng Zhao[1], MSc; Jo-Fan Pan[1], PhD; Benjamin C Brown[3], PhD; Dawn Dowding[1], PhD

[1]Division of Nursing, Midwifery and Social Work, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Oxford Road, Manchester, England, United Kingdom

[2]School of Nursing, Faculty of Applied Sciences, University of British Columbia, Vancouver, BC, Canada

[3]Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, Manchester, England, United Kingdom

Corresponding Author:
Ashley J S Scott, MSc
Division of Nursing, Midwifery and Social Work, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Oxford Road, Manchester, England, United Kingdom

Abstract

Background: Artificial intelligence (AI) tools are increasingly used in clinical settings, yet most syntheses focus on nurses' attitudes or readiness rather than experiences after direct use in practice.

Objective: This study aimed to synthesize registered nurses' experiences of using AI in clinical practice and to identify perceived benefits, barriers, and implementation implications.

Methods: We conducted a systematic literature review of empirical studies reporting nurses' experiences of AI use in clinical settings. Searches were performed in CINAHL, Embase, MEDLINE, PsycINFO, and PubMed (last search: September 13, 2023). Moreover, 2 reviewers (AJSS and QZ) independently screened titles, abstracts, and full texts and appraised included studies using the Mixed Methods Appraisal Tool. We used thematic synthesis with a primarily deductive framework based on the Technology Acceptance Model 2, with the addition of facilitating conditions from the unified theory of acceptance and use of technology.

Results: In total, 20 studies met the inclusion criteria. Perceived usefulness and facilitating conditions were most frequently reported; nurses described AI as supporting decision-making, workflow efficiency, and confidence when implementation included adequate training, interoperability, and technical infrastructure. Ease of use was closely tied to interface design and training. Job relevance and output quality were rated more positively when nurses described AI as aligning with nursing tasks and as producing interpretable, reliable outputs. Common barriers included usability issues, limited integration into workflows and electronic systems, privacy and trust concerns, and inconsistent or poorly contextualized outputs. Across studies, nurses often described adoption as conditional on organizational readiness and meaningful involvement of nurses in design and implementation.

Conclusions: Nurses' accounts suggest that AI may augment clinical work, but the perceived benefits reported in the included literature were contingent on workflow alignment, usable interfaces, training, and supportive infrastructure. These findings suggest potential value in involving nurses in the co-design and iterative refinement of AI tools, grounded in the consistent evidence that facilitating conditions, such as training, interoperability, and organizational readiness, are the primary determinants of whether AI is experienced positively. Prospective evaluation using nursing-relevant outcomes, such as usability, workflow integration, and trust, is needed to move beyond post hoc experiential accounts and establish what implementation conditions reliably produce benefit.

Trial Registration: PROSPERO CRD42022308051; https://www.crd.york.ac.uk/PROSPERO/view/CRD42022308051

(JMIR Nursing 2026;9:e91238) doi:10.2196/91238

KEYWORDS

artificial intelligence; nursing; clinical practice; user experience; decision support systems; technology acceptance model; implementation; systematic review; PRISMA

Introduction

The integration of artificial intelligence (AI) into health care has transformed clinical practice, significantly impacting the roles and responsibilities of registered nurses (RNs). As the largest group of health professionals, RNs are at the forefront of these technological shifts and have historically been an afterthought in technology development [1]. AI holds significant potential to enhance nursing practice in administration, clinical practice, policy, and research, but it remains an area of ongoing exploration [2]. For example, natural language processing

XSL•FO
RenderX

systems are increasingly used to assist with clinical documentation, decision support algorithms enhance diagnostic reasoning, and in nursing administration, AI has been applied to workload prediction and staff scheduling, while in research, it supports pattern recognition in large datasets [3].

AI's emergence in health care has altered nursing roles, workflows, and the nurse-patient relationship [4,5]. Technologies like predictive analytics, virtual health care assistants, and robotics are influencing nursing practice in substantive ways: (1) predictive analytics supports early identification of deteriorating patients, enabling more proactive clinical intervention; (2) virtual health care assistants facilitate patient education and remote monitoring; and (3) robotic technologies are increasingly used for medication dispensing and patient mobility support, reshaping how physical tasks are distributed [6-8]. While these AI-driven technologies offer improvements in health care delivery, they also challenge the maintenance of person-centered, compassionate care [2]. AI already contributes significantly to decision-making support and skill augmentation within clinical practice [9]. Nurses must proactively ensure AI technologies align with nursing's core values, maintaining ethical and compassionate care [10].

The integration of AI into nursing practice is accelerating, and existing systematic reviews predominantly explore general perceptions and theoretical readiness or educational readiness among nurses [11,12]. These studies offer valuable insight into anticipated benefits and challenges, yet conclusions are derived from nurses who have not directly interacted with AI in clinical settings. For example, 1 recent review synthesized qualitative studies focused on nurses' perceptions without requiring previous AI use [12]. Several recent reviews have examined AI specifically within nursing education contexts; one explored AI use in nursing training [13], while another synthesized empirical evidence on AI implementation in hospital settings using a strengths, weaknesses, opportunities, and threats framework [14], and another mapped AI applications, such as generative tools and virtual patient simulators, in university-level nursing programs [15]. All 3 address AI as a pedagogical resource for nursing students and trainees, rather than clinical practice. There remains a lack of synthesis of studies capturing nurses' attitudes after engaging directly with AI tools.

Several theoretical frameworks can be used to understand individuals' acceptance and use of technology, such as AI. The Technology Acceptance Model 2 (TAM2) is among the most extensively validated, proposing that perceived usefulness, perceived ease of use, and social and cognitive influences shape technology adoption [16]. TAM2 was developed in general organizational technology contexts, and we acknowledge that it may underweight dimensions that are particularly salient in nursing, including professional judgment, the nurse-patient relationship, ethical responsibility, accountability for patient outcomes, and the emotional and moral dimensions of clinical work. We selected TAM2 nonetheless because our review question concerned nurses' acceptance-relevant experiences after direct AI use, where TAM2's validated constructs provide an established and transparent analytic vocabulary. To address TAM2's primary acknowledged limitation, its limited treatment of organizational and infrastructural context, we incorporated the Facilitating Conditions construct from the unified theory of acceptance and use of technology (UTAUT) [17], which captures the technical, training, and resource conditions that nurses' accounts suggested were central to their experiences. We did not adopt the full UTAUT to avoid construct duplication with TAM2 (the implications of this deductive choice for what the synthesis could and could not surface are discussed in the Limitations section). This review explores the experiences of RNs navigating modern health care complexities, with a focus on AI's impact on nursing roles, the challenges in adapting to AI technologies, and how these changes are perceived to influence patient care using TAM2 and UTAUT's Facilitating Conditions as the analytical framework.

The research question guiding this systematic review is "What are the experiences of nurses using AI in their clinical work?" The objectives are to (1) develop a better understanding of how nurses interact with AI in clinical practice, and (2) provide insights into nurses' perceptions of AI's impact on efficiency, satisfaction, and nurse-patient relationships.

Methods

Overview

The primary objective of this systematic review is to understand the experiences of nurses using AI in clinical practice.

A systematic review methodology was selected because the research question required a comprehensive, transparent, and reproducible synthesis of empirical evidence from multiple studies. This approach is appropriate for identifying and integrating findings across a body of literature to address a focused question about a specific population's experiences, and is consistent with established guidance for evidence synthesis in health professions research [18].

Protocol and Registration

The study protocol was registered with PROSPERO (CRD42022308051) [19]. The manuscript title has been refined from the registered title for clarity; the review question, eligibility criteria, and methods are consistent with the registered protocol.

Ethical Considerations

Ethics approval was not required because this study synthesizes data from previously published literature and did not involve collection of identifiable participant data.

Eligibility Criteria

Eligibility was determined using predefined inclusion and exclusion criteria (Textbox 1).

XSL•FO
RenderX

Textbox 1. Eligibility criteria.

Inclusion criteria

- Empirical research published in English: to ensure consistency and enable thorough analysis, only studies published in English were included.
- Peer-reviewed journal articles: studies published in peer-reviewed journals were prioritized for their methodological rigor and reliability.
- Studies where registered nurses describe their experiences of using artificial intelligence (AI) in clinical practice: the review focused on direct, real-world experiences of registered nurses with AI technologies. For inclusion, studies were required to report on nurses' experiences or perceptions arising from actual use of or interaction with an AI tool in a clinical setting. Studies reporting only on anticipated, hypothetical, or theoretical attitudes toward AI, without confirmed direct use, were excluded.
- Non–peer-reviewed sources (such as poster presentations and conference abstracts): these would have been considered only if they contained unique empirical data not yet published in a peer-reviewed format and demonstrated methodological rigor as assessed using the Mixed Methods Appraisal Tool (MMAT) criteria applied to all included studies.

Exclusion criteria

- Studies published in languages other than English: non-English language publications were excluded due to language barriers affecting interpretation.
- Studies reporting the experiences of nurses using technology not underpinned by AI: studies reporting on technologies that were not underpinned by AI were excluded.
- Studies reporting the experiences of nurses working in non-clinical settings: research focusing on administrative, academic, or educational settings outside of clinical practice were excluded.
- Studies related to health care professionals (eg, midwives and doctors): studies were excluded unless they specifically reported data related to registered nurses.
- Duplicate reports of the same study: the most comprehensive versions of duplicate studies were included.

Information Sources

A comprehensive literature search was conducted on September 13, 2023 across 5 electronic databases—PubMed, Embase, MEDLINE, CINAHL, and PsycInfo.

Search Strategy

A search strategy was developed and refined through repeated testing and adjustment, using both keywords and controlled vocabulary tailored to each database. Concept blocks, illustrated in Textbox 2, provided the structure for the search, with keywords and controlled vocabulary terms merged using Boolean logic (AND/OR) to ensure comprehensiveness. No date restrictions were applied, allowing for the inclusion of all eligible studies up to the date of the search. Searches were not restricted by language, but inclusion was limited to studies published in English. Reference lists of identified articles were also reviewed, but no additional studies met the inclusion criteria. Full search strategies are provided in Multimedia Appendix 1.

Textbox 2. Search strategy concept blocks.

Block 1: Nursing

- Keywords: Nurse* (Title or Abstract), "Nursing staff" (Title or Abstract), "Nursing professional*" (Title or Abstract), "Registered nurse*" (Title or Abstract)
- CINAHL Headings: Nurses, Nursing Staff

Block 2: Artificial intelligence

- Keywords: "Artificial Intelligence" (Title or Abstract), "Machine Learning" (Title or Abstract), "Neural Network*" (Title or Abstract), "Deep Learning" (Title or Abstract), "Computer Vision" (Title or Abstract), "Natural Language Processing" (Title or Abstract)
- CINAHL Headings: Artificial Intelligence, Machine Learning

Block 3: Experiences or perceptions

- Keywords: Experience* (Title or Abstract), Perception* (Title or Abstract), View* (Title or Abstract), Attitude* (Title or Abstract), Feedback (Title or Abstract), Opinion* (Title or Abstract)
- CINAHL Headings: Attitude, Perception

Study Selection

The review was conducted following the Joanna Briggs Institute (JBI) methodology for systematic reviews, using the "JBI Manual for Evidence Synthesis" [20] as the guiding framework. The reporting of this review adheres to PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines [21]. Study management and screening were managed with Covidence software (Veritas Health Innovation Ltd). Furthermore, 2 reviewers (AJSS and QZ) independently

XSL•FO
RenderX

screened titles, abstracts, and full texts, resolving discrepancies through discussion and consensus. In cases where multiple reports of the same study were identified, the most comprehensive version was prioritized, and duplicates were merged accordingly. The screening and selection process are presented in the PRISMA flow diagram in the Results section. The completed PRISMA 2020 checklist is also provided in Checklist 1.

Data Extraction

Data extraction followed the JBI mixed methods data extraction form following a convergent integrated approach [20]. Where studies included multidisciplinary samples, nurse-specific findings were extracted separately where the primary data permitted. In cases where profession-specific disaggregation was not possible, this was noted explicitly in the data extraction table (Multimedia Appendix 2).

Quality Appraisal

Study quality was assessed using the Mixed Methods Appraisal Tool (MMAT) [22] to accommodate the diversity of included study designs. This aligned with the convergent integrated approach used in this review [20]. Additionally, 2 reviewers (AJSS and QZ) independently assessed study quality, and any discrepancies were resolved through discussion to reach consensus. To capture the complexity of interpretive and critical understandings of AI use in nursing, an inclusive approach to study selection was maintained. As noted by Hannes and Lockwood [23], qualitative evidence can simultaneously illuminate both broad patterns and context-specific detail, and excluding studies entirely based on quality grounds risks narrowing this breadth. Including all studies regardless of MMAT scores allowed for a representative synthesis across diverse clinical contexts. Quality appraisal results were therefore used to inform the interpretation of findings, rather than as a basis for exclusion. Where lower-quality studies contributed to a construct, this is noted within the relevant Results section.

Synthesis Approach

Thematic analysis was conducted using a predefined deductive framework based on TAM2, supplemented by the single UTAUT construct of Facilitating Conditions. All included studies were read in full, and relevant text segments, including direct participant quotations, authors' qualitative summaries, and quantitative findings, were coded line-by-line into the framework constructs. Coding was undertaken by the first author (AJSS), with regular review and discussion of coding decisions with 2 supervising coauthors (BCB and DD) throughout the analysis period. These supervisory discussions served as a calibration mechanism for the application of the codebook; ambiguous extracts and decisions about construct boundaries were brought to these meetings, discussed against operational definitions, and resolved by consensus before being applied across the dataset. The codebook (Table 1) was iteratively refined through these discussions to clarify operational boundaries; construct categories themselves remained fixed throughout. Formal interrater agreement was not calculated, as only a single author (AJSS) coded the data; this is acknowledged as a limitation. Quantitative findings relevant to TAM2 constructs, such as survey results, were summarized and integrated narratively with the qualitative data under relevant headings. Where frequency of reporting is described, this refers to the number of included studies in which relevant data were coded to a given construct. Frequency counts indicate how commonly a construct appeared across the included studies and should not be interpreted as a measure of the depth or strength of any individual theme.

XSL•FO
RenderX

Table . Codebook adapted from TAM2[a]+UTAUT[b].

Construct	Definition (as applied in this review)
Perceived usefulness	• Excerpts where the AI[c] is seen as having an impact on job performance. • Statements indicating that the AI impacts the effectiveness of task completion or enables more task opportunities. • Examples of how AI is perceived to enhance patient care or nursing outcomes.
Perceived ease of use	• Descriptions of how the AI is straightforward and easy to use or interact with. • Quotes about the learning curves in relation to using AI in practice. • Facilitators or challenges related to technical aspects of interacting with the AI.
Subjective norms	• Comments about how the organization, management, or colleagues influence attitudes toward AI. • Encouragement or pressure from peers to use or not use AI. • The role of policy supporting or discouraging AI use.
Intention to use	• Nurses' statements related to plans or indications to use AI in the future. • Motivating or deterring factors relating to the intention to use AI in practice. • Conditional intentions (eg, willingness to use AI if specific supports or changes are implemented).
Usage behavior	• Discussions of actual use or engagement with AI in practice. • Context, duration, and frequency of AI use in practice. • Examples of tasks or procedures where AI is being implemented.
Facilitating conditions[d]	• Technical or organizational infrastructure supporting the use of AI. • Availability or lack of resources like support staff, training, and technical expertise. • Compatibility of the AI with existing systems and workflows.
Job relevance	• To what degree do nurses think AI is relevant to their job. • Specific responsibilities or tasks that are impacted by AI. • Changes in responsibilities or job roles as a result of AI integration.
Output quality	• Perceptions of quality of the results provided by AI systems. • AI's impact on the overall quality, timeliness, and accuracy of nursing care. • Consistency and reliability of AI outputs.
Result demonstrability	• Visibility and tangibility of results of using AI in nursing practice. • Examples of observable and clearly articulated benefits or drawbacks of AI use. • Stories demonstrating the impact of AI use on patient outcomes or nursing practice.

[a]TAM2: Technology Acceptance Model 2.

[b]UTAUT: unified theory of acceptance and use of technology.

[c]AI: artificial intelligence.

[d]Added from UTAUT.

Results

Study Selection

The PRISMA flow diagram (Figure 1) is presented to illustrate the screening and selection process. The initial searches yielded 952 records. After 424 duplicate or merged records were removed in Covidence, 528 articles were screened by title and abstract. Of these, 60 reports were sought for retrieval and assessed for eligibility, resulting in the exclusion of 40 reports for reasons such as the lack of AI experience (n=15), lack of experiential data (n=11), a focus on non-AI technology (n=6), and other contextual or methodological reasons as detailed in Multimedia Appendix 3. Ultimately, 20 studies [24-43] met the inclusion criteria and were included in the final synthesis.

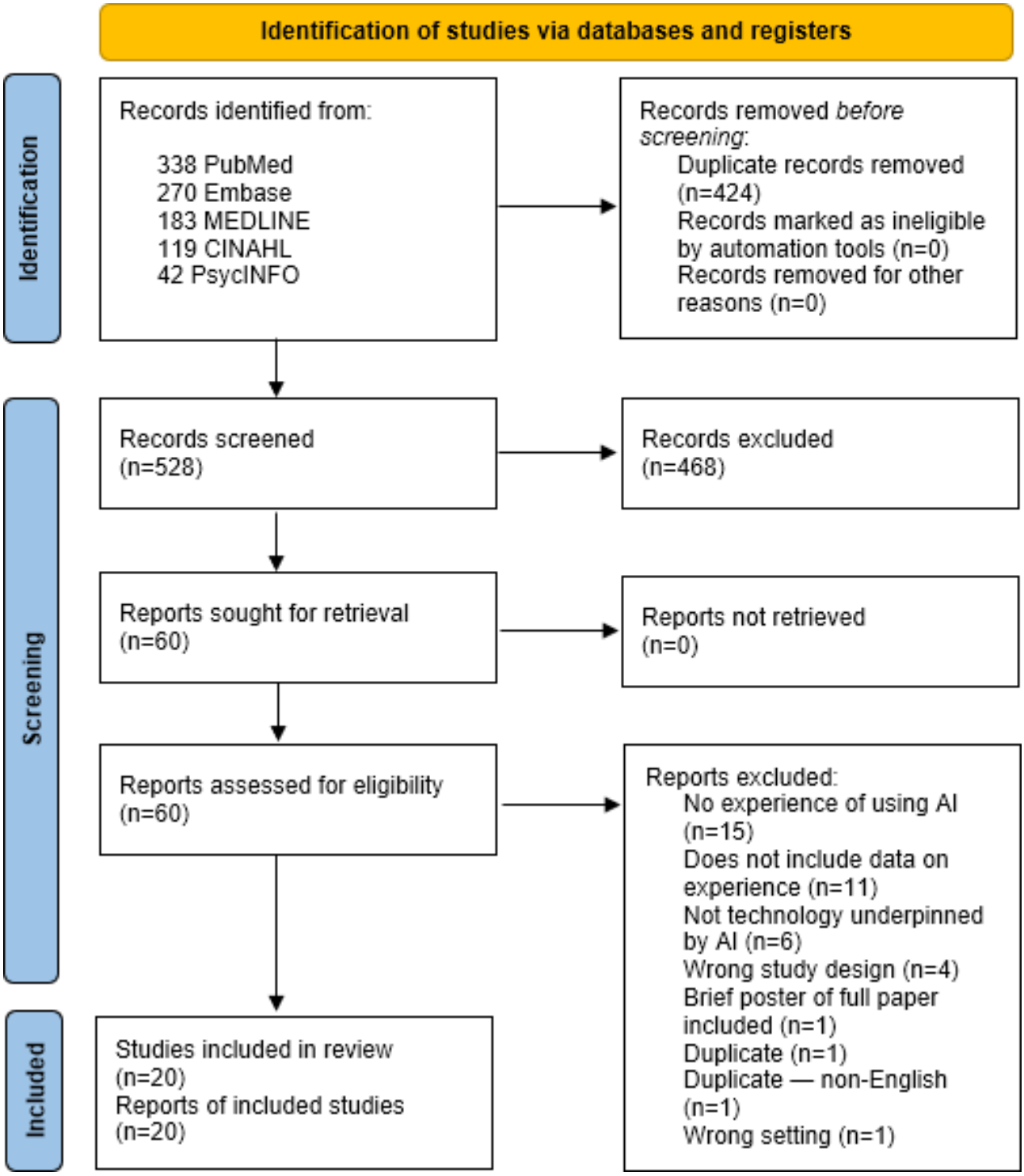
XSL•FO
RenderX

Of the 20 included studies, 13 were appraised as high quality [31-43], 5 as medium [24-28], and 2 as low [29,30] (Multimedia Appendix 4). The 5 medium-rated studies were predominantly quantitative surveys with limitations relating to sampling representativeness and nonresponse bias [24-28]. The 2 low-rated studies had insufficient methodological reporting to fully assess rigor [29,30]. Job Relevance and Output Quality were supported exclusively or predominantly by high-rated studies. Perceived Usefulness was well-represented across quality levels but anchored by a majority of high-rated studies. Facilitating Conditions drew contributions from the broadest quality range, including both low-rated studies; findings for this construct are therefore interpreted with appropriate caution below.

Of the 20 included studies, the proportion of nurse participants varied substantially. Of the total, 9 studies reported that nurses comprised the majority of participants [24,25,29,31-36]. In the remaining studies, nurses participated alongside other professional groups, and in some cases were a small minority. While there was no requirement for studies to have a majority of nurse participants, this variability affects the extent to which findings directly reflect nursing perspectives. Where possible, data specific to nurses was extracted for synthesis. However, in several studies (eg, [26,34,37]) nursing data were supplementary or aggregated with other professions. A full summary of participant demographics and professional backgrounds is presented in Table 2.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of study selection.



XSL•FO
RenderX

Table . Characteristics of included studies.

Study (year)	Country and setting	Study design	Participants	AI[a] and technology focus
Alanzi (2023) [31]	Saudi Arabia, health care sector (government hospitals)	Qualitative, focus group interviews	Total: 54 participants • Physicians: 11 • Nurses: 24 • Dieticians: 8 • Pharmacists: 6 • Physiotherapists: 5 • Sex: 33 females, 21 males • Majority have more than 6 years of experience • Professions: physicians, nurses, dieticians, pharmacists, and physiotherapists	• ChatGPT
Belmin et al (2022) [27]	France, multicenter trial with community-dwelling adults receiving home care	Quantitative, uncontrolled multicenter pragmatic trial	• 206 community-dwelling older adults • Mean age: 85 (SD 8) y • Gender: 78% women • Age eligibility: ≥75 years • Health status: community-dwelling older adults with mild or moderate dependency (GIR 3 or 4) • Receiving care from 109 home aides	• eHealth system with a machine learning algorithm for predicting ED[b] visits • Decision support for predicting and reducing emergency department visits among older adults, patient monitoring, and optimizing care pathways
Boggiss et al (2023) [38]	Aotearoa, New Zealand, pediatric diabetes clinics and online communities	Qualitative, online focus groups, and one-on-one interviews	Total: 30 • Adolescents: 19 (adolescents aged 12-16 y with type 1 diabetes) • 58% female, 68% Aotearoa New Zealand European, 11% Māori • Living in Aotearoa, New Zealand • Diagnosis of type 1 diabetes for more than 6 months • No serious developmental or psychiatric disorders • Diabetes health care professionals: 11 (health care professionals from various disciplines, such as diabetes nurse specialists, health psychologists, dieticians, and endocrinologists)	COMPASS Chatbot (Expert System) • Psychological support and self-management assistance for adolescents with type 1 diabetes
Castagno and Khalifa (2020) [37]	London, United Kingdom, Royal Free London NHS Foundation Trust hospitals	Qualitative, web-based survey	Total: 98 • Medical doctor: 34 • Nurse: 23 • Manager: 11 • Therapist: 7 • Other: 23 • No demographic information was collected	General AI applications in medicine.

XSL•FO
RenderX

Study (year)	Country and setting	Study design	Participants	AI[a] and technology focus
Catalina et al (2023) [28]	Central Catalonia, Spain, primary care settings	Quantitative, observational cross-sectional study using a validated survey	Total: 301 • Mean age: 46.0 (SD 11.0) year • Sex: female (81.1%), male (18.9%) • Occupation: Nursing (47.8%), Medicine (48.5%), Others (3.65%) • Years of experience: More than 10 years (62.5%)	AI as a health care tool, with a specific focus on its impact on PC radiology.
Gardner and Lundsgaarde (1994) [24]	United States, LDS Hospital	Quantitative, observational study using a questionnaire survey	Total: 1320 • Physicians: 360 (aged 27-82 y, mean age 45.5 y) • Nurses: 960 (aged 20 to 67 y, mean age 33.57 y) • Nurses average professional experience of 9.52 years, with 6.96 years at LDS Hospital	Health Evaluation through Logical Processing (HELP). Expert system • Decision support, patient monitoring, and documentation assistance.
Ginestra et al (2019) [29]	Philadelphia, United States, 782-bed academic hospital	Mixed methods, prospective observational study. Surveys were deployed to collect qualitative data on clinicians' perceptions, and quantitative analysis was performed on the survey results	Total: 287 • Nurses: 180 • Providers: 107	Early Warning Score tool (EWS 2.0). Machine Learning. Decision Support and patient monitoring for sepsis.
Gonçalves et al (2020) [32]	Brazil	Experience report, direct observation, and semistructured interviews	Not clearly mentioned	Robot Laura. Machine learning. Decision support for identification of sepsis.

XSL•FO
RenderX

Study (year)	Country and setting	Study design	Participants	AI[a] and technology focus
Haugsten et al (2023) [39]	Denmark	Qualitative, focus group interviews, and thematic analysis	Total: 14 • Doctor: 8 • Nurse: 6 • Regularly work in a pigmented lesions clinic	ATBM Master (FotoFinder), deep learning, assessment and diagnosis of skin lesions suspicious of melanoma through total body dermoscopy and estimation of malignancy probability.
Helman et al (2022) [33]	Pittsburgh, United States	Qualitative, focus group interviews	Total: 23 • Nurses: 14 • Nurse Practitioners: 4 • Physicians: 5 • Median age: approximately 35 years • Sex: 60% female • Median clinical experience: 8 years	Predictive algorithm for cardiorespiratory insufficiency (CRI) as part of an intelligent decision support system, machine learning, decision support, and patient monitoring.
Im and Chee (2006) [25]	United States	Mixed methods, prospective observational study, and questionnaires	Total: 122 nurses • Mean age: 40.26 years • Predominantly female (93%) • Majority White (75%) • 40% have graduate degrees • 49% Protestant • 59% married	Decision Support Computer Program (DCSP) for cancer pain management. Expert system.
Jauk et al (2021) [34]	Austria	Mixed methods. Convergent parallel design combining questionnaire-based assessments and expert group meetings.	Total: 47 • Nurses: 37 • Physicians: 10 • Sex: male (14), female (33) • Age: median 29 (IQR 26 - 42) years	Predictive algorithm for delirium management decision support. Machine learning.
Jordan et al (2023) [35]	United States	Qualitative, single-site. Small group and individual semistructured interviews and comparative analysis	Total: 13 • ED triage nurses • 72.7% female, 27.3% male • Age ranges: 45.5% aged 25 - 34 years, 27.3% aged 35 - 44 years, 18.2% aged 45 - 54 years, 9.1% aged 55 - 64 years • Highest educational degree in nursing: 9.1% nursing diploma, 36.4% associate, 45.5% bachelor, 9.1% master's • Employment status: 9.1% per diem, 27.3% part-time, 63.6% full-time • Years of nursing experience: mean 11.5 (SD 10.5, range 1 - 35) years	KATE, decision support for emergency nursing triage to improve patient acuity determination. Machine learning.
Koech et al (2022) [40]	Kenya	Qualitative, cross-sectional descriptive, interviews, and focus groups		TraCer Device, gestational age assessment. Computer vision.

XSL•FO
RenderX

Study (year)	Country and setting	Study design	Participants	AI[a] and technology focus
			Total: 70 • Focus group discussions: 52 • In-depth interviews: 18 • Health care workers: diverse group of professionals from antenatal care (ANC) clinic, maternity, outpatient, and radiology departments • Community members: pregnant women, partners, and parents • Sex: health care workers—7 male, 11 female; Community members—15 male partners, 37 female (pregnant women and mothers) • Age groups: health care workers—<35 years (n=7), 35 - 44 years (n=6), 45+ years (n=5); Community members—<35 years (n=30), 35 - 44 years (n=8), 45+ years (n=14)	
LeBaron and Wang (2023) [41]	United States	Mixed methods, surveys	Total: 40 • Clinicians: 5 • Nursing students: 19 • Medical students: 16	CommSense, real-time feedback to clinicians on communication performance metrics such as medical jargon, interruptions, and speech dominance to improve patient-clinician communication and health care delivery. Natural language processing.
Lintz (2023) [26]	Texas, United States, Rural medical center	Quantitative, cross-sectional questionnaire	Total: 48 • Returned questionnaires: 36 • Majority younger than 39 years old • 70% physicians, 20% registered nurses, 10% other kinds of providers • Male: 58%, female: 42%	Hand hygiene monitoring system. Machine learning.
Petitgand et al (2020) [30]	Quebec	Qualitative. Semistructured interviews, informal conversations, nonparticipant observations, with thematic content analysis	Total: 20 • DSS developers: 5 • AHC managers: 5 • Emergency physicians: 7 • Nurses: 3	Decision support for diagnostic decision-making in emergency care. Deep learning.
Rui et al (2023) [42]	Massachusetts, United States	Mixed methods. Usability study including observations, think-aloud		DynaMed and Micromedex with Watson (DynaMedex). Decision support for clinicians related to drug and disease management. Natural language processing and machine learning.

XSL•FO
RenderX

Study (year)	Country and setting	Study design	Participants	AI[a] and technology focus
			Total: 43 • Physicians: 21 • Pharmacists: 14 • Registered nurses: 2 • Nurse practitioners: 1 • Physician's assistants: 5 • Specialties: internal medicine, neurology, cardiology, oncology or hematology, infectious diseases, endocrinology • Average years of practice: 10 years	
Sandhu et al (2020) [36]	Duke University Hospital, United States	Qualitative. Semistructured interviews, modified grounded theory approach	Total: 15 • ED Physicians: 7 • RRT Nurses: 8	Sepsis Watch. Decision support for early sepsis detection and monitoring in emergency departments. Machine learning.
Schwartz et al (2022) [43]	Massachusetts, United States	Qualitative. Semistructured interviews with content analysis	Total: 17 • Prescribing providers: 9 • Nurses: 8	CONCERN, predictive clinical decision support for in-hospital deterioration.

[a]AI: artificial intelligence.

[b]ED: emergency department.

Quality Appraisal

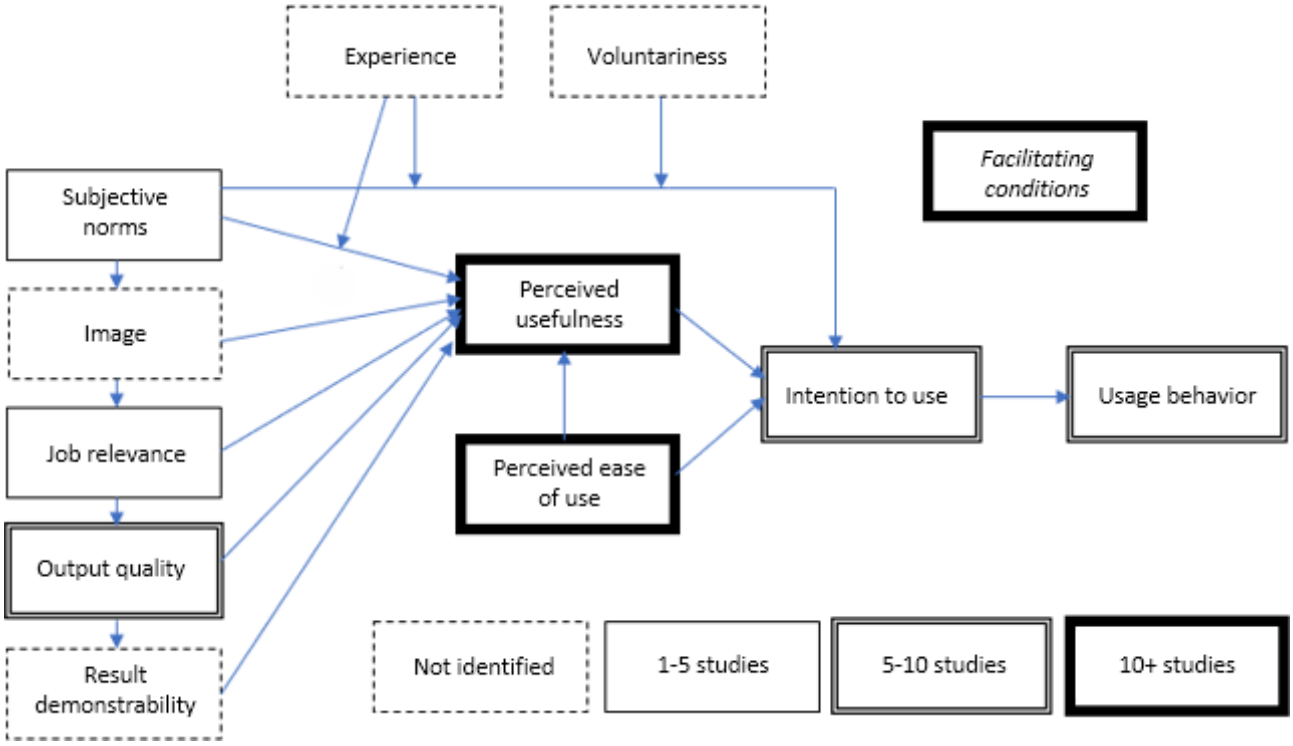
Using the MMAT [22], adherence to quality criteria ranged from strong methodological coherence [31,33] to studies with notable limitations such as unclear sampling strategies [27], lack of representativeness [26], and insufficient reporting of nonresponse bias [24,25]. Common issues included limited discussion of divergences between qualitative and quantitative findings in mixed methods studies [29,34], and a reliance on single data sources [37]. Detailed MMAT findings for each study are presented in Table S1 in Multimedia Appendix 4.

Synthesis of Findings

Findings from the included studies, grouped by TAM2 constructs, are summarized below and presented in detail in Multimedia Appendix 2. Figure 2 illustrates a conceptual map of the TAM2 and UTAUT constructs weighted by frequency of them being identified in the included studies.

XSL•FO
RenderX

Figure 2. Conceptual map of constructs weighted by frequency.



Perceived Usefulness

A total of 17 studies reported nurses' perspectives of AI as enhancing the effectiveness and quality of their work [24-27,29,31,33-43]. Reported benefits included earlier detection of clinical deterioration, improved triage and decision-making [29,35,36], more structured and efficient documentation [24,43], and enhanced patient-clinician communication [41]. As 1 nurse explained,

> [ChatGPT can help me in learning new developments in the telenursing process, especially in relation to the standards of practices, treatment or monitoring procedures. This helps me to improve my skills and capabilities and be up-to-date with the developments in the telenursing sectors, and provide high quality care to the patients remotely.] [31]

In some cases, AI tools contributed to professional confidence and visibility [32,35], with nurses describing the technology as a "safety net" in relation to their decision-making [35]. Perceived usefulness was mixed where the tangible benefits to workflow were unclear or where implementation was limited [34,39]. Furthermore, 3 contributing studies [26,34,37] involved multidisciplinary samples from which nurse-specific data could not be disaggregated; these studies are highlighted in Multimedia Appendix 2, and the nurse-attributed findings above are therefore drawn predominantly from the remaining 14 studies.

Facilitating Conditions

In total, 16 studies reported on facilitating conditions [24-26,28-31,33,35-41,43]. Common themes included access to adequate training, system interoperability, hardware availability, data privacy, and sufficient time within workflows. Willingness to adopt AI was often expressed as being conditional on meeting these facilitating factors [26,28,40]. For example, 1 nurse noted that,

> *Training was yet another organizational matter… not all participants had taken part in these sessions.* [39]

Nurses also requested integrated feedback mechanisms, such as dashboards and trend data [33,41], which highlights the importance of embedding technical support and continuous learning into implementation strategies.

This construct drew contributions from the broadest quality range of any major theme, including both low-rated studies [29,30] and 4 medium-rated surveys. Petitgand et al [30] additionally only involved 3 nurse participants within a predominantly physician sample. While the consistency of facilitating conditions findings across a large number of high-quality studies lends the construct overall coherence, the specific findings from lower-quality studies should be treated as indicative rather than definitive, and the low-rated studies should be interpreted in the context of their methodological limitations as documented in Multimedia Appendix 4.

Perceived Ease of Use

A total of 11 studies addressed perceived ease of use [24-26,28,30,34,36,38,40,42,43]. Nurses valued clear, straightforward interfaces and outputs [36,42], with structured displays facilitating interpretation [24]. As 1 nurse commented,

> *The nurses commented that their computer-printed charts were readable and understandable.* [24]

Barriers included slow system responses, limited interoperability, and difficulty editing records [24,26], as well as challenges interpreting model outputs, with nurses questioning "what have you done to change our processes? What have you done to improve the medical history?" [30].

XSL•FO
RenderX

Several studies noted that ease of use was closely tied to training and resources [28,38,40], which suggests that structured exposure supports both usability and uptake. Moreover, 4 of the contributing studies were medium-rated quantitative surveys, and 1 was rated low [30]; meaning findings on ease of use from these studies are less methodologically robust than those drawn from the high-rated qualitative studies, which provide the primary evidential basis for this construct.

Intention to Use

In total, 9 studies described nurses' intentions to use AI tools [24,27,33,35,39-43]. Intentions were generally reported as positive, with participants expressing interest in continuing or expanding use of AI, particularly with tools that supported decision-making, communication, or workflow efficiency. However, this willingness to adopt was often conditional on receiving adequate training and support, reflecting the critical role of preparatory education in facilitating adoption. All 9 contributing studies reported nurse-specific findings on intention to use, providing a clear nurse-attributable evidence base for this construct; quality ratings of contributing studies were predominantly high (7 high and 2 medium).

Use Behavior

In total, 8 studies described nurses' actual AI system use [24,26,27,29,34-36,43]. Usage ranged from integration into usual practice [36] to limited uptake during pilot testing [34]. Where AI was used, nurses described making changes to patient management and blending system outputs with their own expertise and clinical judgment [27,29]. Lower usage was usually linked to technical limitations, time constraints, or incomplete integration with existing systems. Of the 8 contributing studies, 6 reported nurse-specific use patterns; in 2 studies [26,34], profession-specific use could not be disaggregated and findings reflect aggregated clinician use behavior.

Output Quality

A total of 8 studies addressed perceptions of output quality [26,27,31,32,34,35,38,43]. High-quality outputs were described as accurate, relevant, and timely, supporting decision-making and patient care [27,35]. In contrast, neutral or mixed views emerged when outputs were perceived as unclear, lacked contextualization, or when usability issues were present [26,34]. Of the 8 contributing studies, 2 [26,34] involved multidisciplinary samples in which nurses' perceptions of output quality could not be separated from those of other clinical groups; the findings on output quality therefore reflect a combination of nurse-specific and broader clinician views.

Job Relevance

Of the total, 4 studies talked about job relevance [31,32,35,38]. Nurses could see a clear alignment with some core professional tasks, such as maintaining up-to-date clinical knowledge [31], supporting patient self-management [38], aiding early identification of sepsis [32], and enhancing triage decisions without replacing professional nursing expertise [35]. AI systems that aligned with existing nursing responsibilities appeared to reinforce acceptance and perceived value. All 4 contributing studies [31,32,35,38] were appraised as high quality on the MMAT and reported job relevance findings specific to nurses, providing the most consistently nurse-attributable evidence base of any construct in this review, albeit drawn from a small number of studies.

Subjective Norms

Of all, 2 studies contained discussion of subjective norms [27,32]. Positive attitudes toward AI among peers and team leaders were associated with greater acceptance, while involvement in the development process contributed to professional satisfaction and perceived status. These findings highlight the influence of workplace culture and peer endorsement in shaping AI adoption. Given that this construct was evidenced by only 2 studies, 1 high-quality qualitative study [32] and 1 medium-quality quantitative study [27], the latter with participant sample limitations including a low response rate, and given the limited evidential base overall, these findings should be treated as preliminary, with low confidence, and interpreted with caution.

Across the included studies, perceived usefulness and facilitating conditions emerged most frequently, highlighting nurses' focus on whether AI tools tangibly improve patient care and whether the necessary infrastructure and education are in place to support their integration. Perceived ease of use was also a fairly common theme, closely tied to training and ongoing support, suggesting that structured exposure to AI systems is essential for successful adoption. Job relevance and output quality were less frequently discussed but were positive when present, which reinforces the importance of aligning AI applications with core nursing tasks, as well as the need for them to deliver accurate, contextually relevant outputs. In contrast, subjective norms were only sparsely reported, indicating limited exploration of how peer attitudes and organizational culture shape adoption.

Discussion

Overview

This systematic review synthesized findings from 20 studies reporting RNs' experiences of AI in clinical practice. Nurses described AI as having the potential to support decision-making, workflow efficiency, and clinical confidence, although these reported benefits were consistently contingent on implementation quality. Perceived usefulness and facilitating conditions emerged as the most frequently reported constructs, with ease of use, workflow alignment, and organizational support recurring as determinants of whether nurses experienced AI as a benefit or a burden. Findings most directly attributable to nurses come from studies with majority nurse samples (notably for Job Relevance and several studies contributing to Perceived Usefulness), whereas constructs such as Output Quality, Subjective Norms, and elements of Facilitating Conditions drew partially on multidisciplinary samples; these are interpreted with caution below. Confidence in each finding is therefore not uniform; Job Relevance and Output Quality are anchored by high-quality studies, Perceived Usefulness is supported by a mostly high-quality evidence base, Perceived Ease of Use and Facilitating Conditions draw on a wider range of methodological quality, and Subjective Norms rests on the smallest and most heterogeneous evidence base.

XSL•FO
RenderX

Principal Results

Across the 20 included studies, nurses generally viewed AI as having the potential to improve clinical workflows and to contribute positively to care delivery, although the breadth of these perceptions varied by AI type and clinical setting. However, challenges were also evident, such as training needs, workload concerns, interoperability issues, and ethical considerations. While constructs such as perceived usefulness and facilitating conditions were frequently discussed, subjective norms and job relevance were less often explored, suggesting underresearched areas with potential to influence adoption.

Comparison With Previous Work

Perceived Usefulness

Findings from this review align with the broader AI in health care literature, in which clinicians have expressed optimism about AI's perceived capacity to support clinical tasks, improve efficiency, and contribute to patient care [44-46]. Nurses in the included studies valued AI for specific applications where they perceived benefits as tangible, such as clinical decision support or rapid information retrieval. However, these perceptions were contingent upon the technology being implemented in a way that integrated meaningfully into existing workflows and practice.

Perceived Ease of Use

Consistent with previous research, user-friendly design was a key determinant of willingness to adopt AI in nursing practice [5,47]. Positive experiences were associated with technologies that nurses described as streamlining workflows and reducing burden [48,49]. Conversely, steep learning curves, significant training demands, and inadequate integration led to frustration and resistance [48,50]. Ethical and social concerns, including privacy, trust, and algorithmic bias, were additional barriers [51,52]. These findings suggest that even with positive initial intentions to use AI, without adequate support, integration, and training, this may not translate into long-term adoption.

Subjective Norms

Although less frequently addressed in the included studies, subjective norms play a recognized role in technology adoption across contexts [53-55]. Positive cultural attitudes, leadership endorsement, and peer support could foster AI uptake among nurses. Educational settings represent an opportunity to shape these norms early by embedding AI exposure into curricula, normalizing its use, and developing confidence before clinical deployment.

Facilitating Conditions

Facilitating conditions emerged as critical to successful implementation. Key barriers included interoperability, data sharing, algorithm transparency, and infrastructure readiness [44]. Education and training were consistently identified as enablers [48,56,57]. Academic institutions can play a pivotal role by equipping future nurses with AI competencies and fostering a culture of informed, confident use. Embedding AI in training environments may help bridge the gap between intention and practice.

Job Relevance

While discussed less frequently, job relevance is important for long-term adoption. Concerns included overreliance on AI potentially diminishing nursing roles and the need to ensure AI complements rather than replaces nursing judgment. A proactive approach, where nurses adopt a positive stance toward technological integration [58], may enhance perceived relevance and strengthen the profession's role in shaping AI use.

Implications

Policy and Practice

Findings from this review reinforce that nurses view AI as a tool to supplement rather than replace clinical decision-making, a perspective echoed in previous research [1,59]. Buchanan et al [59] further call for urgent curricular reform to prepare nurses for AI-enabled health care. Embedding AI into nursing education can demystify the technology, highlight its clinical benefits, and provide hands-on experience that builds confidence and competence.

At a policy level, priorities should include establishing clear interoperability standards, developing robust ethical guidance, and ensuring ongoing professional development for AI use in nursing. At a practice level, health care organizations should actively involve nurses in the design, testing, and evaluation of AI tools to ensure they meet usability and clinical relevance standards. Structured training programs can reduce resistance, improve confidence, and normalize AI use before nurses enter practice, which could help to translate positive perceptions into long-term adoption in real-world clinical settings.

Nursing Research

The findings of this review highlight several key areas for future nursing research. First, there is a need for studies exploring the optimal integration of AI systems into nursing workflows, focusing on usability, interoperability, and the impact on patient care outcomes. Research should investigate how AI can supplement clinical decision-making without replacing nursing judgment, which may go some way to addressing concerns about overreliance on technology. Second, given the critical role of education in fostering AI adoption, studies should examine effective strategies for incorporating AI education into nursing curricula and continuing professional development. This includes exploring the use of innovative educational tools to prepare nurses for the evolving health care landscape. Third, research is needed on the long-term effects of AI integration on nursing roles, job satisfaction, and patient-nurse relationships. Finally, studies should address the ethical implications of AI in nursing, including issues of data privacy, algorithmic bias, and the potential transformation of nursing practice. By focusing on these areas, future research can guide the development of AI technologies that enhance nursing practice while preserving the core values of compassionate, patient-centered care.

Limitations

A key strength of this review is its comprehensive synthesis across diverse AI applications and settings, supported by a transparent search strategy and appraisal process using the MMAT [22]. Another strength is the structured use of the TAM2

XSL•FO
RenderX

framework [16], which provided a consistent lens for analysis and allowed for thematic comparison across studies, although, as noted below, this carries trade-offs for inductive emergence.

There are 6 notable limitations. First, the included studies encompassed a heterogeneous range of AI technologies and contexts, from sepsis prediction tools to natural language processing documentation aids, which limits the generalizability of specific implementation recommendations. In particular, conclusions regarding workflow alignment and facilitating conditions are likely to be highly context-dependent and should not be applied uniformly across all AI tool types or care settings.

Second, many studies involved multidisciplinary participant samples in which nurses were not always the majority, and in several cases, nursing data were aggregated with data from other professions. Where studies included multidisciplinary samples, nurse-specific data were extracted separately where the primary data permitted, as documented in Multimedia Appendix 2. In a small number of studies where profession-specific disaggregation was not possible, findings may reflect broader health care professional perspectives rather than distinctly nursing experiences, and these cases are flagged explicitly in Multimedia Appendix 2.

Third, the deductive analytical framework carries an inherent trade-off. TAM2 was selected a priori and registered with PROSPERO, which ensures transparency and reduces post hoc analytical drift; however, coding line-by-line into prespecified constructs may have constrained the emergence of nursing-specific themes that did not map cleanly onto these categories. We did not encounter substantial extracts that fell entirely outside the framework, although themes relating to professional identity, relational care, or moral distress may be more likely to be visible under inductive analysis. Dimensions such as professional identity, relational care, ethical responsibility, and moral distress in AI-assisted practice may therefore be underrepresented in the synthesis not because they were absent from the primary literature, but because the framework did not create space for them. Inductive or hybrid reanalysis of this literature could potentially surface a complementary picture and is recommended as a priority for future qualitative synthesis.

Fourth, coding of the included studies was undertaken by a single reviewer (AJSS), with regular consultation and consensus discussion of coding decisions with 2 supervising coauthors (BCB and DD) but without independent double-coding of the data. While the consultative process was designed to support consistent application of the codebook, the absence of independent dual coding means that formal interrater agreement could not be calculated, and we cannot fully exclude the possibility of single-coder interpretative bias. This limitation is partially mitigated by the use of a transparent, preregistered framework and by the iterative supervisory review process, but readers should weigh the synthesis findings with this constraint in mind.

Fifth, although the search syntax was not restricted by language, inclusion criteria were limited to studies published in English. This decision was made on practical grounds of interpretive validity, but we acknowledge that it may have excluded relevant empirical work from non-English speaking health care contexts, which could have captured different cultural and organizational dimensions of nurses' AI experiences.

Finally, the rapid pace of AI development means that nurses' perceptions are likely to evolve over time, potentially affecting the long-term applicability of these findings.

Conclusions

This systematic review synthesizes nurses' experiences with AI in clinical practice, identifying both the opportunities and challenges associated with its integration. While nurses often view AI as a means to enhance workflows, support decision-making, and contribute to patient care, adoption is reported to be hindered by barriers such as usability issues, limited training, and concerns about trust and clinical relevance. It is important to note that this review captures nurses' reported perceptions and experiences, rather than direct evidence that AI improves patient outcomes, workflow efficiency, or clinical safety; such claims require prospective evaluation. Addressing these challenges requires aligning AI development with user needs and embedding robust education and training into both preregistration curricula and continuing professional development. The consistent findings that facilitating conditions are the primary determinant of positive experience suggest potential value in involving nurses meaningfully in the co-design and iterative refinement of AI tools; this inference is drawn from the weight of evidence on implementation prerequisites rather than direct evidence on participatory design processes, which were only sparsely reported in the included literature. Prospective evaluation using nursing-relevant outcomes, including usability, workflow integration, and trust calibration, is needed to establish which specific implementation conditions reliably produce benefit, and to test whether co-design approaches improve adoption in practice.

By closing these gaps, nursing can play a central role in shaping AI tools that enhance, rather than erode, the human elements of care. As AI capabilities evolve, the profession has an opportunity, and a responsibility, to lead their ethical, effective, and patient-centered implementation.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

Full search strategies and supporting materials are provided in Multimedia Appendices 1-4. Additional extracted data are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: AJSS
Formal analysis: AJSS, QZ, BCB, DD
Methodology: AJSS, BCB, DD
Supervision: BCB, DD
Writing - original draft: AJSS
Writing - review & editing: AJSS, QZ, JFP, BCB, DD

Conflicts of Interest

None declared.

Multimedia Appendix 1
Full search strategies.
[DOCX File, 13 KB - nursing_v9i1e91238_app1.docx]

Multimedia Appendix 2
Results summary table.
[XLSX File, 37 KB - nursing_v9i1e91238_app2.xlsx]

Multimedia Appendix 3
Excluded full-text articles and reasons for exclusion.
[DOCX File, 20 KB - nursing_v9i1e91238_app3.docx]

Multimedia Appendix 4
Quality assessment.
[DOCX File, 37 KB - nursing_v9i1e91238_app4.docx]

Checklist 1
PRISMA checklist.
[DOCX File, 277 KB - nursing_v9i1e91238_app5.docx]

References

1. Seibert K, Domhoff D, Huter K, Krick T, Rothgang H, Wolf-Ostermann K. Application of digital technologies in nursing practice: results of a mixed methods study on nurses' experiences, needs and perspectives. Z Evid Fortbild Qual Gesundhwes 2020 Dec;158-159:94-106. [doi: 10.1016/j.zefq.2020.10.010] [Medline: 33223491]
2. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Nursing in the age of artificial intelligence: protocol for a scoping review. JMIR Res Protoc 2020 Apr 16;9(4):e17490. [doi: 10.2196/17490] [Medline: 32297873]
3. El Arab RA, Al Moosa OA, Sagbakken M, et al. Integrative review of artificial intelligence applications in nursing: education, clinical practice, workload management, and professional perceptions. Front Public Health 2025;13:1619378. [doi: 10.3389/fpubh.2025.1619378] [Medline: 40823249]
4. Rosa ND, Vaz TA, Lucena ADF. Nursing workload: use of artificial intelligence to develop a classifier model. Rev Lat Am Enfermagem 2024;32:e4239. [doi: 10.1590/1518-8345.7131.4239] [Medline: 38985046]
5. Seibert K, Domhoff D, Fürstenau D, Biessmann F, Schulte-Althoff M, Wolf-Ostermann K. Exploring needs and challenges for AI in nursing care – results of an explorative sequential mixed methods study. BMC Digit Health 2023 May 4;1(1):13. [doi: 10.1186/s44247-023-00015-2]
6. Epelde F. Revolutionizing patient safety: the economic and clinical impact of artificial intelligence in hospitals. Hospitals 2024 Dec 12;1(2):185-194. [doi: 10.3390/hospitals1020015]
7. Lukkahatai N, Joseph Dino MN, Saligan L. Empowering care: transforming nursing through artificial intelligence. Artif Intell Med Surg 2025. [doi: 10.5772/intechopen.1010323]
8. Yelne S, Chaudhary M, Dod K, Sayyad A, Sharma R. Harnessing the power of AI: a comprehensive review of its impact and challenges in nursing science and healthcare. Cureus 2023 Nov;15(11):e49252. [doi: 10.7759/cureus.49252] [Medline: 38143615]
9. Tiase V, Cato K. From artificial intelligence to augmented intelligence: practical guidance for nurses. Online J Issues Nurs 2021 Sep 30;26(3). [doi: 10.3912/OJIN.Vol26No03Man04]
10. McGrow K. Artificial intelligence in nursing: a journey from data to wisdom. Nursing 2025 Apr 1;55(4):16-24. [doi: 10.1097/NSG.0000000000000165] [Medline: 40122866]

XSL•FO
RenderX

11. Alqaissi N, Qtait M. Knowledge, attitudes, practices, and barriers regarding the integration of artificial intelligence in nursing and health sciences education: a systematic review. SAGE Open Nurs 2025;11:23779608251374185. [doi: 10.1177/23779608251374185] [Medline: 40919284]
12. Joo JY, Liu MF, Ho MH. Nurses' perceptions of artificial intelligence adoption in healthcare: a qualitative systematic review. Nurse Educ Pract 2025 Oct;88:104542. [doi: 10.1016/j.nepr.2025.104542] [Medline: 40925173]
13. Ma J, Wen J, Qiu Y, et al. The role of artificial intelligence in shaping nursing education: a comprehensive systematic review. Nurse Educ Pract 2025 Mar;84:104345. [doi: 10.1016/j.nepr.2025.104345] [Medline: 40168750]
14. Lifshits I, Rosenberg D. Artificial intelligence in nursing education: a scoping review. Nurse Educ Pract 2024 Oct;80:104148. [doi: 10.1016/j.nepr.2024.104148] [Medline: 39405792]
15. Cucci F, Marasciulo D, Romani M, et al. The contribution of artificial intelligence in nursing education: a scoping review of the literature. Nurs Rep 2025 Aug 1;15(8):283. [doi: 10.3390/nursrep15080283] [Medline: 40863670]
16. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
17. Venkatesh V, Morris MG, Davis GB, Davis FD. Unified Theory of Acceptance and Use of Technology. APA PsycTests 2021. [doi: 10.1037/t57185-000]
18. Purssell E, McCrae N. How to Perform A Systematic Literature Review: A Guide for Healthcare Researchers, Practitioners and Students, 1st edition: Springer International Publishing; 2020. [doi: 10.1007/978-3-030-49672-2]
19. Scott A, Pan JF, Dowding D, Brown B. What are the experiences of nurses using artificial intelligence in clinical practice? A systematic review. PROSPERO 2022 [FREE Full text]
20. Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: JBI; 2020.
21. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
22. Hong QN, Fàbregues S, Bartlett G, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. EFI 2018 Dec 18;34(4):285-291. [doi: 10.3233/EFI-180221]
23. Hannes K, Lockwood C. Pragmatism as the philosophical foundation for the Joanna Briggs meta-aggregative approach to qualitative evidence synthesis. J Adv Nurs 2011 Jul;67(7):1632-1642. [doi: 10.1111/j.1365-2648.2011.05636.x] [Medline: 21466579]
24. Gardner RM, Lundsgaarde HP. Evaluation of user acceptance of a clinical expert system. J Am Med Inform Assoc 1994;1(6):428-438. [doi: 10.1136/jamia.1994.95153432] [Medline: 7850568]
25. Im EO, Chee W. Nurses' acceptance of the decision support computer program for cancer pain management. Comput Inform Nurs 2006;24(2):95-104. [doi: 10.1097/00024665-200603000-00008] [Medline: 16554693]
26. Lintz J. Provider satisfaction with artificial intelligence-based hand hygiene monitoring system during the COVID-19 pandemic: study of a rural medical center. J Chiropr Med 2023 Mar 27;22(3):197-203. [doi: 10.1016/j.jcm.2023.03.004] [Medline: 37360943]
27. Belmin J, Villani P, Gay M, et al. Real-world implementation of an eHealth system based on artificial intelligence designed to predict and reduce emergency department visits by older adults: pragmatic trial. J Med Internet Res 2022 Sep 8;24(9):e40387. [doi: 10.2196/40387] [Medline: 35921685]
28. Catalina QM, Fuster-Casanovas A, Vidal-Alaball J, et al. Knowledge and perception of primary care healthcare professionals on the use of artificial intelligence as a healthcare tool. Digit Health 2023;9:20552076231180511. [doi: 10.1177/20552076231180511] [Medline: 37361442]
29. Ginestra JC, Giannini HM, Schweickert WD, et al. Clinician perception of a machine learning-based early warning system designed to predict severe sepsis and septic shock. Crit Care Med 2019 Nov;47(11):1477-1484. [doi: 10.1097/CCM.0000000000003803] [Medline: 31135500]
30. Petitgand C, Motulsky A, Denis JL, Régis C. Investigating the barriers to physician adoption of an artificial intelligence-based decision support system in emergency care: an interpretative qualitative study. Stud Health Technol Inf 2020;270:1001-1005. [doi: 10.3233/SHTI200312]
31. Alanzi TM. Impact of ChatGPT on teleconsultants in healthcare: perceptions of healthcare experts in Saudi Arabia. J Multidiscip Healthc 2023;16:2309-2321. [doi: 10.2147/JMDH.S419847] [Medline: 37601325]
32. Gonçalves LS, Amaro MDM, Romero ADL, Schamne FK, Fressatto JL, Bezerra CW. Implementation of an artificial intelligence algorithm for sepsis detection. Rev Bras Enferm 2020;73(3):e20180421. [doi: 10.1590/0034-7167-2018-0421] [Medline: 32294705]
33. Helman S, Terry MA, Pellathy T, et al. Engaging clinicians early during the development of a graphical user display of an intelligent alerting system at the bedside. Int J Med Inform 2022 Mar;159:104643. [doi: 10.1016/j.ijmedinf.2021.104643] [Medline: 34973608]
34. Jauk S, Kramer D, Avian A, Berghold A, Leodolter W, Schulz S. Technology acceptance of a machine learning algorithm predicting delirium in a clinical setting: a mixed-methods study. J Med Syst 2021 Mar 1;45(4):48. [doi: 10.1007/s10916-021-01727-6] [Medline: 33646459]

XSL•FO
RenderX

35. Jordan M, Hauser J, Cota S, Li H, Wolf L. The impact of cultural embeddedness on the implementation of an artificial intelligence program at triage: a qualitative study. J Transcult Nurs 2023 Jan;34(1):32-39. [doi: 10.1177/10436596221129226] [Medline: 36214065]
36. Sandhu S, Lin AL, Brajer N, et al. Integrating a machine learning system into clinical workflows: qualitative study. J Med Internet Res 2020 Nov 19;22(11):e22421. [doi: 10.2196/22421] [Medline: 33211015]
37. Castagno S, Khalifa M. Perceptions of artificial intelligence among healthcare staff: a qualitative survey study. Front Artif Intell 2020;3:578983. [doi: 10.3389/frai.2020.578983] [Medline: 33733219]
38. Boggiss A, Consedine N, Hopkins S, et al. Improving the well-being of adolescents with type 1 diabetes during the COVID-19 pandemic: qualitative study exploring acceptability and clinical usability of a self-compassion chatbot. JMIR Diabetes 2023 May 5;8:e40641. [doi: 10.2196/40641] [Medline: 36939680]
39. Haugsten ER, Vestergaard T, Trettin B. Experiences regarding use and implementation of artificial intelligence-supported follow-up of atypical moles at a dermatological outpatient clinic: qualitative study. JMIR Dermatol 2023 Jun 23;6:e44913. [doi: 10.2196/44913] [Medline: 37632937]
40. Koech A, Musitia PM, Mwashigadi GM, et al. Acceptability and feasibility of a low-cost device for gestational age assessment in a low-resource setting: qualitative study. JMIR Hum Factors 2022 Dec 27;9(4):e34823. [doi: 10.2196/34823] [Medline: 36574278]
41. LeBaron V, Flickinger T, Ling D, et al. Feasibility and acceptability testing of CommSense: a novel communication technology to enhance health equity in clinician-patient interactions. Digit Health 2023;9:20552076231184991. [doi: 10.1177/20552076231184991] [Medline: 37456129]
42. Rui A, Garabedian PM, Marceau M, et al. Performance of a web-based reference database with natural language searching capabilities: usability evaluation of DynaMed and Micromedex with Watson. JMIR Hum Factors 2023 Apr 17;10:e43960. [doi: 10.2196/43960] [Medline: 37067858]
43. Schwartz JM, George M, Rossetti SC, et al. Factors influencing clinician trust in predictive clinical decision support systems for in-hospital deterioration: qualitative descriptive study. JMIR Hum Factors 2022 May 12;9(2):e33960. [doi: 10.2196/33960] [Medline: 35550304]
44. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019 Jan;25(1):30-36. [doi: 10.1038/s41591-018-0307-0] [Medline: 30617336]
45. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med 2019 Oct 29;17(1):195. [doi: 10.1186/s12916-019-1426-2] [Medline: 31665002]
46. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial intelligence in precision cardiovascular medicine. J Am Coll Cardiol 2017 May 30;69(21):2657-2664. [doi: 10.1016/j.jacc.2017.03.571] [Medline: 28545640]
47. Ng ZQP, Ling LYJ, Chew HSJ, Lau Y. The role of artificial intelligence in enhancing clinical nursing care: a scoping review. J Nurs Manag 2022 Nov;30(8):3654-3674. [doi: 10.1111/jonm.13425] [Medline: 34272911]
48. Garvey KV, Thomas Craig KJ, Russell R, Novak LL, Moore D, Miller BM. Considering clinician competencies for the implementation of artificial intelligence-based tools in health care: findings from a scoping review. JMIR Med Inform 2022 Nov 16;10(11):e37478. [doi: 10.2196/37478] [Medline: 36318697]
49. O'Connor S, Yan Y, Thilo FJS, Felzmann H, Dowding D, Lee JJ. Artificial intelligence in nursing and midwifery: a systematic review. J Clin Nurs 2023 Jul;32(13-14):2951-2968. [doi: 10.1111/jocn.16478] [Medline: 35908207]
50. Chew HSJ, Achananuparp P. Perceptions and needs of artificial intelligence in health care to increase adoption: scoping review. J Med Internet Res 2022 Jan 14;24(1):e32939. [doi: 10.2196/32939] [Medline: 35029538]
51. Bélisle-Pipon JC, Couture V, Roy MC, Ganache I, Goetghebeur M, Cohen IG. What makes artificial intelligence exceptional in health technology assessment? Front Artif Intell 2021;4:736697. [doi: 10.3389/frai.2021.736697] [Medline: 34796318]
52. Lainjo B, Tmouche H. The impact and implication of artificial intelligence on thematic healthcare and quality of life. International Journal of Applied Research on Public Health Management 2023 Jan;8(1):1-17. [doi: 10.4018/IJARPHM.318140]
53. Lee C, Wan G. Including subjective norm and technology trust in the technology acceptance model. SIGMIS DATABASE 2010 Nov 23;41(4):40-51. [doi: 10.1145/1899639.1899642]
54. Perry A. Consumers' acceptance of smart virtual closets. Journal of Retailing and Consumer Services 2016 Nov;33:171-177. [doi: 10.1016/j.jretconser.2016.08.018]
55. Schepers J, Wetzels M. A meta-analysis of the technology acceptance model: investigating subjective norm and moderation effects. Information & Management 2007 Jan;44(1):90-103. [doi: 10.1016/j.im.2006.10.007]
56. Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing artificial intelligence training in medical education. JMIR Med Educ 2019 Dec 3;5(2):e16048. [doi: 10.2196/16048] [Medline: 31793895]
57. Russell RG, Lovett Novak L, Patel M, et al. Competencies for the use of artificial intelligence-based tools by health care professionals. Acad Med 2023 Mar 1;98(3):348-356. [doi: 10.1097/ACM.0000000000004963] [Medline: 36731054]
58. Archibald MM, Barnard A. Futurism in nursing: technology, robotics and the fundamentals of care. J Clin Nurs 2018 Jun;27(11-12):2473-2480. [doi: 10.1111/jocn.14081] [Medline: 28940657]
59. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on nursing education: scoping review. JMIR Nurs 2021;4(1):e23933. [doi: 10.2196/23933] [Medline: 34345794]

XSL•FO
RenderX

Abbreviations

AI: artificial intelligence
JBI: Joanna Briggs Institute
MMAT: Mixed Methods Appraisal Tool
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RN: registered nurse
TAM2: Technology Acceptance Model 2
UTAUT: unified theory of acceptance and use of technology

Edited by E Borycki, J Sarvestan; submitted 11.Jan.2026; peer-reviewed by D Rosenberg, J Xi; revised version received 12.May.2026; accepted 14.May.2026; published 25.Jun.2026.

Please cite as:
Scott AJS, Zhao Q, Pan JF, Brown BC, Dowding D
Nurses' Experiences Using AI in Clinical Practice: Systematic Review
JMIR Nursing 2026;9:e91238
URL: https://nursing.jmir.org/2026/1/e91238
doi:10.2196/91238

© Ashley J S Scott, Qimeng Zhao, Jo-Fan Pan, Benjamin C Brown, Dawn Dowding. Originally published in JMIR Nursing (https://nursing.jmir.org), 25.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Facilitating Digital Transformation in Nursing Through Nursing Development Units: Scoping Review

Annabell Gangnus[1*], Dr rer med; Pascal Müller[1*], MSc; Katrin Gayen[1], Dip; Uta Kirchner-Heklau[1], MSc; Sebastian Hofstetter[1,2], PhD; Patrick Jahn[1], Prof Dr

[1]Department of Internal Medicine, Health Service Research Working Group, Acute Care, Faculty of Medicine, University Medicine Halle (Saale), Martin-Luther-University Halle-Wittenberg, Magdeburger Straße 12, Halle (Saale), Germany

[2]Dorothea-Erxleben-Learning-Centre, Faculty of Medicine, University Medicine Halle (Saale), Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

[*]these authors contributed equally

Corresponding Author:
Pascal Müller, MSc
Department of Internal Medicine, Health Service Research Working Group, Acute Care, Faculty of Medicine, University Medicine Halle (Saale), Martin-Luther-University Halle-Wittenberg, Magdeburger Straße 12, Halle (Saale), Germany

Abstract

Background: Nursing Development Units (NDUs) are structured clinical environments designed to enhance professional development, collaboration, and organizational learning. While NDUs have been widely studied for their impact on nursing practice, their role in supporting digital transformation in health care has been less explicitly examined.

Objective: This scoping review aimed to map the characteristics, structures, and processes of NDUs and identify how these elements may relate to digital readiness and the integration of digital tools in nursing practice.

Methods: We conducted a scoping review of 40 publications describing NDUs. Data were coded iteratively to identify patterns at the individual, interpersonal, and organizational levels. Core elements and subcategories of NDUs were mapped, highlighting potential connections to capacities required for digital transformation.

Results: NDUs operate across three levels: (1) individual—fostering professional role development, leadership, and resilience; (2) interpersonal—promoting participatory cultures, collaborative practice, and shared unit visions; and (3) organizational—providing structured frameworks for practice philosophy, role development, resource and change management, and embedded evaluation systems. While most primary publications did not explicitly address digital competencies, the structures, reflective practices, and collaborative processes described represent foundational capacities for engaging with digital tools, innovation, and organizational change.

Conclusions: NDUs are characterized by multilevel structures and processes that enhance professional and organizational capacities. Although explicit evidence on digital transformation is limited, these capacities align with key prerequisites for the adoption and effective use of digital technologies in nursing practice. Further research is needed to examine how NDUs directly support digital innovation in clinical settings.

(JMIR Nursing 2026;9:e89051) doi:10.2196/89051

KEYWORDS

nursing development unit; professional development; innovation in nursing; reflective practice; digital transformation; interdisciplinary collaboration; organizational learning; PRISMA

Introduction

Digital transformation is no longer a future scenario in health care but a present reality driven by artificial intelligence, remote monitoring, electronic health records, and data-informed decision-making [1-4]. While these innovations promise improved care quality, efficiency, and patient engagement, their implementation in nursing faces persistent challenges, including digital burnout, workflow disruption, resistance to change, and insufficient staff-centered design [5,6]. These barriers are not merely technical but rooted in organizational culture, leadership, and professional identity. Consequently, digital transformation in nursing often fails due to misalignment between technological innovations and clinical practice [6]. In this review, digital transformation in nursing is understood as a sociotechnical process. It involves technological adoption, changes in professional roles and competencies, and the adaptation of organizational routines and culture.

Nursing Development Units (NDUs) have emerged as a promising model to foster innovation and support digital

XSL•FO
RenderX

transformation in nursing. NDUs are nursing-led, practice-based innovation hubs designed to integrate research, education, and clinical practice through reflective, participatory, and evidence-informed development cycles [7-9]. Unlike isolated pilot projects, NDUs are embedded within health care organizations and supported by dedicated leadership, resources, and evaluation systems, enabling long-term iterative learning and continuous improvement [10,11]. Their core mechanisms—structured reflection, participatory decision-making, and embedded evaluation—directly address known barriers to digital transformation by empowering nurses as cocreators of practice change.

Despite this potential, NDUs remain undertheorized in the context of digital health. Systematic mapping of their conceptual features, functional roles, and organizational frameworks—particularly regarding how they enable or constrain digital transformation—is lacking. Moreover, terminological variation (eg, "Clinical Development Unit," "Practice Development Unit," and "Nursing Professional Unit") complicates cross-contextual comparisons and limits the transferability of insights [11-13]. This conceptual and methodological heterogeneity highlights a key research gap: the absence of a coherent, internationally applicable framework to understand how NDUs function as catalysts for digital innovation in nursing.

This scoping review aimed to address this gap by systematically mapping the international literature on NDUs, synthesizing their core conceptual features, functional roles, and organizational frameworks. The review also explored how NDUs engage with digital transformation—not as passive recipients of technology but as active spaces of knowledge translation [14], organizational change [15], and co-design [16]. A scoping review methodology was chosen due to the conceptual complexity, terminological diffusion, and limited existing synthesis of the NDU concept [17,18]. This approach allows for a comprehensive, transparent, and structured mapping of the literature, providing a foundation for future implementation models and strategies for digital innovation in nursing.

Methods

This scoping review followed the Joanna Briggs Institute methodology [17] and the framework by Arksey and O'Malley [18], which encompasses 4 phases: identifying relevant publications, selecting publications, extracting data, and synthesizing findings. Reporting adhered to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist to ensure transparency and reproducibility [19].

Search Strategy

A systematic literature search was conducted in May 2025 in the databases MEDLINE (via PubMed), ScienceDirect, EBSCO, Cochrane Library, and Europe PMC. Supplementary searches in Google Scholar and LIVIVO captured gray literature and publications not indexed in biomedical databases. Results were sorted by relevance, and all retrieved records were screened without a stopping rule to maximize sensitivity.

The primary database search used the exact phrase "Nursing Development Unit" to prioritize conceptual precision. No additional filters (eg, publication year, study design, or language) were applied to capture the complete literature on NDUs. This decision was grounded in the historically established use of the term. In the literature, "NDU" often denotes a defined organizational model rather than a generic practice development activity. We anticipated that broader search strings (eg, including "Practice Development Unit" or "Clinical Development Unit") would markedly increase retrieval in the main databases. Much of this literature is conceptually adjacent but heterogeneous (eg, general practice development, quality improvement, or clinical education units) and does not necessarily reflect the NDU model targeted in this review.

To reduce the risk of missing relevant publications due to terminological variation, we complemented the core database search with supplementary searches and screened for records that provided structural or process descriptions consistent with NDUs (eg, "Clinical Development Unit," "Practice Development Unit," "Nursing Clinical Development Unit," and "Nursing Professorial Unit"). The full MEDLINE search syntax is provided in Multimedia Appendix 1.

Publications were included if they described structural, functional, or organizational characteristics of NDUs, as well as development processes and implementation strategies. Publications primarily evaluating effectiveness or outcomes (eg, patient metrics and clinical performance) were excluded as the focus was on conceptual mapping rather than outcome assessment. Disease- or population-specific NDUs were included only if they provided insights into NDU structures or processes. Only English- or German-language publications with available full texts were considered. English was used to ensure coverage of the international literature, whereas German-language publications were included for feasibility given the review team's proficiency. Textbox 1 summarizes the eligibility criteria.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Empirical and conceptual publications with abstracts available
- Language: German or English
- Description of content-related aspects, progressive development process, and implementation strategies of Nursing Development Units (NDUs)

Exclusion criteria

- Disease-specific focus (unless NDU structures or processes were described)
- Evaluation of NDUs (effectiveness or outcomes)
- Full text not available

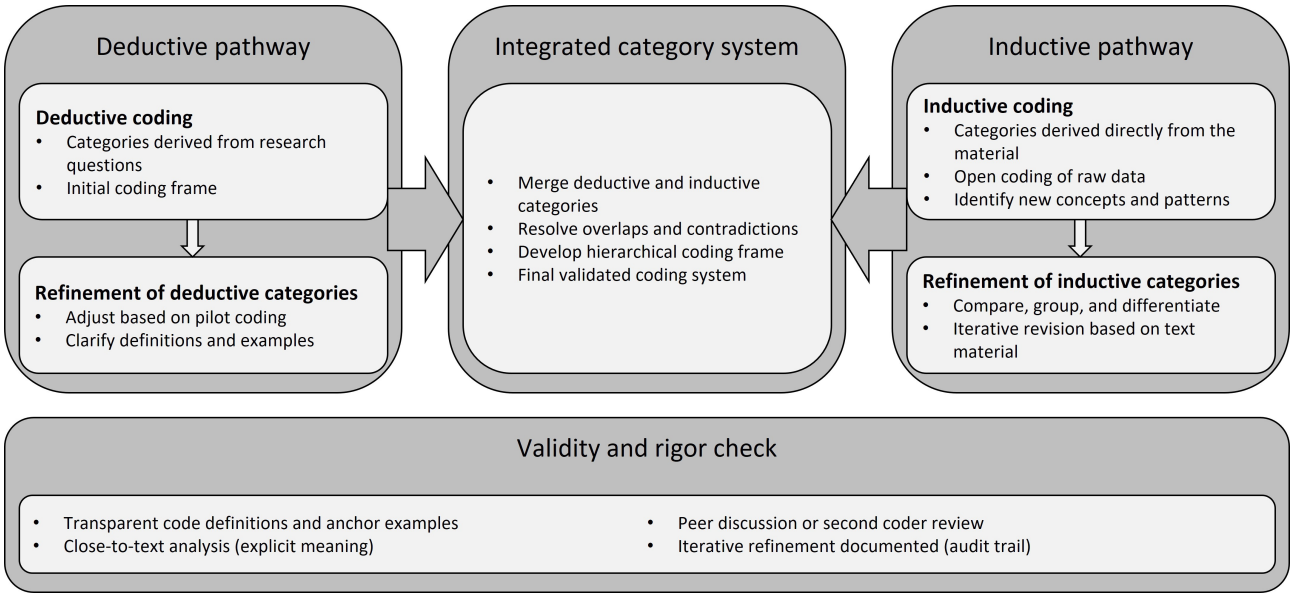
Publication Selection and Data Extraction

Duplicate records were manually removed based on title, author, digital object identifier, journal, and year. Three reviewers (AG, PM, and KG) independently screened titles and abstracts, and any record deemed eligible by at least one reviewer progressed to full-text review. The involvement of 3 reviewers at this stage was chosen to efficiently manage the large number of retrieved records.

Full-text screening was performed independently by 2 reviewers (AG and KG) using predefined inclusion and exclusion criteria, with disagreements resolved through discussion. All included publications were full-text documents (peer-reviewed articles or gray literature reports) that underwent thematic analysis. Data extraction followed a standardized charting framework capturing structural, functional, and organizational characteristics; development and implementation processes; and references to digital transformation where present. Each publication was extracted by a single reviewer, with the framework ensuring consistency. Given the more in-depth nature of full-text assessment, the involvement of 2 reviewers was considered sufficient to ensure consistency and methodological rigor.

A qualitative synthesis was conducted using the structured qualitative content analysis by Kuckartz and Rädiker [20], as recommended by Levac et al [21]. Thematic categories emerged through an iterative process: (1) deductive mapping of initial categories from the research questions (eg, “leadership identity”), (2) inductive coding of all publications to identify recurring patterns, and (3) aggregation of overlapping patterns into the final thematic structure through team consensus. Category definitions, anchor examples, and iterative coding refinement ensured analytical rigor. Figure 1 [20] illustrates the process.

Figure 1. Qualitative content analysis process according to Kuckartz and Rädiker [20] (applied in this review).



Ethical Considerations

As this study was a scoping review of published literature, it did not involve human participants, and therefore, ethics approval was not required.

Results

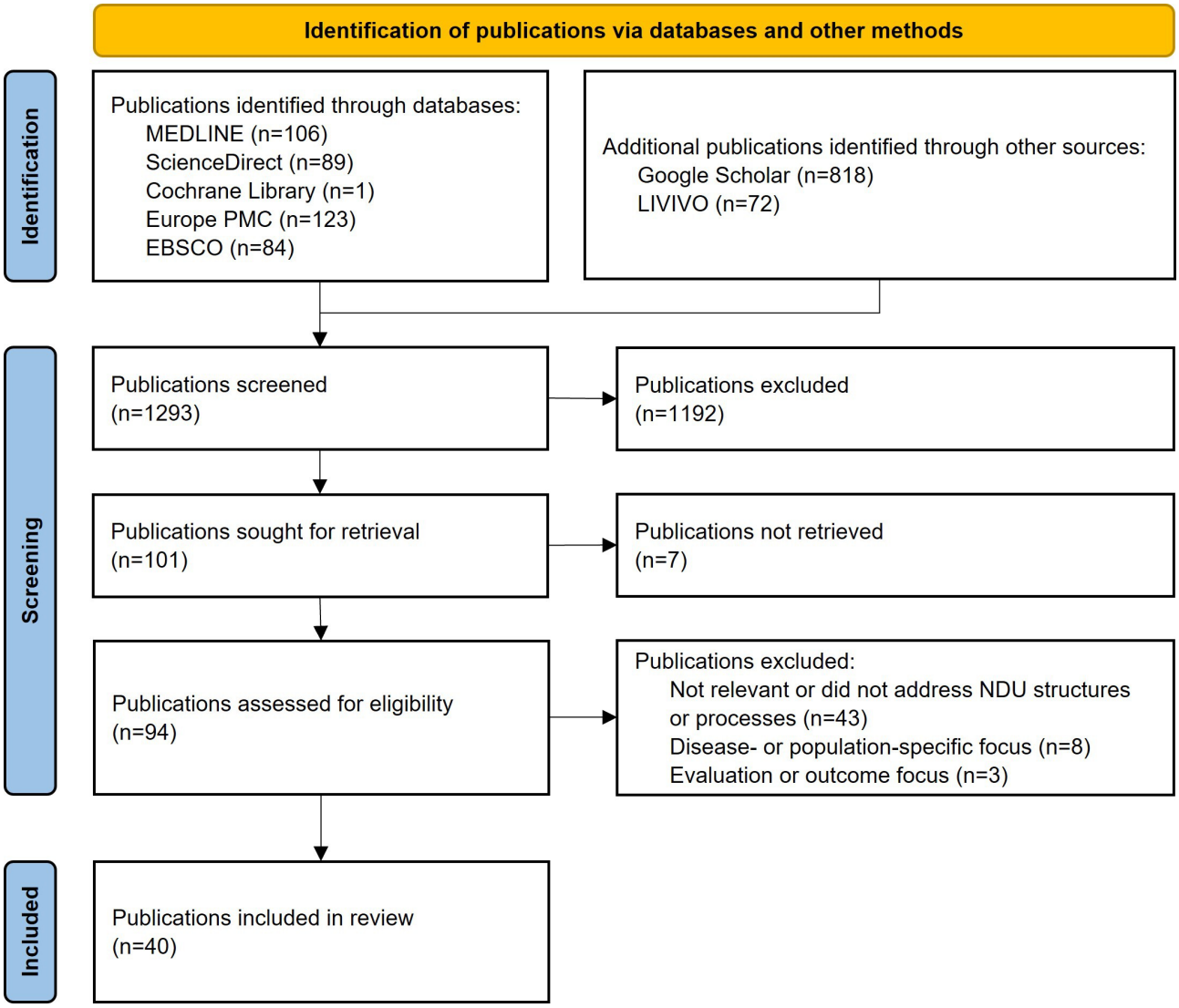
Publication Selection

The database search conducted in May 2025 yielded 1293 records: 106 (8.2%) from MEDLINE (via PubMed), 818 (63.3%) from Google Scholar, 72 (5.6%) from LIVIVO, 89

XSL•FO
RenderX

(6.9%) from ScienceDirect, 1 (0.1%) from the Cochrane Library, 123 (9.5%) from Europe PMC, and 84 (6.5%) from EBSCO. After removing duplicates and screening titles and abstracts, of the 1293 publications, 101 (7.8%) underwent full-text review. Applying the predefined eligibility criteria resulted in 40 publications included for in-depth analysis. No unresolved disagreements occurred during the selection process. The publication selection process is illustrated in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart (Figure 2 [22]).

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart illustrating the study selection process (adapted from Moher et al [22]). NDU: Nursing Development Unit.



Main Characteristics of the Included Publications

The 40 included publications originated from the United Kingdom (n=26, 65%) and Australia (n=14, 35%) and covered the period from 1991 to 2017. Most publications were in English (n=39, 97.5%), with 2.5% (n=1) in German. Study designs comprised qualitative, quantitative, mixed methods, and conceptual approaches. The publications examined how professional roles, leadership, collaborative practices, organizational structures, and reflective processes shape nursing practice. Publication characteristics are summarized in Table 1.

XSL•FO
RenderX

Table . Characteristics of the included publications (N=40).

Study	Country	Design	Focus	Purpose
Atsalos and Greenwood [8], 2001	Australia	Qualitative	Leadership experiences in NDUs[a]	To explore leadership experiences of nurses working in NDUs under conditions of organizational stress
Atsalos et al [23], 2007	Australia	Qualitative	Experiences of nurse leaders in NDUs	To explore the experiences of nurse leaders involved in developing Clinical Development Units into centers of nursing excellence
Avallone and Gibbon [24], 1998	United Kingdom	Cross-sectional survey	Nurses' perceptions of their work environment in an NDU	To examine nurses' perceptions of their work environment within an NDU
Bell and Procter [25], 1998	United Kingdom	Qualitative	Experiences of nurses working in an NDU	To explore nurses' experiences of involvement in nursing research within an NDU
Bland [10], 1997	United Kingdom	Case study	Development of an emergency nurse practitioner role	To describe the development of the emergency nurse practitioner role within an accident and emergency NDU
Bowles and Bowles [26], 2000	United Kingdom	Quantitative	Transformational leadership in NDUs vs conventional units	To compare transformational leadership in NDUs and conventional clinical settings
Cannard [27], 1996	United Kingdom	Quantitative	Aromatherapy to promote relaxation and stress reduction	To describe the implementation of aromatherapy as a practice development initiative within an NDU
Christensen and Craft [28], 2017	Australia	Conceptual	Translating research into nursing practice	To discuss strategies for translating nursing research into practice through an NDU
Christian and Norman [29], 1998	United Kingdom	Qualitative	Clinical leadership in NDUs	To examine approaches to clinical leadership development within NDUs in England
Draper [30], 1996	United Kingdom	Narrative review	Opportunities for evaluation of NDUs	To critically examine the purposes, characteristics, and effectiveness of NDUs
Duffield [31], 2005	Australia	Evaluation study	Masterclass for unit managers	To evaluate the design and delivery of a leadership masterclass for nursing unit managers
Gerrish [7], 2001	United Kingdom	Evaluation study	Evaluation of NDUs	To evaluate a Nursing and Practice Development Unit accreditation program
Gerrish and Ferguson [32], 2000	United Kingdom	Qualitative	Factors influencing NDU progress	To identify factors influencing the development and progress of NDUs
Graham [33], 1996	United Kingdom	Conceptual	Conceptual framework for NDUs	To present a conceptual framework supporting reflective practice and practice development within NDUs
Graham [34], 2000	United Kingdom	Qualitative	Reflective practice in mental health nurses	To examine reflective practice processes among mental health nurses working in an NDU

XSL•FO
RenderX

Study	Country	Design	Focus	Purpose
Graham [35], 2003	United Kingdom	Qualitative	Leadership perspectives in NDUs	To explore leadership perspectives within an NDU from academic and clinical viewpoints
Greenwood [36], 2000	Australia	Qualitative	Clinical Development Units: challenges and issues	To describe the characteristics, achievements, and challenges of Clinical Development Units
Greenwood [12], 1997	Australia	Qualitative	Accreditation of NDUs	To describe NDUs and examine models of accreditation
Greenwood [37], 1999	Australia	Case study	Western Sydney approach	To describe the establishment and leadership preparation approach of a Clinical Development Unit network in western Sydney, New South Wales
Greenwood and Kearns [38], 1996	Australia	Qualitative	Establishing a transcultural NDU	To describe the establishment of a transcultural NDU in an Australian context
Greenwood [13], 2000	Australia	Qualitative	Issues surrounding establishment and survival of NDUs	To examine challenges related to the establishment and sustainability of NDUs
Greenwood and Parsons [39], 2002	Australia	Evaluation study	Evaluation of a leadership preparation program	To evaluate a leadership preparation program for Clinical Development Unit leaders
Happell and Martin [40], 2002	Australia	Qualitative	Changing mental health nursing culture	To describe the implementation of a nursing Clinical Development Unit program in mental health nursing
Happell and Martin [41], 2004	Australia	Evaluation study	Evaluation of the Nursing Clinical Development Unit program	To evaluate outcomes of a Nursing Clinical Development Unit program in mental health nursing
Happell and Martin [42], 2005	Australia	Evaluation study	Changing the culture of mental health nursing	To assess the impact of a Nursing Clinical Development Unit program on mental health nursing culture
Johns [43], 1991	United Kingdom	Conceptual	Holistic model of nursing practice	To present a holistic model of nursing practice developed within an NDU
Keatinge and Scarfe [44], 1998	United Kingdom	Qualitative	NDUs in dementia care	To describe the establishment of an NDU in dementia care
Keatinge et al [45], 2000	United Kingdom	Action research	Nursing management of agitation in institutionalized residents with dementia	To examine nursing management of agitation in dementia care through participatory action research within an NDU
Malby [46], 1996	United Kingdom	Qualitative	Overview of NDUs in the United Kingdom	To provide an overview of the development and implementation of NDUs in the United Kingdom
Manley [47], 1997	United Kingdom	Action research	Advanced practitioner and consultant nurse role	To develop and examine an advanced practitioner role through action research within an NDU context

XSL•FO
RenderX

Study	Country	Design	Focus	Purpose
Parsons and Mott [11], 2003	Australia	Qualitative	Toward Clinical Development Units	To describe the principles and processes underpinning Clinical Development Units
Pearson [48], 1997	United Kingdom	Evaluation study	King's Fund Centre NDU network	To assess the progress and direction of the King's Fund Centre NDU network
Redfern and Stevens [49], 1998	United Kingdom	Qualitative	NDU structure and orientation	To describe the structure, aims, and organization of NDUs
Redfern and Murrells [50], 1998	United Kingdom	Qualitative	Research, audit, and networking activity in NDUs	To compare research, audit, and networking activities between NDUs and non-NDUs
Redfern et al [51], 1997	United Kingdom	Qualitative	Evaluation of NDUs	To evaluate the value and core characteristics of NDUs
Ryan [52], 1994	United Kingdom	Qualitative	Improving discharge planning	To describe practice development activities aimed at improving discharge planning within an NDU
Schiereck [53], 2000	United Kingdom	Qualitative	Social interaction in NDUs	To examine social interactions between nurses and patients within an NDU
Scholes [54], 1996	United Kingdom	Qualitative	Role transition and emotional labor	To explore the impact of working in an NDU on practitioners' role transition and emotional labor
Vaughan [9], 1998	United Kingdom	Qualitative	History of NDU programs	To compare the development trajectories of 2 NDU programs
Wright [55], 2007	United Kingdom	Qualitative	Contribution to quality	To examine the concept of NDUs and their contribution to quality in nursing practice

[a]NDU: Nursing Development Unit.

Core Elements of NDUs

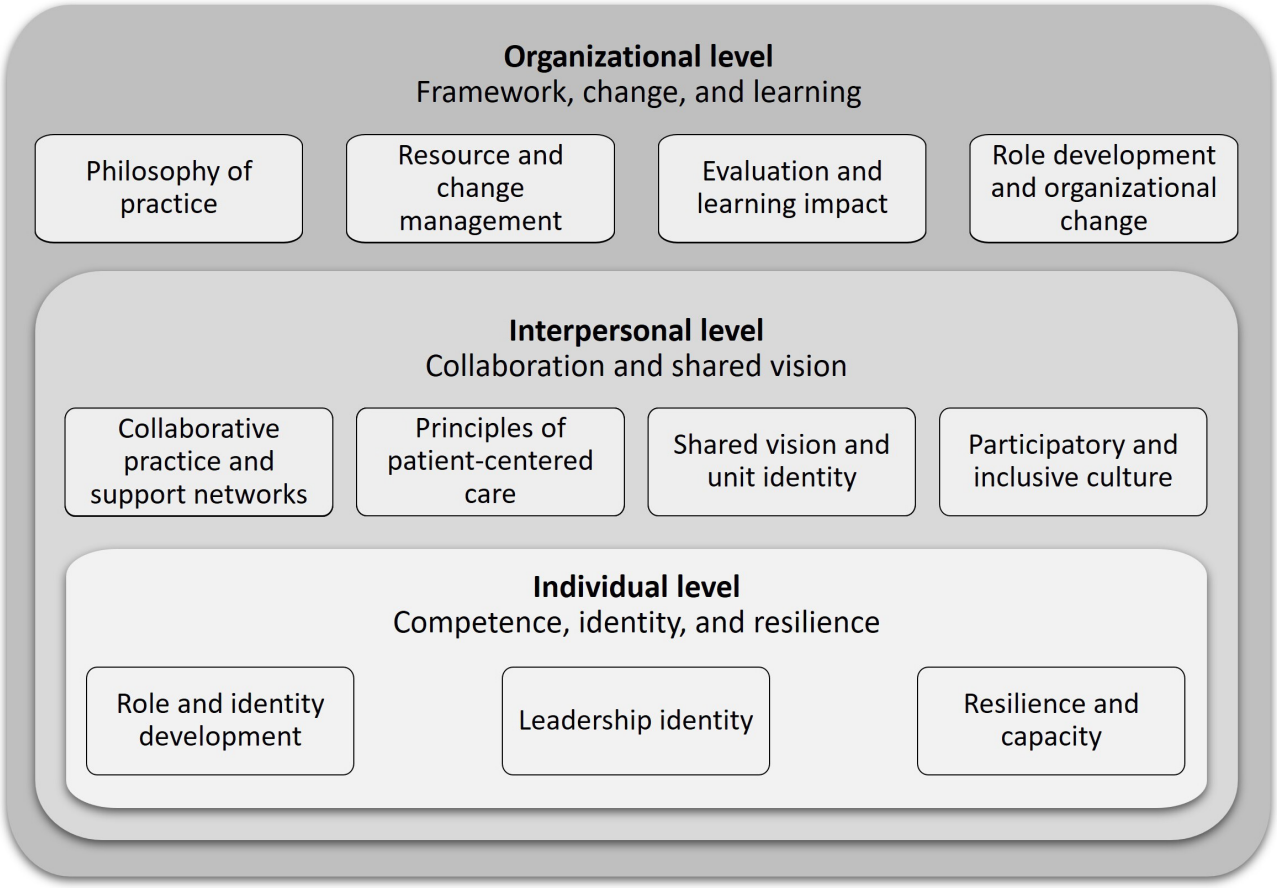
Overview

The literature consistently described NDUs as operating across 3 interrelated levels: individual, interpersonal, and organizational (Figure 3). These levels structure descriptions of professional development, collaboration, and organizational learning. This framework emerged inductively from the data through iterative coding of all included publications, reflecting patterns observed in the literature rather than a pre-established theoretical model.

Several characteristics identified at these levels—such as competence development, participatory practices, and embedded learning structures—are discussed in the broader literature on organizational readiness for change and digital transformation. While explicit engagement with digital technologies was seldom reported in the included publications, the identified capacities and practices indicated potential enablers for adopting digital tools and supporting digital innovation in nursing.

XSL•FO
RenderX

Figure 3. Core elements of a Nursing Development Unit inductively synthesized from the included publications.



Individual Level: Competence, Identity, and Resilience

NDUs were described as settings in which professional roles, leadership capabilities, and personal resilience are addressed. Activities included reflective practice, integration of ethical and evidence-based principles, and adaptation to evolving clinical demands. Table 2 maps publications contributing to individual-level components.

A clear understanding of professional roles, the integration of personal and shared philosophies of practice, and the development of self-efficacy and professional coherence were described as central elements of role and identity development [23,43]. Several publications highlighted intrinsic motivation and self-awareness in relation to personal growth and competence development [23].

Table . Mapping publications to individual-level components of Nursing Development Units (n=17).

Level and subcategory	Publications
Individual	
Role and identity development	[8,10,23,25,33-35,43,45,47,54]
Leadership identity	[8,23,26,29,31,35,39]
Resilience and capacity	[27,34,44,45,54]

Reflective practice was described as integrating ethical considerations, clinical judgment, and personal values into everyday care [43]. The literature reported structured forms of reflection, including written protocols, workshops, and reflective models, which were used to support critical inquiry into values, roles, and routines and articulate emerging professional identities within NDUs [33,34]. Modular training programs addressing clinical and research competencies were described across several publications and commonly linked to supervision and assessment processes [10].

Leadership identity was addressed in a subset of publications (7/17, 41.2%), which described clinical leaders as taking on facilitative and innovative roles and as supporting team processes. At the same time, challenges such as variable staff motivation and uncertainty regarding the NDU concept were reported [8,23]. Leadership support structures, including mentorship, networking, and formal development programs, were described in relation to team support during periods of change [31].

Several publications (5/17, 29.4%) described NDUs as contexts in which professional resilience is addressed, particularly in relation to iterative development processes and challenges associated with time constraints, emotional labor, and organizational obstacles [23,43]. Staff perceptions reported in

XSL•FO
RenderX

the literature included high expectations, impatience for visible progress, and difficulties adjusting to evolving roles [43].

Across the included publications, reflective practice, continuous competence development, and leadership identity were recurrently described at the individual level. While the reviewed publications did not explicitly address digital competencies or digital practices, they outlined professional capacities—such as critical reflection, adaptability, and role clarity—that are frequently discussed in the broader literature in relation to organizational and technological change.

Interpersonal Level: Collaboration and Shared Vision

Across the included publications, NDUs were described as fostering participatory and inclusive cultures, supporting collaborative practices, and facilitating the development of shared visions and unit identities. Several publications reported that team interactions and decision-making processes emphasized principles of patient-centered care. Table 3 summarizes the publications contributing to the interpersonal-level subcategories.

Participatory and inclusive cultures were described as involving collaboration, shared decision-making, and ongoing improvement [32,53]. Publications reported that traditional nursing boundaries were sometimes transcended, enabling staff to engage in collective reflection, articulate unit philosophies, formulate action plans, and evaluate progress. Bottom-up approaches, including staff participation as coresearchers, were described in relation to collective ownership of practice change [32,45].

Table . Mapping publications to interpersonal-level components of Nursing Development Units (n=30).

Level and subcategory	Publications
Interpersonal	
Participatory and inclusive culture	[8,23-25,32,34,40,42,43,53,54]
Collaborative practice and support networks	[24,25,34,37,39,41,45,47,50,52,53]
Shared vision and unit identity	[9,11-13,28,33,36-38,40,42,44,46,49]
Principles of patient-centered care	[8,27,38,43,44,52,53]

Collaborative practices and support networks were described as developing through mentorship and interdisciplinary teamwork, supporting staff in addressing challenges and continuing professional growth [23,49]. Strategic planning processes were reported to clarify team needs, anticipate obstacles, and link practice activities to evidence generation. Several publications described collaborative planning as fostering shared ownership and a collective vision for change [54]. Partnerships with academic institutions were reported to support training, research, and innovation [9,49].

Shared vision and unit identity were described as contributing to team coherence. Staff involvement in defining the NDU's purpose, intended outcomes, and guiding principles was frequently reported [43,45,54]. The literature highlighted that shared philosophies of practice grounded in clinical realities and staff perspectives can reinforce team coherence and collective identity [43]. Exposure to experiences from other NDUs was reported to help manage expectations, facilitate collective problem-solving, and support ongoing development [44].

Patient-centered care was described as emphasizing trust, sensitivity, shared decision-making, and patient autonomy [8,43]. Evidence-based approaches were reported across NDUs, although implementation varied. Some publications reported extensive evaluation frameworks, whereas others focused on routine questioning and evidence-informed decisions [9]. Integration with relational care models such as primary nursing was described as an example of applying patient-centered principles in practice [8,43,53].

Across the reviewed publications, participatory cultures, collaborative planning, and shared decision-making were consistently reported as central interpersonal characteristics of NDUs. Although digital technologies were not explicitly addressed, these interactional patterns reflected approaches commonly discussed in the broader literature for co-designing and implementing complex interventions, including digital tools in health care settings.

Organizational Level: Framework, Change, and Learning

At the organizational level, NDUs were described as involving frameworks and structures that support practice development, change, and learning. Several publications highlighted overarching philosophies of practice, structured role development, and embedded evaluation systems as recurring elements. Table 4 summarizes the publications contributing to the organizational-level subcategories.

XSL•FO
RenderX

Table . Mapping publications to organizational-level components of Nursing Development Units (n=31).

Level and subcategory	Publications
Organizational	
Philosophy of practice	[7,9,11-13,28,33,35-38,40-44,46,48-51,55]
Resource and change management	[9,12,13,26,31,32,36-39,41,44,46,48,49,51,55]
Role development and organizational change	[10,11,29,33,35,37,43,47,51]
Evaluation and learning impact	[7,28,30-32,39,41,48,50-53,55]

Philosophies of practice were described as providing frameworks of values, guiding principles, and standards that inform unit culture and practice [51]. Role development and organizational change were reported in relation to clarifying responsibilities, establishing decision-making processes, and supporting structural adjustments to avoid duplication and facilitate innovation [33,53].

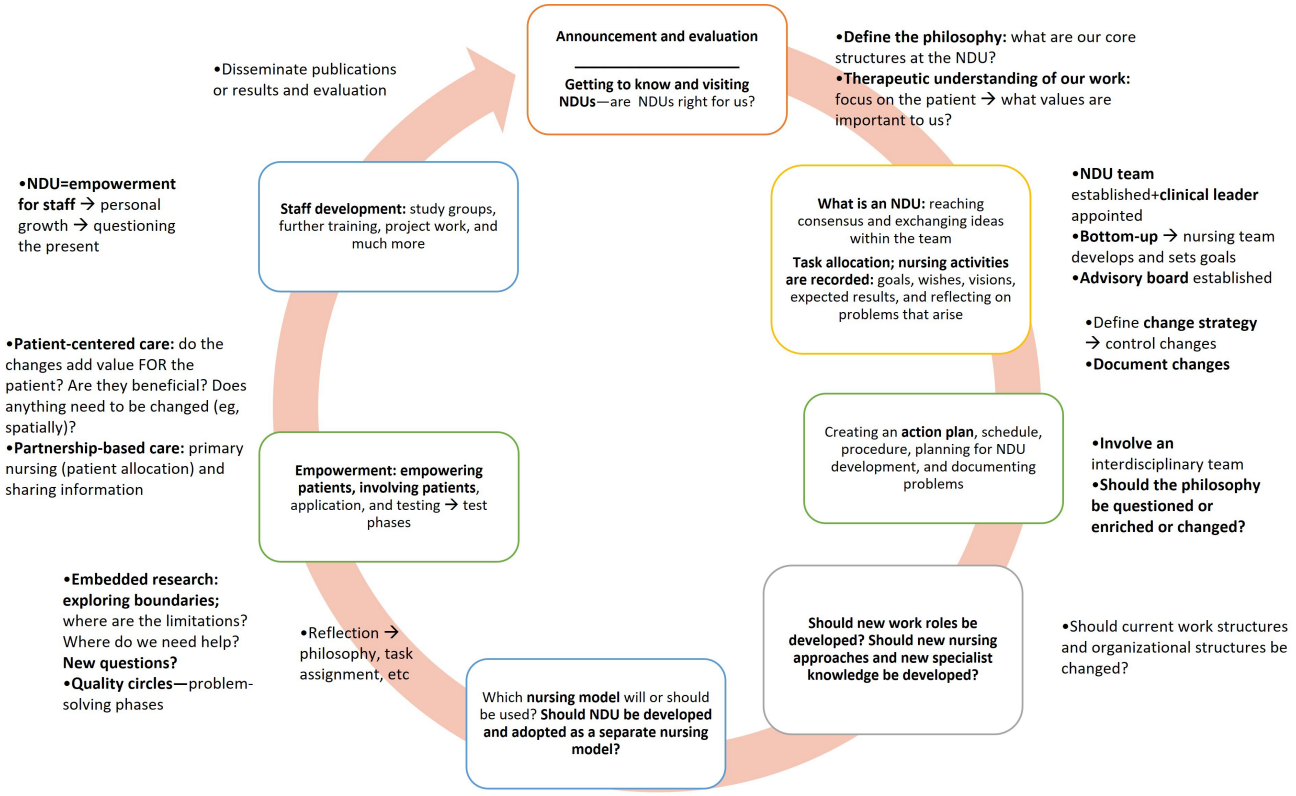
Management of resources and change was described in the literature as including strategies to allocate time and materials, address resistance, and respond to operational pressures [49,53]. Evaluation and learning systems were reported to support reflective processes, facilitate improvements based on evidence, and connect practice to research and learning across disciplines [50,51,53]. Iterative development processes using structured strategies, participatory approaches, and strategic planning were described as approaches to respond to local needs while maintaining growth and improvement [10,11,29,50,51,53,54].

Embedded evaluation systems and iterative learning cycles were frequently reported as integral to NDU development. While these publications did not specifically address digital innovations, the described organizational infrastructures and learning mechanisms were consistent with approaches that are widely considered relevant for implementing and adapting complex interventions, including digital tools in health care settings.

Development and Implementation Process of NDUs

NDU development was reported as a multiyear, iterative process involving alignment of organizational structures, professional identity formation, and collaborative teamwork (Figure 4). Implementation evolves through cycles of reflection, strategic planning, practice-based experimentation, and evidence-informed evaluation grounded in a shared philosophy of practice that incorporates diverse staff perspectives and emphasizes holistic, patient-centered nursing [8,11,43,55].

Figure 4. Progressive development toward a Nursing Development Unit (NDU) synthesized inductively from the included publications.



Implementation practices reported included client-centered care, nurse empowerment, and evidence-based routines [11,29]. Adaptations to local cultural, staffing, and regional conditions were described as supporting flexibility while maintaining core principles [49]. Staff engagement in data collection and analysis was reported to facilitate learning, support innovation uptake, and foster interdisciplinary thinking [50,51,53].

XSL•FO
RenderX

Challenges reported in the literature included staff turnover, limited acceptance by other health care professionals, high workloads, and gaps between theory and practice. Sustained leadership support and team commitment were described as important factors for maintaining NDU development over time [28].

Discussion

Core Characteristics of NDUs: A Foundation for Innovation

This scoping review mapped the structural and procedural characteristics of NDUs across individual, interpersonal, and organizational levels based on 40 international publications from the United Kingdom and Australia spanning 1991 to 2017. The findings confirm NDUs as multilevel, practice-based innovation spaces that foster professional development, collaborative cultures, and structured learning. These characteristics are not merely descriptive but reflect a coherent framework for supporting continuous improvement and evidence-informed practice.

At the individual level, NDUs are not merely settings for skill development but spaces for professional identity formation. The emphasis on reflective practice, leadership identity, and resilience suggests that NDUs support nurses in navigating complex clinical realities and evolving roles [10,23,33,43]. This process fosters psychological safety and self-efficacy—conditions that are increasingly recognized as essential for adaptive learning and innovation [56,57].

At the interpersonal level, NDUs function as cocreation platforms. The documented practices of shared decision-making, collaborative planning, and peer learning indicate a culture where knowledge is coproduced rather than imposed in a top-down manner [23,32,45,50]. This participatory orientation aligns with contemporary models of human-centered design and may reduce resistance to change—especially when introducing new technologies [16,58].

At the organizational level, NDUs are defined by structured learning cycles and embedded evaluation systems. Their iterative development process—rooted in reflection, experimentation, and feedback—mirrors established frameworks for organizational change [10,11,29,50,51,53]. This suggests that NDUs are not static entities but dynamic systems capable of continuous adaptation, a quality that is critical in rapidly evolving digital health environments [59,60].

Crucially, these characteristics are not isolated features but interconnected processes that reinforce one another. For example, individual reflection is amplified through team dialogue, team collaboration is guided by organizational frameworks, and organizational learning is sustained through individual and team engagement. This synergistic interplay positions NDUs as potential organizational incubators for innovation—although the extent to which they support digital transformation remains to be empirically tested.

Implications for Digital Transformation: A Cautious Interpretive Framework

While the included publications did not explicitly address digital technologies, the identified characteristics of NDUs suggest potential conditions that may support digital transformation in nursing. This section presents these connections as interpretive transfer hypotheses—plausible inferences based on conceptual alignment with known prerequisites for successful technology adoption.

NDUs may support digital transformation by fostering the professional and organizational capacities required for technology integration. For instance, reflective practice and continuous competence development—core to NDUs—can help nurses critically engage with digital tools, align them with ethical and clinical values, and integrate them meaningfully into care routines [33,34,43]. This supports the development of a digital professional identity where technology use is not seen as an external imposition but as an extension of nursing expertise [5,6,61,62].

The participatory culture of NDUs—characterized by shared decision-making, co-design, and peer learning—may reduce resistance to change and improve the fit of digital tools with clinical workflows [16,63-65]. By involving nurses in the design, testing, and evaluation of digital solutions, NDUs could serve as safe spaces for experimentation, where failures are reframed as learning opportunities [15,16]. This aligns with human-centered design principles and may enhance user acceptance and long-term sustainability [58].

At the organizational level, the structured frameworks for change management, iterative evaluation, and evidence-based practice in NDUs mirror the adaptive and learning-oriented processes needed for digital implementation [7,15,50,51]. The emphasis on feedback loops, strategic planning, and continuous improvement provides a robust infrastructure for managing the uncertainties and disruptions associated with digital innovation [60,66].

However, these implications must be interpreted with caution. The evidence base is limited in both temporal and technological scope: most included studies (39/40, 97.5%) were published before 2007, and none explicitly examined digital tools or competencies. Therefore, the claim that NDUs "support digital transformation" should not be understood as a proven causal relationship but rather as a plausible hypothesis that requires empirical validation in contemporary digital contexts.

Strengths and Limitations

This review's strength lies in its systematic and transparent approach guided by established frameworks (Joanna Briggs Institute and PRISMA-ScR guidelines). The inclusion of diverse study types and international literature provides a broad understanding of NDU development and implementation. The independent screening and iterative data extraction process enhances rigor, transparency, and generalizability.

Several limitations must be acknowledged. First, while this review focused on publications explicitly labeled as "Nursing Development Unit," it is possible that related concepts or

XSL•FO
RenderX

alternative labels (eg, “Clinical Development Unit” and “Practice Development Unit”) exist in more recent (digital) nursing literature and may share similar mechanisms for supporting innovation and technology adoption. Although supplementary searches were conducted, our decision to anchor the database search to the exact phrase “Nursing Development Unit” may have led to missed publications describing comparable models under different terminology. Consequently, our synthesis should be interpreted as mapping the literature that explicitly self-identified as “NDU” rather than exhaustively capturing all international Practice Development Unit models. Second, the age of the included publications (mostly published before 2007) limits the direct applicability of the findings to current digital health landscapes, where technologies such as artificial intelligence, real-time monitoring, and interoperable electronic health records are increasingly central. Third, the lack of outcome evaluations and the heterogeneity of definitions and implementation contexts reduce the comparability of the findings. Fourth, the exclusion of non–English- and non–German-language publications may have introduced language bias. Fifth, the absence of a formal quality appraisal—typical in scoping reviews that prioritize conceptual mapping over effect estimation—limits the comprehensive assessment of the robustness of individual findings.

These limitations underscore the need for future research to explicitly link NDU principles with digital transformation using contemporary case studies, mixed methods designs, and outcome-focused evaluations.

Conclusions: A Foundation for Future Inquiry

This review demonstrates that NDUs are structured environments that promote professionalization, collaboration, and learning through multilevel processes. While the evidence base does not directly support claims about digital transformation, the core characteristics of NDUs—reflective practice, participatory cultures, and iterative learning—may offer enabling conditions for the adoption and integration of digital technologies in nursing.

However, these connections remain theoretical and require empirical testing. Future research should investigate how NDUs can be adapted to support digital innovation, with a focus on digital competencies, workflow integration, ethical considerations, and staff well-being. Only through such research can we determine whether NDUs are indeed a viable foundation for digital transformation in nursing or whether new models are needed to meet the demands of the digital age.

Funding

This research was funded by the Federal Ministry of Research, Technology, and Space (grant ID 01ZZ2401A). The study has been funded for a 5-year period starting in November 2024. The study was funded for €4,559,301.88 (US $5,291,298). The funder had no involvement in the study design; data collection, analysis, and interpretation; or writing of the manuscript. The authors acknowledge the financial support of the Open Access Publication Fund of Martin Luther University Halle-Wittenberg.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 13 KB - nursing_v9i1e89051_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 104 KB - nursing_v9i1e89051_app2.pdf]

References

1. Canfell OJ, Woods L, Meshkat Y, et al. The impact of digital hospitals on patient and clinician experience: systematic review and qualitative evidence synthesis. J Med Internet Res 2024 Mar 11;26:e47715. [doi: 10.2196/47715] [Medline: 38466978]
2. Schlicht L, Wendsche J, Melzer M, Tschetsche L, Rösler U. Digital technologies in nursing: an umbrella review. Int J Nurs Stud 2025 Jan;161:104950. [doi: 10.1016/j.ijnurstu.2024.104950] [Medline: 39603090]
3. Ruksakulpiwat S, Thorngthip S, Niyomyart A, et al. A systematic review of the application of artificial intelligence in nursing care: where are we, and what’s next? J Multidiscip Healthc 2024;17:1603-1616. [doi: 10.2147/JMDH.S459946] [Medline: 38628616]

XSL•FO
RenderX

4. Chua M, Lau XK, Ignacio J. Facilitators and barriers to implementation of telemedicine in nursing homes: a qualitative systematic review and meta-aggregation. Worldviews Evid Based Nurs 2024 Jun;21(3):318-329. [doi: 10.1111/wvn.12711] [Medline: 38340069]
5. Bimerew M. Barriers and enablers of nurses' adoption of digital health technology to facilitate healthcare delivery in resource-limited settings. Stud Health Technol Inform 2024 Jul 24;315:64-68. [doi: 10.3233/SHTI240107] [Medline: 39049227]
6. Walzer S, Armbruster C, Mahler S, Farin-Glattacker E, Kunze C. Factors influencing the implementation and adoption of digital nursing technologies: systematic umbrella review. J Med Internet Res 2025 Jul 31;27:e64616. [doi: 10.2196/64616] [Medline: 40743516]
7. Gerrish K. A pluralistic evaluation of nursing/practice development units. J Clin Nurs 2001 Jan;10(1):109-118. [doi: 10.1046/j.1365-2702.2001.00439.x] [Medline: 11820228]
8. Atsalos C, Greenwood J. The lived experience of clinical development unit (nursing) leadership in Western Sydney, Australia. J Adv Nurs 2001 May;34(3):408-416. [doi: 10.1046/j.1365-2648.2001.01758.x] [Medline: 11328446]
9. Vaughan B. The story of NDUs — how the nursing, midwifery and health visiting development unit programme began. NT Res 1998;3(4):272-274. [doi: 10.1177/174498719800300406]
10. Bland A. Developing the emergency nurse practitioner role in accident and emergency: a bottom-up approach. Accid Emerg Nurs 1997 Jan;5(1):42-47. [doi: 10.1016/s0965-2302(97)90063-1] [Medline: 9069735]
11. Parsons M, Mott S. Royal Rehabilitation Centre Sydney: towards Clinical Development Units (Nursing). Collegian 2003 Jan;10(1):27-29. [doi: 10.1016/s1322-7696(08)60617-4] [Medline: 15481508]
12. Greenwood J. Should nursing development units be accredited? Collegian 1997;4(1):6-11. [doi: 10.1016/s1322-7696(08)60199-7]
13. Greenwood J. Clinical development units (nursing): Issues surrounding their establishment and survival. Int J Nurs Pract 2000 Dec;6(6):338-344. [doi: 10.1046/j.1440-172x.2000.00234.x]
14. Graham ID, Logan J, Harrison MB, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26(1):13-24. [doi: 10.1002/chp.47] [Medline: 16557505]
15. Pomare C, Churruca K, Long JC, Ellis LA, Braithwaite J. Organisational change in hospitals: a qualitative case-study of staff perspectives. BMC Health Serv Res 2019 Nov 14;19(1):840. [doi: 10.1186/s12913-019-4704-y] [Medline: 31727067]
16. Lyon AR, Munson SA, Renn BN, et al. Use of human-centered design to improve implementation of evidence-based psychotherapies in low-resource communities: protocol for studies applying a framework to assess usability. JMIR Res Protoc 2019 Oct 9;8(10):e14990. [doi: 10.2196/14990] [Medline: 31599736]
17. von Elm E, Schreiber G, Haupt CC. Methodische anleitung für scoping reviews (JBI-methodologie) [Article in German]. Z Evid Fortbild Qual Gesundhwes 2019 Jun;143:1-7. [doi: 10.1016/j.zefq.2019.05.004] [Medline: 31296451]
18. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32. [doi: 10.1080/1364557032000119616]
19. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
20. Kuckartz U, Rädiker S. Qualitative Inhaltsanalyse. Methoden, Praxis, Umsetzung mit Software und Künstlicher Intelligenz: Juventa Verlag; 2024.
21. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
22. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097. [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
23. Atsalos C, O'Brien L, Jackson D. Against the odds: experiences of nurse leaders in Clinical Development Units (Nursing) in Australia. J Adv Nurs 2007 Jun;58(6):576-584. [doi: 10.1111/j.1365-2648.2007.04249.x] [Medline: 17442031]
24. Avallone I, Gibbon B. Nurses' perceptions of their work environment in a Nursing Development Unit. J Adv Nurs 1998 Jun;27(6):1193-1201. [doi: 10.1046/j.1365-2648.1998.00653.x] [Medline: 9663871]
25. Bell M, Procter S. Developing nurse practitioners to develop practice: the experiences of nurses working on a nursing development unit. J Nurs Manag 1998 Mar;6(2):61-69. [doi: 10.1046/j.1365-2834.1998.00051.x] [Medline: 9582779]
26. Bowles A, Bowles NB. A comparative study of transformational leadership in nursing development units and conventional clinical settings. J Nurs Manag 2000 Mar;8(2):69-76. [doi: 10.1046/j.1365-2834.2000.00151.x] [Medline: 11051962]
27. Cannard G. The effect of aromatherapy in promoting relaxation and stress reduction in a general hospital. Complement Ther Nurs Midwifery 1996 Apr;2(2):38-40. [doi: 10.1016/s1353-6117(96)80062-x] [Medline: 9439271]
28. Christensen M, Craft J. The nursing professorial unit: translating acute and critical care nursing research. Int Pract Dev J 2017;7(2):1-10. [doi: 10.19043/ipdj.72.009]
29. Christian SL, Norman IJ. Clinical leadership in nursing development units. J Adv Nurs 1998 Jan;27(1):108-116. [doi: 10.1046/j.1365-2648.1998.00500.x] [Medline: 9515615]
30. Draper J. Nursing development units: an opportunity for evaluation. J Adv Nurs 1996 Feb;23(2):267-271. [doi: 10.1111/j.1365-2648.1996.tb02666.x] [Medline: 8708238]

XSL•FO
RenderX

31. Duffield C. A Master Class for nursing unit managers: an Australian example. J Nurs Manag 2005 Jan;13(1):68-73. [doi: 10.1111/j.1365-2834.2004.00489.x] [Medline: 15613096]
32. Gerrish K, Ferguson A. Nursing development units: factors influencing their progress. Br J Nurs 2000;9(10):626-630. [doi: 10.12968/bjon.2000.9.10.6273] [Medline: 11235272]
33. Graham I. A presentation of a conceptual framework and its use in the definition of nursing development within a number of nursing development units. J Adv Nurs 1996 Feb;23(2):260-266. [doi: 10.1111/j.1365-2648.1996.tb02665.x] [Medline: 8708237]
34. Graham IW. Reflective practice and its role in mental health nurses' practice development: a year-long study. J Psychiatr Ment Health Nurs 2000 Apr;7(2):109-117. [doi: 10.1046/j.1365-2850.2000.00269.x] [Medline: 11146906]
35. Graham I. Leading the development of nursing within a Nursing Development Unit: the perspectives of leadership by the team leader and a professor of nursing. Int J Nurs Pract 2003 Aug;9(4):213-222. [doi: 10.1046/j.1440-172x.2003.00427.x] [Medline: 12887373]
36. Greenwood BJ. On donkeys, ingratiates, thinkers and...clinical development units. Collegian 2000 Apr;7(2):34-39. [doi: 10.1016/s1322-7696(08)60363-7] [Medline: 11855453]
37. Greenwood J. Clinical Development Units (Nursing): the Western Sydney approach. J Adv Nurs 1999 Mar;29(3):674-679. [doi: 10.1046/j.1365-2648.1999.00936.x] [Medline: 10210465]
38. Greenwood J, Kearns E. Establishing a transcultural nursing development unit: the South-Western Sydney experience. Collegian 1996 Jan;3(1):27-31. [doi: 10.1016/s1322-7696(08)60142-0] [Medline: 9265481]
39. Greenwood J, Parsons M. The evaluation of a clinical development unit leadership preparation program by focus group interviews - part 2: negative aspects. Nurse Educ Today 2002 Oct;22(7):534-541. [doi: 10.1016/s0260-6917(02)90761-6] [Medline: 12384034]
40. Happell B, Martin T. Changing the mental health nursing culture: the nursing clinical development unit approach. Int J Ment Health Nurs 2002 Mar;11(1):54-60. [doi: 10.1046/j.1440-0979.2002.00226.x] [Medline: 12400108]
41. Happell B, Martin T. Exploring the impact of the implementation of a nursing clinical development unit program: what outcomes are evident? Int J Ment Health Nurs 2004 Sep;13(3):177-184. [doi: 10.1111/j.1440-0979.2004.0330.x] [Medline: 15361173]
42. Happell B, Martin T. Changing the culture of mental health nursing: the contribution of nursing clinical development units. Issues Ment Health Nurs 2005 Nov;26(9):921-933. [doi: 10.1080/01612840500248213] [Medline: 16203646]
43. Johns C. The Burford Nursing Development Unit holistic model of nursing practice. J Adv Nurs 1991 Sep;16(9):1090-1098. [doi: 10.1111/j.1365-2648.1991.tb03370.x] [Medline: 1939922]
44. Keatinge D, Scarfe C. Creating a nursing development unit in a dementia care context. Int J Nurs Pract 1998 Jun;4(2):120-125. [doi: 10.1046/j.1440-172x.1998.00003.x] [Medline: 9748941]
45. Keatinge D, Scarfe C, Bellchambers H, et al. The manifestation and nursing management of agitation in institutionalised residents with dementia. Int J Nurs Pract 2000 Feb;6(1):16-25. [doi: 10.1046/j.1440-172x.2000.00177.x] [Medline: 10839037]
46. Malby R. Nursing development units in the United Kingdom. Adv Pract Nurs Q 1996;1(4):20-27. [Medline: 9447040]
47. Manley K. A conceptual framework for advanced practice: an action research project operationalizing an advanced practitioner/consultant nurse role. J Clin Nurs 1997 May;6(3):179-190. [Medline: 9188335]
48. Pearson A. An evaluation of the King's Fund Centre Nursing Development Unit network 1989-91. J Clin Nurs 1997 Jan;6(1):25-33. [doi: 10.1111/j.1365-2702.1997.tb00280.x] [Medline: 9052107]
49. Redfern S, Stevens W. Nursing development units: their structure and orientation. J Clin Nurs 1998 May;7(3):218-226. [doi: 10.1046/j.1365-2702.1998.00137.x] [Medline: 9661383]
50. Redfern S, Murrells T. Research, audit and networking activity in nursing development units. NT Research 1998 Jul;3(4):275-288. [doi: 10.1177/174498719800300407]
51. Redfern S, Normand C, Christian S, et al. An evaluation of nursing development units. NT Res 1997 Jul;2(4):292-303. [doi: 10.1177/174498719700200410]
52. Ryan A. Improving discharge planning. Nurs Times 1994;90(20):33-34. [Medline: 8008577]
53. Schiereck S. Social interaction between nurses and female patients in the course of organization of a Nursing Development Unit [Article in German]. Pflege 2000 Aug;13(4):234-241. [doi: 10.1024/1012-5302.13.4.234] [Medline: 11022572]
54. Scholes J. Role transition and emotional labour: understanding the impact of a nursing development unit on staff and on the therapeutic milieu. NT Res 1996 Jul;1(4):286-293. [doi: 10.1177/174498719600100407]
55. Wright SG. Developing nursing: the contribution to quality. 1988. Int J Health Care Qual Assur 2007;20(1):53-60. [doi: 10.1108/09526860710721222] [Medline: 18240508]
56. Bektaş G, Ünal B, Nal M. The effect of transformational leadership on nurses' innovative behaviors: the mediating effect of psychological empowerment. BMC Nurs 2025 Jul 1;24(1):684. [doi: 10.1186/s12912-025-03467-1] [Medline: 40597113]
57. Elbus LM, Mostafa MG, Mahmoud FZ, Shaban M, Mahmoud SA. Nurse managers' managerial innovation and it's relation to proactivity behavior and locus of control among intensive care nurses. BMC Nurs 2024 Jul 16;23(1):485. [doi: 10.1186/s12912-024-02084-8] [Medline: 39014395]

XSL•FO
RenderX

58. Dopp AR, Parisi KE, Munson SA, Lyon AR. Integrating implementation and user-centred design strategies to enhance the impact of health services: protocol from a concept mapping study. Health Res Policy Syst 2019 Jan 8;17(1):1. [doi: 10.1186/s12961-018-0403-0] [Medline: 30621706]
59. Øygarden O, Olsen E, Mikkelsen A. Changing to improve? Organizational change and change-oriented leadership in hospitals. J Health Organ Manag 2020 Aug 25;ahead-of-print(ahead-of-print):687-706. [doi: 10.1108/JHOM-09-2019-0280] [Medline: 32830931]
60. van den Hoed MW, Daniëls R, Beaulen A, Hamers JP, van Exel J, Backhaus R. Perspectives on managing innovation readiness in long-term care: a Q-methodology study. BMC Geriatr 2024 Dec 19;24(1):1017. [doi: 10.1186/s12877-024-05572-3] [Medline: 39702004]
61. De Leeuw JA, Woltjer H, Kool RB. Identification of factors influencing the adoption of health information technology by nurses who are digitally lagging: in-depth interview study. J Med Internet Res 2020 Aug 14;22(8):e15630. [doi: 10.2196/15630] [Medline: 32663142]
62. Ramadan OM, Elsharkawy NB, Hafiz AH, et al. Neonatal nurses' e-health literacy and technology-mediated clinical practice: a cross-sectional analysis of digital health competencies and practice patterns. BMC Nurs 2025 Sep 26;24(1):1199. [doi: 10.1186/s12912-025-03839-7] [Medline: 41013614]
63. Kyrarini M, Lygerakis F, Rajavenkatanarayanan A, et al. A survey of robots in healthcare. Technologies 2021;9(1):8. [doi: 10.3390/technologies9010008]
64. Silvera-Tawil D. Robotics in healthcare: a survey. SN Comput Sci 2024;5:189. [doi: 10.1007/s42979-023-02551-0]
65. Trainum K, Tunis R, Xie B, Hauser E. Robots in assisted living facilities: scoping review. JMIR Aging 2023 Mar 6;6:e42652. [doi: 10.2196/42652] [Medline: 36877560]
66. Ignatowicz A, Tarrant C, Mannion R, El-Sawy D, Conroy S, Lasserson D. Organizational resilience in healthcare: a review and descriptive narrative synthesis of approaches to resilience measurement and assessment in empirical studies. BMC Health Serv Res 2023 Apr 19;23(1):376. [doi: 10.1186/s12913-023-09242-9] [Medline: 37076882]

Abbreviations

NDU: Nursing Development Unit
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR : Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by L Donelle; submitted 05.Dec.2025; peer-reviewed by N Nestler, S Walzer; revised version received 11.Mar.2026; accepted 22.Apr.2026; published 04.Jun.2026.

Please cite as:
Gangnus A, Müller P, Gayen K, Kirchner-Heklau U, Hofstetter S, Jahn P
Facilitating Digital Transformation in Nursing Through Nursing Development Units: Scoping Review
JMIR Nursing 2026;9:e89051
URL: https://nursing.jmir.org/2026/1/e89051
doi:10.2196/89051

© Annabell Gangnus, Pascal Müller, Katrin Gayen, Uta Kirchner-Heklau, Sebastian Hofstetter, Patrick Jahn. Originally published in JMIR Nursing (https://nursing.jmir.org), 4.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Simulation-Based Training for Nursing Students to Improve Patient Safety: Systematic Review

Azucena González-Sanz[1], MSN; José Carlos López-García[2], BSN; Elena Sutil-Rodríguez[3,4], PhD; Raúl Juárez-Vela[5], PhD; José Ángel Santos-Sánchez[6], PhD; Noelia Navas-Echazarreta[5], PhD; Antonio Martínez-Sabater[5,7,8], PhD; Consuelo Sancho-Sanchez[6], PhD

[1]Primary Care Center Toro, Gerencia de Asistencia Sanitaria de Zamora (GAZSA), Zamora, Castille and León, Spain

[2]Primary Care Center Virgen Concha, Gerencia de Asistencia Sanitaria de Zamora (GAZSA), Zamora, Castille and León, Spain

[3]Primary Care Center Benavente Sur, Gerencia de Asistencia Sanitaria de Zamora, SACYL, Zamora, Castille and León, Spain

[4]CECAVIS Research Group, Instituto de Investigación Biosanitaria de León (IBIOLEON LEÓN), Leon, Castille and León, Spain

[5]Research Group in Care (GRUPAC), Pre-departmental Biomedical Sciences and Health Specialties Unit, Faculty of Health Sciences, Universidad of La Rioja, C/ Duquesa de la Victoria, 88, Logroño, La Rioja, Spain

[6]Faculty of Medicine, Universidad de Salamanca, C. Alfonso X el Sabio, s/n, Salamanca, Castille and León, Spain

[7]Nursing Care and Education Research Group (GRIECE),Nursing Department, Universitat de València, Valencia, Valencia, Spain

[8]Care Research Group (INCLIVA), Hospital Clínico Universitario de Valencia, Valencia, Valencia, Spain

Corresponding Author:
Noelia Navas-Echazarreta, PhD
Research Group in Care (GRUPAC), Pre-departmental Biomedical Sciences and Health Specialties Unit, Faculty of Health Sciences, Universidad of La Rioja, C/ Duquesa de la Victoria, 88, Logroño, La Rioja, Spain

Abstract

Background: Patient safety is a fundamental pillar of health care quality. Simulation-based training provides a controlled environment for nursing students to develop safety competencies and error-recognition skills before clinical practice.

Objective: This systematic review aimed to describe and characterize the simulation-based education features and modalities used to address patient safety outcomes in undergraduate nursing students, identifying the strategies that contribute to improvements in safety-related competencies.

Methods: A systematic review was conducted following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines across PubMed, Web of Science, Scopus, CINAHL, Cochrane, and Lilacs (2019 - 2024). Inclusion criteria focused on original studies involving undergraduate nursing students and simulation interventions measuring patient safety outcomes. Studies in languages other than English, Spanish, or Portuguese were excluded. Two reviewers independently performed study selection and data extraction. Methodological quality was assessed using Joanna Briggs Institute tools, applying a 60% quality threshold for inclusion. Results were synthesized through a narrative approach.

Results: A total of 20 studies from 12 countries were included. The methodological quality was high (n=14) and moderate (n=6). Findings revealed that high-fidelity simulation and virtual reality are the primary strategies used. Simulation proved effective in enhancing both technical skills (medication administration accuracy) and nontechnical skills (communication via SBAR [Situation, Background, Assessment, Recommendation] and ISBAR [Identification, Situation, Background, Assessment, Recommendation] tools, teamwork, and adverse event reporting). Key strategies contributing to safety included repetitive practice and interprofessional simulation, which significantly improved error detection and clinical judgment.

Conclusions: Simulation is an essential pedagogical strategy for preparing nursing students to deliver safe care. Practical implications include the need to integrate structured simulation into nursing curricula to bridge the theory-practice gap. Future research should prioritize longitudinal designs to assess the retention of these safety skills in clinical settings and develop standardized metrics for measuring patient safety outcomes.

Trial Registration: PROSPERO CRD420250653469; https://www.crd.york.ac.uk/PROSPERO/view/CRD420250653469

(JMIR Nursing 2026;9:e87898) doi:10.2196/87898

KEYWORDS

nursing students; simulation training; patient safety; nursing; systematic review; PRISMA; Preferred Reporting Items for Systematic Reviews and Meta-Analyses

XSL•FO
RenderX

Introduction

Overview

Patient safety is a fundamental pillar of quality health care, aimed at minimizing risks and preventing harm during medical care [1]. According to the World Health Organization (WHO), this concept is defined as "the absence of avoidable harm to a patient and the reduction to an acceptable minimum of the risk of unnecessary harm associated with healthcare" [2]. Despite its importance, safety failures remain a critical global issue; more than 1 in 10 patients suffer harm during treatment, leading to more than 3 million deaths annually. Consequently, unsafe care must be recognized as both a major public health crisis and a significant economic burden [3].

The study of these failures gained international prominence following the landmark report "To Err is Human," published in 2000 by the National Academy of Sciences. This study revealed that the cost of medical errors extends beyond economics, deeply affecting patient satisfaction, professional morale, and overall trust in the health system [4]. To address these challenges, a comprehensive monitoring approach is required. This involves the proactive identification of hazards, continuous process improvement, and the promotion of a safety culture that actively engages both health care professionals and patients [1].

Patient safety culture is a complex phenomenon in which values, attitudes, competencies, and behaviors influence how health care professionals perceive and manage safety risks [2]. A comprehensive framework for patient safety culture consists of leadership, teamwork, evidence-based practice, communication, learning, just culture, and patient-centered care [5]. This concept is integrated with WHO resolution WHA72.6 [6], which prioritizes concrete measures to prevent avoidable harm in health care, such as strategies for safe medication, error-free surgery, infection control, sepsis management, reliable diagnostics, and adequate hygiene in facilities. This multidimensional framework highlights the systemic nature of safety culture and the central role of leadership and communication in preventing adverse events.

The occurrence of safety incidents is rarely the result of a single individual's actions but rather the culmination of various systemic factors [4]. These include rapid technological advances, increasingly complex care processes, inadequate policies, and an aging population with multiple chronic comorbidities [2]. Because incidents often stem from multiple overlapping causes, focusing on individual blame fails to address the underlying systemic issues, making it likely that the same errors will recur [2,4]. To mitigate these risks, the Global Patient Safety Action Plan 2021‐2030 outlines strategic activities designed to identify, evaluate, and manage risks throughout the health care continuum, with the ultimate goal of preventing harm or minimizing its impact [7].

Training and education in safe practices are essential to maintaining and improving standards of care and ensuring that health systems can adapt and respond to emerging challenges in health care [1]. In this context, simulation, defined as any technique that extends or replaces real-world experiences to promote reflective learning, has established itself as a key strategy for improving research on patient safety incidents [8].

Clinical simulation has proven effective as a teaching strategy for developing both technical and nontechnical skills, including communication, teamwork, leadership, and critical thinking [9,10]. This approach allows for the creation of structured, meaningful, and reflective learning environments that facilitate the effective and safe resolution of complex situations, in line with the competencies to be acquired [8-10].

Clinical scenarios form the basis of clinical simulation, as they allow learning experiences to be structured with different levels of complexity in line with training objectives [11]. A key element in their design is fidelity, understood as the degree to which the simulation reproduces the conditions of actual clinical practice [12], which can be classified as low, medium, or high depending on the resources used and the educational goals. Low-fidelity simulation typically focuses on the acquisition of basic technical skills using task trainers, simple manikins, or case studies, while medium and high levels incorporate more complex scenarios and greater realism. Similarly, the simulation modality refers to the format used to develop the training experience. According to the Healthcare Simulation Dictionary, these modalities can include role-playing, virtual or online simulations, task trainers, high-tech mannequins, or immersive simulations with standardized patients (SPs) or trained actors [12]. Thus, the combination of modality and fidelity level allows simulation to be adapted to different educational contexts and clinical competencies [11]. Immersive simulation, particularly when based on realistic clinical scenarios, offers strong opportunities to develop competencies and promote changes in thinking and practice. Its effectiveness is enhanced when followed by structured debriefing, enabling reflective analysis and identification of improvement strategies [9].

Nursing professionals, from their formative stage, need to be actively trained to provide safe care, and effective communication is essential to reduce health errors [13,14]. A culture of open communication facilitates enhanced team interactions and equips students with efficacious strategies to deploy in their practice, thereby fostering critical thinking and emergency management skills. The formation of effective interprofessional teams through simulation-based learning has been linked to improved patient outcomes, a reduction in medical errors, and the delivery of high-quality care [15]. The students have acquired knowledge and have been instructed in the guidelines for conducting practices more safely. However, they encounter obstacles when attempting to communicate and report errors, as they perceive themselves to have a lower status and are sometimes fearful of the potential consequences [16].

A coalition of international scientific societies with expertise in simulation has issued a call for the integration of simulation-based learning into the curricula of undergraduate and postgraduate programs. This initiative is designed to foster a culture of safety [17]. It is also imperative that the simulation be designed in such a way that it does not compromise the psychological and physical integrity of the participants, while also ensuring the safety of the training process [18].

XSL•FO
RenderX

Previous reviews on educational interventions in patient safety highlight considerable methodological heterogeneity and limited evidence regarding the specific impact of simulation [19]. One review found that although patient safety competence was frequently assessed, only 2 studies used simulation, and none examined behavioral changes, limiting conclusions about real clinical impact [18]. A published protocol targeting nursing students also considers simulation among various teaching methods, but anticipates substantial variability that may hinder conclusions about its differential effectiveness [20].

Another systematic review reported improvements in knowledge, clinical skills, and confidence through simulation-based learning; however, it included varying fidelity levels and did not specifically address patient safety outcomes, limiting applicability to this domain [21]. Similarly, quantitative syntheses confirm the benefits of simulation but do not analyze its transfer to patient safety in real health care contexts [22]. A comparison between SPs and role-playing showed improvements in communication skills, yet without linking findings to objective patient safety indicators [23]. Finally, an exploratory review associated simulation with greater competence and confidence but did not differentiate between levels of evidence or modalities, limiting the strength of its recommendations [24].

Overall, despite suggested benefits, the literature shows methodological heterogeneity, limited analysis of behavioral outcomes, and insufficient connection to specific patient safety indicators, raising the question of the actual effect of simulation-based learning on patient safety outcomes in nursing students.

Objective

Since 2020, with the closure of in-person training activities and the impossibility of clinical practices, justified by immediate public health needs: preventing the transmission of infections, facilitating social distancing, and responding to government orders, the need for education and training that adapts to the new demands of the health system has become evident [25,26].

The use of simulation methodology in the education of nursing students can be a strategy that improves the application of the theoretical concepts learned and clinical practice, to provide safe and quality care. The objectives of this systematic review were (1) to determine the effect of simulation-based education on patient safety outcomes in nursing students, (2) to identify the aspects related to patient safety that have been the subject of the use of simulation, and (3) to describe and characterize simulation-based education features and modalities used to address patient safety outcomes in undergraduate nursing students.

Methods

Design

A systematic review was carried out during the months of February and March 2024, following the checklist of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [27]. The review followed the PRISMA 2020 guidelines and was designed and presented as a systematic review rather than an exploratory review, to accurately determine the effect of simulation-based education on patient safety outcomes, establishing strict predefined eligibility criteria, quality assessment, and structured synthesis of evidence (Checklist 1).

Search Strategy

A systematic search was carried out from 2019 to 2024 in the databases: PubMed, Web of Science, Scopus, CINAHL, Cochrane, and Lilacs. These databases were selected to ensure broad coverage of biomedical, nursing, educational, and multidisciplinary research, as well as to minimize publication and indexing bias. To guide the search, we started from the PIO (Population, Intervention, Outcomes) question (Table 1).

The terms were adapted to the controlled language of the MeSH (Medical Subject Headings) and Descriptors in Health Sciences, in addition to CINAHL subject headings (Table 2).

Table . Keywords PIO[a].

PIO question	Keywords	MeSH[b]	DeCS[c]
Population	Nursing students	Nursing students	Estudiantes de enfermería
Intervention	Training through simulation techniques	Simulation training	Simulación
Outcomes	Improving patient safety	Patient safety	Seguridad del paciente

[a]PIO: Population, Intervention, Outcomes.

[b]MeSH: Medical Subject Headings.

[c]DeCS: Descriptors in Health Sciences in Spanish.

XSL•FO
RenderX

Table . Search strategy.

Base	Search strategy	Number of articles
PubMed	((“education, nursing”[MeSH Terms] OR (“education”[All Fields] AND “nursing”[All Fields]) OR “nursing education”[All Fields] OR “education nursing”[All Fields] OR (“nursing education research”[MeSH Terms] OR (“nursing”[All Fields] AND “education”[All Fields] AND “research”[All Fields]) OR “nursing education research”[All Fields])) AND "2019/02/27 00:00”:“3000/01/01 05:00”[Date - Publication] AND ((“students, nursing”[MeSH Terms] OR (“students”[All Fields] AND “nursing”[All Fields]) OR “nursing students”[All Fields] OR (“nursing”[All Fields] AND “students”[All Fields])) AND (“simulation training”[MeSH Terms] OR (“simulation”[All Fields] AND “training”[All Fields]) OR “simulation training”[All Fields]) AND (“patient safety”[MeSH Terms] OR (“patient”[All Fields] AND “safety”[All Fields]) OR “patient safety”[All Fields]) AND "2019/02/27 00:00”:“3000/01/01 05:00”[Date - Publication] AND "2019/02/27 00:00”:“3000/01/01 05:00”[Date - Publication] AND "2019/02/27 00:00”:“3000/01/01 05:00”[Date - Publication])) AND (y_5[Filter])	92
Scopus	(TITLE-ABS-KEY (nursing AND students) AND TITLE-ABS-KEY (simulation AND training) AND TITLE-ABS-KEY (nursing AND education) AND TITLE-ABS-KEY (patient AND safety)) AND PUBYEAR >2018	109
WoS[a]	(((TS=(nursing students)) AND TS=(simulation training)) AND TS=(nursing education)) AND TS=(patient safety) Last 5 years	136
LILACS	(estudiantes de enfermería) AND (simulación) AND (educación en enfermería) AND (seguridad del paciente) AND (year_cluster:[2019 TO 2024])	14
CINAHL	nursing students AND nursing education AND (simulation training or simulation education or simulation learning) AND (patient safety or patient outcomes or quality of care) Limiters - Publication Date: 20190101 - 20241231	133
Cochrane	(*simulation training) AND (nursing students):ti,ab,kw AND (patient safety):ti,ab,kw (Word variations have been searched) con año de publicación de 2019 hasta 2024, fecha de publicación en la Biblioteca Cochrane Entre Jan 2019 y Feb 2024, en Ensayos	32

[a]WoS: Web of Science.

Inclusion and Exclusion Criteria

The inclusion criteria were articles written in English, Spanish, and Portuguese, which were primary studies on the use of simulation in the undergraduate stages of nursing to maintain patient safety, such as communication, teamwork, medication safety, clinical performance linked to safety, without differentiating the type of simulator used. Although initially it was decided to include only nursing students, it was later decided to also include studies involving medical students in the case of communication or teamwork.

Exclusion criteria were considered qualitative studies, reviews, editorials, personal experiences, and those quantitative articles that were not peer reviewed, studies that dealt exclusively with student satisfaction and perception with the simulation methodology, studies in which among the participants there were already graduated professionals, and those in which at least one of the objectives did not deal with safety, studies on the use of simulators that were specific for learning specialized care, studies referring to the design of simulation scenarios or the development of modifications of educational strategies related to simulation.

XSL•FO
RenderX

Article Selection, Methodological Quality Assessment, and Risk of Bias

The articles obtained were exported to the Zotero bibliographic manager. After removing duplicates, articles were selected by title and abstract during online sessions by 2 researchers. When discrepancies arose, another revisor was contacted for a final decision. Full texts of the remaining articles were obtained to assess eligibility. Two reviewers thoroughly assessed the full texts to determine study inclusion, resolving discrepancies through discussion and, when necessary, with the involvement of a third reviewer. Two reviewers determined that articles met eligibility criteria. References from secondary articles, such as reviews and scoping reviews, were also reviewed to assess their potential inclusion of those articles that met the objectives of this study. Those selected underwent a quality evaluation by 2 authors independently using the Joanna Briggs Institute (JBI) Critical Appraisal Tools [28] according to study design (the checklist for randomized controlled trials [RCTs] [29], the checklist for quasi-experimental studies [30], and the checklist for analytical cross-sectional studies [31]).

The methodological quality of all included studies was assessed independently by 2 reviewers using the JBI Critical Appraisal Tools according to the study design: the checklist for RCTs, quasi-experimental studies, and analytical cross-sectional studies. Discrepancies were resolved by consensus with a third reviewer. Studies meeting ≥60% of the applicable JBI criteria were included in the final synthesis. Considering that the JBI critical appraisal tools do not establish universal cutoff points for classifying methodological quality, leaving this decision to the discretion of review teams, a 60% threshold was applied as a pragmatic criterion to balance the inclusion of methodologically acceptable studies while avoiding overly restrictive exclusions, thereby ensuring the availability of sufficient evidence for analysis [28].

Data Extraction and Synthesis

The authors independently extracted data from the studies into a Microsoft Excel spreadsheet. The sheet recorded data on the author, year of publication, country, age of participants, type of study, objective, scales and instruments used, results and conclusions, limitations, safety aspects covered, type of simulation, types of interventions performed, and simulated cases (Table S7 in Multimedia Appendix 1 [32-51] and Table S8 in Multimedia Appendix 2 [32-51]). Given the diversity among the included studies, clinical and methodological heterogeneity (diverse domains, modalities, and durations), noncomparable instruments, and incomplete statistical reporting (missing effect sizes and variances), a narrative synthesis was conducted, guided by the principles of thematic content analysis.

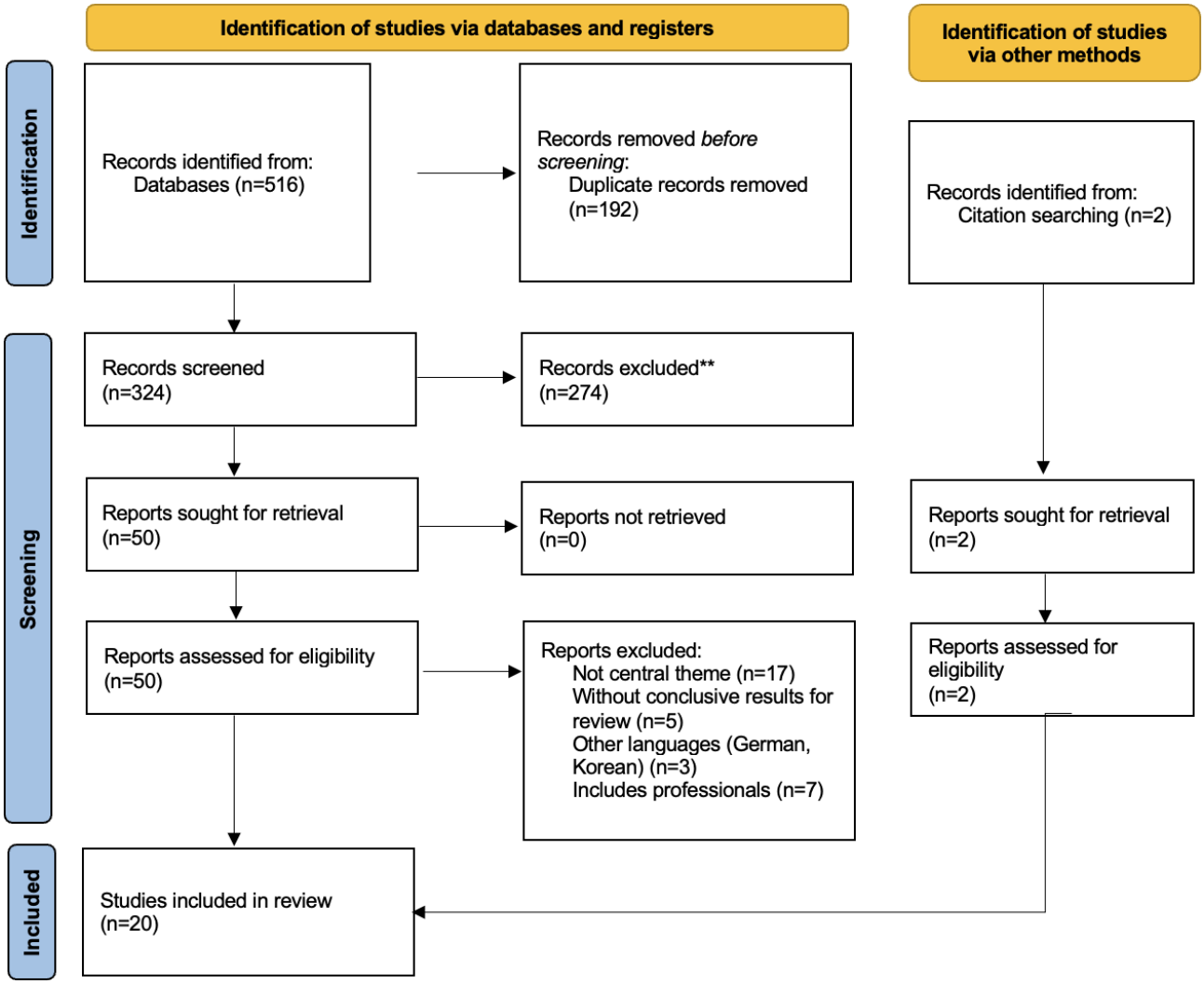
Results

Study Selection

The selection process is shown in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram of study selection.



Characteristics of Included Studies

A total of 20 studies [32-51] met the inclusion criteria for this review. The primary characteristics of these studies are detailed in Table 3. The included articles were published between 2019 and 2024. The country with the highest number of studies was Korea with 6 studies [32,34,36,37,41,48], followed by the United States [32,40,45] and China [39,43,49], with 3 studies each. Spain [42,50] and Taiwan [46,51] contributed 2 studies each. Germany [35], Ireland [47], and Indonesia [38] contributed 1 study each. One study was led by an international team (Canada, England, Scotland, and Australia) [44]. For the purpose of counting, each of the 20 studies [32-51] was assigned to the country of the coordinating institution, and the multicountry study was considered as a single study with international participation.

XSL•FO
RenderX

Table . Characteristics of included studies.

Author (year), country	Participants	Type of study and assignment	Objectives	Data collection scales and instruments	Results and conclusions	Limitations
Breen et al [47] (2019), Ireland	Total 90: 45 nursing students 3rd course, 45 last-year medical students. Year: 2016.	ECA 3 groups: e-learning (E)[a], E+standard simulation (E+S)[b], competency-based progression simulation and e-learning (E+PBP)[c]. Randomized	To determine the effectiveness of a competency-based training (PBP)[d] for clinical communication.	National Early Warning Score (NEWS)[e] based on the ISBAR[f] tool	Communication competence was achieved by 2/29 (7%) in E, 3/23 (13%) in E+S, and 15/25 (60%) in competency-based progression. Competency-based training was significantly more effective (P<.001)	Single center; undergraduate sample; short training duration
Son and Kim [36] (2019), Korea	98 Nursing students. CG[g]: 41, IG[h]: 57 Course 3rd Year: 2017 Age: 22 - 23 Gender: Women	Quasi-experimental. Pretest-posttest design of nonequivalent control group. Convenience sampling	Effectiveness of communication education based on SEGUE[i] between students and patients.	Communication competence and effectiveness measured by a self-report questionnaire and by teacher and standardized patient ratings of the students	Intervention students showed significantly greater improvements in communication competence and efficacy than controls (In terms of performance, as assessed by lecturers and students, the P value is <.001, the standardized patient score is 0.042, and the communication effectiveness score is 0.004).	Single institution; small control group
Jeong and Kim [37] (2020), Korea	54 nursing students. IG: 26, CG: 28 Grade: 3rd Year: 2018 Age: 20 - 40 average 23 Gender: Women, only 1 man	Clinical trial (pre-post). Simple blind	Develop a fall simulation program using the SBAR[j] communication technique	Knowledge of falls and attitude toward falls scale, post-fall evaluation protocol (AHRQ[k]). GICC-15[l] and Dunsford adapted structured communication	SBAR-based simulation significantly improved structured fall-related communication compared with controls.	No long-term follow-up; single-blind design
Liaw et al [49] (2020), China	120 medical and nursing students. IG: 60, CG: 60 Grade: 3rd-4th year. Year: 2018. Mean age 22.17 (SD 2.07) years. Gender: Women (81, 67.5%)	RCT[m] (pretest-posttest) Randomized	To evaluate a team training program using VR[n] versus conventional live simulations on the performance of communication skills and teamwork	ATHCT[o] and ISVS[p]. Baseline, posttest, and 2 months after	Both groups improved teamwork attitudes; VR training was noninferior with higher follow-up ISVS.	Self-reported outcomes; single-center design
Sanko and McKay [33] (2020), United States	231 nursing students from 2016 (CG: 68), 2017 (IG: 85), and 2018 (CG: 78) patient safety courses.	Clinical trial (2 cohorts years different intervention and control groups). Convenience sampling	To assess whether exposure to simulation scenarios to enhance systematic thinking influenced adverse event reporting and the type	Simulated AERS[q]. System Thinking Scale (STS)[r]	Systems thinking scores increased significantly after simulation, and adverse event reporting improved (P<.001).	Single-site sample; simulated reporting system
Lee and Kim [34] (2020), Korea	194 nursing students. 47 teams (4 and 3 components) Course: higher level. Year: 2015 - 2016	Prospective observational convenience sampling	To examine the relationships between nursing students' team task performance and SBAR-R[s] communication.	Measurement of team task performance and communication using developed checklists based on SBAR-R	Higher SBAR-R scores were significantly associated with better team task performance (P=.004).	Single university; limited scope

XSL•FO
RenderX

Author (year), country	Participants	Type of study and assignment	Objectives	Data collection scales and instruments	Results and conclusions	Limitations
Wai et al [48] (2021), Korea	46 students: 19 medical, 27 nursing. Interprofessional teams. Grade: Final year medicine (5th), 3rd-4th year nursing. Gender: 63% female students.	Mixed methods convenience sampling	To compare the effectiveness of combined classroom plus clinical simulation versus clinical simulation alone on attitudes, perceptions, and teamwork performance.	HFAS[t] Survey. Teamwork performance using TBL-SAI[u] focus group interview	Both groups improved teamwork attitudes (P=.04) with no added benefit from classroom teaching	Small sample; limited power
Musharyanti et al [38] (2021), Indonesia	95 nursing students IG:55. CG: 40.	Quasi-experimental. Nonequivalent randomized control group.	To compare drug safety knowledge and skills after safety training with the 4C/ID[v] teaching method.	MCQs[w] and 2 checklists developed ad hoc on patient safety and medication administration	Intervention students achieved significantly higher safety knowledge and skills than controls (P<.001).	No baseline testing
Du et al [39] (2021), China	47 nursing students.CG:21, IG:26 Grade: 2nd Year 2019. Age 17 - 27 Gender: 44 women -3 men	Controlled trial. Simple blind.	To assess the risk of pressure ulcer development in 3 different scenarios	OSCE[x] adapted to different scenarios Pressure Ulcer Knowledge Assessment Tool (PUKAT 2.0)[y]	OSCE-based simulation significantly improved assessment performance (P<.001).	Small sample. Single center
Lee and Lim [32] (2021), Korea	30 nursing students. Grade: Final year. Year: 2018. Average age 22.17 Gender: women	Quasi-experimental (pre-post). Convenience sampling	Develop, implement, and verify the effectiveness of a simulation-based handover education program.	Communication tool adapted from SBAR. Knowledge questionnaire developed by authors. Self-efficacy (adapted questionnaire).PASS-BAR[z]	Simulation significantly improved handover knowledge, self-efficacy, and performance (P≤.001).	No control group; all-female sample
Craig et al [40] (2021), United States	83 nursing students. CG:35. IG:45 Grade: 3rd.	Quasi-experimental. Convenience sampling	To examine the effects of an educational strategy using a simulation program on medication management.	MSKA[aa] and MSCEC[ab]	Simulation significantly improved medication administration skills (P<.001).	Single-site study
Raurell-Torreda et al [50] (2021), Spain	93 nursing students. IG: 48, CG:45 Course: 3rd, 1 medical student in 5th year. Age: CG: 22.3±5.2- IG: 23.3±6.8 Gender: 78.5% women	Clinical Trial Randomized	Evaluate the impact of SBAR training on interprofessional teamwork skills (role-related and communication) and nontechnical skills.	KidSIM-TPS[ac] and CSET[ad] (nontechnical skills)	Intervention students showed significant improvements in teamwork behaviors (P≤.004).	Partial implementation
Park and Kim [41] (2021), Korea	91 nursing students. IG: 47, CG: 44. Final year of nursing Academic year 2018/19. Age: IG 22.59±1.23, CG 22.86±1.39 Gender:+77% women	Clinical Trial Randomized	Analyze the impact of simulated patient deterioration on situational awareness and patient safety competence-attitude.	SAGAT[ae] modified; PSCSE[af] modified.	Simulation significantly improved situational awareness and safety attitudes (P<.001).	Single university
Chen et al [51] 2022, Taiwan	54 students: 18 medical students and 36 nursing students Grade: 4th year nursing, 5th year medicine Year: 2019 Gender: 70% Women	Mixed methods Randomized	Determine the importance of interprofessional training on competence, teamwork attitudes, and safety.	MTP[ag], TBP[ah], TA[ai], and PSA[aj]	Both groups achieved comparable competence gains; qualitative data supported learning benefits	Small sample

XSL•FO
RenderX

Author (year), country	Participants	Type of study and assignment	Objectives	Data collection scales and instruments	Results and conclusions	Limitations
Pol-Castañeda et al [42] (2022), Spain	179 nursing students. Grade: 2nd Academic year 2018/19 Age: 60% between 18 - 25 Gender: 89% women	Mixed methods Convenience sampling	To assess the acquisition of skills in safe medication administration by nursing students	Questionnaire. adapted from the. MASAT[ak]	Simulation improved most medication skills except documentation.	Assessment variability
Goldsworthy et al (2022), Canada, England, Scotland, and Australia [44]	88 nursing students 5 diverse university sites in 4 countries Grade: 3rd-4th Pandemic year 2020?	Quasi-experimental Convenience sampling	Explore the impact of a virtual simulation on to recognize and respond to a rapidly deteriorating patient.	10-item Clinical Self-Efficacy Survey designed 20-item multiple-choice test on evidence-based practice	Virtual simulation significantly improved knowledge scores (P=.001).	Nonrandomized design
Li et al [43] (2023), China	205 nursing students.IG:103, CG: 102. Grade: 2nd Academic year 2020/21. Mean age CG: 19.65 (SD 0.75) years, IG:19.78 (SD 0.77) years. Gender: women 87% - 89%	Quasi-experimental (2-tailed) Convenience sampling	Exploring the effects of an online course (SPOC)[al] combined with simulation-based training in a patient safety education program.	PSCSE	Patient safety competence scores were significantly higher in intervention students (P<.001).	Single-center study
Haerlin et al [45] (2023), United States	193 nursing students: Group Clinical experience: 51, Group Mannequins:44, Group VR: 57. Academic year 2021/22. Age: 19-53 (median 21). Gender: 81.6% female	Quasi-experimental Convenience sampling	Compare differences in learning and practice in patient care as a function of the learner's training experience.	Competency by CCEI[am] and LCJR[an]. Clinical Learning CLECS[ao] 2.0	Manikin-based simulation produced equal or superior outcomes compared with other modalities.	Regional sample
Chou et al [46] (2024), Taiwan	84 nursing students IG: 42 and CG: 42. Grade: 2nd Year: 2022. Mean age 20.3 (SD 0.46) years. Gender: 80% women	Clinical Trial Randomized	To examine the effectiveness of a VR communication simulation in the acquisition of communication skills in the fundamentals of nursing practice.	Kalamazoo Consensus Statement Essential Elements Communication Checklist, "Communication Self-Assessment Scale" modified, Development and Testing of a Perceived Stress Scale for Nursing Students in Clinical Practice and a learning satisfaction	Kalamazoo Consensus Statement Essential Elements Communication Checklist, "Communication Self-Assessment Scale" modified, Development and Testing of a Perceived Stress Scale for Nursing Students in Clinical Practice and a learning satisfaction	Short intervention; self-report bias

XSL•FO
RenderX

Author (year), country	Participants	Type of study and assignment	Objectives	Data collection scales and instruments	Results and conclusions	Limitations
Heier et al [35] (2024), Germany	221 Students, 154 medical and 67 nursing students (IG: 66 medical, 28 nursing/CG: 88 medical, 39 nursing) Course: 3rd year medical and 1st-2nd year nursing students. Year: October 2021-March 2023. Mean age 24 (SD 3.9) years. Gender: 51.13% women	Mixed methods Convenience sampling	Develop joint communication skills training for nursing and medical students in professional error communication.	Adaptation of G-IPAS[ap] "Teamwork,Roles and Responsibilities," "Patient- centeredness" and a self-developed interprofessional error communication scale	Adaptation of G-IPAS "Teamwork,Roles and Responsibilities," "Patient- centeredness" and a self-developed interprofessional error communication scale	Nonrandomized design

[a]E: e-learning.
[b]E+S: e-learning plus standard simulation group.
[c]E+PBP: e-learning plus competency-based progression simulation group.
[d]PBP: progression-based performance/competency-based progression.
[e]NEWS: National Early Warning Score.
[f]ISBAR: Identification, Situation, Background, Assessment, Recommendation.
[g]CG: control group.
[h]IG: intervention group.
[i]SEGUE: Set the stage, Elicit information, Give information, Understand the patient's perspective, End the encounter.
[j]SBAR: Situation, Background, Assessment, Recommendation.
[k]AHRQ: Agency for Healthcare Research and Quality.
[l]GICC-15: General Interpersonal Communication Competency (15 items).
[m]RCT: randomized controlled trial.
[n]VR: virtual reality.
[o]ATHCT: Attitudes Toward Interprofessional Health Care Team.
[p]ISVS: Interprofessional Socialization and Valuing Scale.
[q]AERS: Adverse Event Reporting System.
[r]STS: Systems Thinking Scale.
[s]SBAR-R: Situation, Background, Assessment, Recommendation, Read-back.
[t]HFAS: Human Factors Attitude Survey.
[u]TBL-SAI: Team-Based Learning Student Assessment Instrument.
[v]4C/ID: Four-Component Instructional Design.
[w]MCQ: multiple choice questionnaire.
[x]OSCE: Objective Structured Clinical Examination.
[y]PUKAT: Pressure Ulcer Knowledge Assessment Tool.
[z]PASS-BAR: Patient Safety Screen-Based Assessment Record (handover tool).
[aa]MSKA: Medication Safety Knowledge Assessment.
[ab]MSCEC: Medication Safety Critical Elements Checklist.
[ac]KidSIM-TPS: KidSIM -Program Team Performance Scale.
[ad]CSET: Clinical Simulation Evaluation Tool.
[ae]SAGAT: Situational Awareness Global Assessment Technique.
[af]PSCSE: Patient Safety Competency Self-Evaluation.
[ag]MTP: Medical Task Performance.
[ah]TBP: Team Behavior Performance.
[ai]TA: Teamwork Attitude.
[aj] PSA: Patient Safety Attitude.
[ak]MASAT: Medication Administration Safety Assessment Tool.
[al]SPOC: Small Private Online Course.
[am]CCEI: Creighton Competency Evaluation Instrument.

XSL•FO
RenderX

[an]LCJR: Lasater Clinical Judgment Rubric.

[ao]CLECS: Clinical Learning Environment Comparison Survey.

[ap]G-IPAS: German Interprofessional Attitudes Scale.

The study designs varied, encompassing 8 RCTs [33,37,39,41,46,47,49,50], 11 quasi-gstudies [32,35,36,38,40,42-45,48,51], and 1 analytical cross-sectional study [34]. The participant samples predominantly consisted of undergraduate nursing and/or medical students at various stages of their training. The objectives of the studies commonly focused on evaluating the use of simulation-based training on competencies essential for patient safety, such as interprofessional communication, teamwork, and clinical reasoning. The studies that included only nursing students were 14 [32-34,36-46] and the remaining 6 were studies with a mixture of nursing and medical students [35,47-51], although one of them only included 1 medical student [50].

The total number of participants included was 2494. The study with the least participation grouped 30 students [41] and the one with the highest participation had 231 [42]. Most participants were women (more than 60%); in one of the studies, all participants belonged to a women's nursing school [41].

The studies were carried out from 2015 [34] to 2023 [35]; 5 of them did not report the age of the participants, and 4 did not report the academic year in which the participants were enrolled, their sex, or the year the study was conducted.

Measurement tools varied across studies. Only 4 studies used ad hoc questionnaires [34,36,38,44]. The remaining 16 studies [32,33,35,37,39-43,45-51] used validated instruments, some with contextual adaptations (eg, cultural and language modifications and scenario-specific items).

Characteristics of Interventions

The simulation-based training interventions detailed in the included studies were heterogeneous, as detailed in Table 4. Interventions ranged from single-session workshops (eg, 3.5 h) to multisession programs integrated into the curriculum.

The modalities of simulation used included high-fidelity simulators (HFS; n=9) [32,34,39-41,43,45,47,48], SPs (n=8) [35-39,48,49,51], virtual reality (VR; n=4) [44-46,49], role play (n=2) [38,50], and board simulations (n=1) [33]. Several studies combined HFS with SP or VR, so the categories are not mutually exclusive. The transfer of communication using a structured methodology between the team uses different strategies, such as SBAR (Situation, Background, Assessment, Recommendation) and SBAR-R [32,34,37,47,50] handover tools and KidSIM-TPS [50] teamwork frameworks, while other less commonly used techniques include the Kalamazoo communication checklist [46]. Structured communication with the patient is guided by the SEGUE (Set the stage, Elicit information, Give information, Understand the patient's perspective, End the encounter) framework [36]. These structured tools were detailed in the instruments and notes of Table 3.

XSL•FO
RenderX

Table . Characteristics of interventions.

Author (year), country	Modalities and type of simulation	Security aspects covered	Direction of effect	Type of intervention in the study	Case
Breen et al [47] (2019), Ireland	HFS[a]-HF[b]	Communication between professionals	Favors competency-based simulation	All groups 15 min training on the ISBAR[c] tool. Group (E[d]): HF room. Group (E+S[e]) work in mixed discipline pairs: telephone calls on 4 standardized cases (3.5 h). Group (E+PBP[f]): same training as group E+S, reached competencies	4 standardized clinical cases of acute patient deterioration
Son and Kim [36] (2019), Korea	SP[g]-HF	Communication with the patient	Favors simulation	CG[h]: Prebriefing + pretest (50 min), simulation 60 min, debriefing+posttest (50 min) afterward, and SEGUE[i]-based communication (30 min). IG[j]: SEGUE communication before	Standardized pediatric scenario, with mothers of 5-year-old children admitted to the hospital for acute gastroenteritis with fever
Jeong and Kim [37] (2020), Korea	SP-HF	Falls and communication	Favors simulation	IG: 3 educational sessions on falls communication and SBAR[k] method (theory, practice, and discussion). CG: training according to guidelines on patient care and transfer of information	Standardized patient falling out of bed
Liaw et al [49] (2020), China	VR[l] (simulated environment with avatar use) and live simulation (SP)-HF	Interprofessional communication and teamwork	Noninferior (VR vs live)	All groups received 3 h of team training on nurse-physician communication. IG: virtual environment and avatar, CG: SP. Each scenario lasted approximately 15 - 20 min and was followed by a 30-min facilitator briefing	Sepsis and septic shock scenario
Sanko and McKay [33] (2020), United States	Board simulation (board game) to teach and develop systematic thinking-LF[m]	Notification of incidents and adverse effects. System failures	Favors simulation	CG: Simulation scenarios. IG: 2017 course included a tabletop simulation“ Friday Night in the ER.”	Emergency care
Lee and Kim [34] (2020), Korea	HFS (IAM[n])-HF	Interprofessional communication	Positive association	Before the scenario training on SBAR-R[o] communication and election of the leader of each team. Development of the scenario in 2 times: before a call to a fictitious doctor and performance of the task in a team after receiving instructions by phone	Acute myocardial infarction emergency care

XSL•FO
RenderX

Author (year), country	Modalities and type of simulation	Security aspects covered	Direction of effect	Type of intervention in the study	Case
Wai et al [48] (2021), Korea	HFS-HF	Teamwork	Both improved	IG: combined classroom+simulation. CG 1: online training on team-based learning, individual test and later in group, with those who would solve the simulation. CG 2: clinical case simulation. All groups are training on patient safety, human factors, and communication	Predetermined critical case scenarios: chest pain and weakness in MMII[p]
Musharyanti et al [38] (2021), Indonesia	Role play and simulated patients (actors)-LF	Medication administration	Favors simulation	IG: 18-h training sessions over 5 wk, overview and training activity with a 4C/ID[q] approach including presentation of real cases, small group discussion, reflection and simulation of oral and intramuscular drug administration. CG: 2 wk, overview, 2 lecture sessions and video playback, and posttesting	Oral and intramuscular drug administration
Du et al [39](2021), China	SP in high fidelity room-HF	PUP[r]	Favors simulation	CG: pressure ulcer training in the conventional classroom 90 min, IG: training through simulated clinical scenarios	Three clinical scenarios from admission, hospitalization and deterioration of the disease
Lee and Lim [32] (2021), Korea	HFS (respiratory)-HF	Communication in transfer	Favors simulation	Students were divided into groups of 3 to 4 for 120-min sessions consisting of 50 min of theoretical training and 70 min of simulation-based training	2 patients with respiratory problems with high-fidelity simulator
Craig et al [40] (2021), United States	HFS with computer package, electronic medical records (identification wristbands, carts, barcodes, and computerized records)	Medication administration	Favors simulation	4 wk of simulation. IG: 1. low-fidelity simulation on medication administration. 2. high-fidelity simulation focused on safe medication administration. 3. clinical rotation. 4. high-fidelity simulation+debriefing. CG: 1. standard training. 2 clinical rotation and 3 - 4 same as IG	Scenario for administration of oral and subcutaneous medication (insulin aspart)

XSL•FO
RenderX

Author (year), country	Modalities and type of simulation	Security aspects covered	Direction of effect	Type of intervention in the study	Case
Raurell-Torreda et al [50] (2021), Spain	Role-playing SBAR. -LF	Interprofessional communication	Favors simulation	IG: divided into subgroups of 20 students for 1-h role play session, learning objectives focused on basic professional health care skills, teamwork, use of SBAR worksheet and role distribution in respiratory tract management, nursing procedures and techniques, and use of documentation. Patient assessment and intervention in 3 nursing roles: procedures, assessment, and follow-up	Patient in shock in an emergency department setting (based on a clinical case from the National League for Nursing) to assess teamwork and nontechnical skills
Park and Kim [41] (2021), Korea	HFS (mannequin)-HF	Systemic and organizational factors. Diagnostic errors	Favors simulation	IG: The PDS-IB[s] and CG: simple PDS. The scenario theme for the simulations in both groups was patient deterioration. The simulations for both groups consisted of 1.5 h	Patient with chronic obstructive pulmonary disease was transferred from the emergency room to the inpatient ward (worsening)
Chen et al [51] (2022), Taiwan	SP-HF	Teamwork	Comparable gains	In group 1 (received IPE[t] training, followed by SPE[u]) and group 2 (received SPE training followed by IPE training). Simulation training was structured for 4 wk (3 h/wk) that incorporated a 2-wk IPE program during which medical and nursing students were trained together	Critically ill patients with AHA[v] guidelines for cardiopulmonary resuscitation and emergency cardiac care
Pol-Castañeda et al [42] (2022), Spain	SP (clinical case)-HF	Medications	Favors simulation	Briefing, simulation scenarios were conducted in 24 groups of 6 to 8 students, each playing a different role	3 scenarios: hypocalcemia due to gastrointestinal disease in the emergency department, respiratory infection, paracentesis
Goldsworthy et al [44] (2022), Canada, England, Scotland, and Australia	VR-HF	Patient deterioration	Favors virtual simulation	The treatment group completed 6 VR of medical surgical nursing case studies over 3 wk (2/wk). Two VR were completed each week that they could repeat	Acute deterioration care: angina and cardiac arrest; anaphylaxis; acute asthma exacerbation; COPD[w] and pneumothorax, pulmonary embolism; and blood transfusion reaction.
Li et al [43] (2023), China	HFS (standardized patient+mannequin)-HF	Patient safety, medication errors, and adverse effects	Favors combined training	All: online course-training adverse effects, types, effects, and communication and teamwork. IG: 2 simulation cases in addition to training. CG: online training only	Scenarios of care in respiratory infection (medication) and hemiplegia (basic care)

XSL•FO
RenderX

Author (year), country	Modalities and type of simulation	Security aspects covered	Direction of effect	Type of intervention in the study	Case
Haerling et al [45] (2023), United States	HFS mannequin and VR displays-HF	Safety risks. Interprofessional and patient communication Medication Administration	Manikin superior and equivalent	Clinical experience: 4 h of traditional clinical experience Mannequin simulation: simulation activities with mannequins in 2 scenarios VR was similar to that of the mannequin-based simulations. The groups varied according to the type of experiential learning activity they completed first	Postoperative discharge care and postoperative emergency room readmission
Chou et al [46] (2024), Taiwan	VR-HF	Communication	Favors VR simulation	IG received a VR training in nurse-patient communication skills 2 wk prior to practice. The program was delivered in 4 sessions for 30 min each time for 2 wk. CG received the 30-min nurse-patient communication teaching video that could be downloaded and viewed	Simulated hospital ward scenarios with 4 learning tasks: self-presentation, establishing a nurse-patient relationship, interaction, and medical history collection
Heier et al [35] (2024), Germany	SP-HF	Notification of medication errors	Favors simulation	Interprofessional communication skills training on acute care medical errors (IG) with a cohort that did not receive interprofessional training (CG)	3 scenarios reported in a critical incident reporting system focused on medication errors caused by a chain of errors. Chemotherapy, wrong antibiotic, and chemotherapy preparation with errors

[a]HFS: high-fidelity simulator.

[b]HF: high-fidelity.

[c]ISBAR: Identity-Situation-Background-Assessment-Recommendation.

[d]E: e-learning.

[e] E+S: e-learning plus standard simulation group.

[f] E+PBP: e-learning plus competency-based progression simulation group.

[g]SP: standardized patient.

[h]CG: control group.

[i]SEGUE: Set the stage, Elicit information, Give information, Understand patient perspective, End the encounter.

[j]IG: intervention group.

[k]SBAR: Situation, Background, Assessment, Recommendation.

[l]VR: virtual reality.

[m]LF: low fidelity.

[n]IAM: High Fidelity Simulator for Acute Myocardial Infarction.

[o]SBAR-R: SBAR with Readback and Response.

[p]MMII: Lower Limbs.

[q]4C/ID: Four Components Instructional Design.

[r]PUP: pressure ulcer prevention.

[s]PDS-IB: Patient Deterioration Simulation with Inattentional Blindness.

[t]IPE: Interprofessional Education.

[u]SPE: Single Profession Education.

[v]AHA: American Heart Association.

[w]COPD: chronic obstructive pulmonary disease.

XSL•FO
RenderX

Methodological Quality Results

The overall methodological quality of the studies was moderate to high. Using the JBI Critical Appraisal Tools, RCTs generally scored between 9 and 13 out of 13 possible criteria, quasi-experimental studies between 7 and 9 out of 9, and cross-sectional studies met 6 to 8 out of 8 criteria, assessed according to JBI methodological standards for evidence synthesis [29-31].

Common strengths included clear description of inclusion criteria, valid measurement of outcomes, and appropriate statistical analyses. The most frequent limitations were related to lack of blinding of participants or assessors, incomplete follow-up, and limited control of confounding factors.

A summary of the appraisal results is presented in Table 5, while full details of the assessment can be found in Tables S2, S3, and S4 in Multimedia Appendices 3-5.

Table . Summary of methodological quality results.

Study design	Number of studies	JBI[a] tool used	Mean criteria met (%)	Common strengths	Common limitations
Randomized controlled trials	8	JBI RCT[b] Checklist (13 items)	85 - 100 (9, 13/13)	Clear objectives, valid measurement tools, and appropriate statistical analysis	Lack of blinding (3), small samples
Quasi-experimental studies	11	JBI Quasi-Experimental Checklist (9 items)	78 - 100 (7, 9/9)	Clear cause-effect design and reliable outcome measures	Absence of control group in some cases (3), incomplete follow-up (2)
Cross-sectional studies	1	JBI Analytical Cross-Sectional Checklist (8 items)	75 (6/8)	Clear inclusion criteria, valid exposure, and outcome measures	Confounding not always controlled

[a]JBI: Joanna Briggs Institute.

[b]RCT: randomized controlled trial.

Topics

The topics they discussed were communication, both interprofessional [32,34,46,47,49,50] as communication with the patient [36], teamwork [48,51] aspects related to medication, such as its administration [38,40,42] and notification of errors or adverse effects [33,35,43], systemic and organizational factors [41], falls [37], pressure ulcer prevention [39], patient deterioration [44], and various security aspects a mix of the above [45].

Mixed groups, with nursing and medical students, address topics such as interprofessional communication and teamwork [48-51]; only one of them addresses the issue of medication [35].

Most simulations aimed at developing communication and teamwork skills used high-fidelity modalities, including SP and mannequins [32,34,36,47,48,51] and VR environments [46,49]. One study used role-playing as the primary simulation modality [50].

To develop skills in aspects of medication administration, HFS has been used [40], role play [38], and simulated patients [38,42]. In the case of notification of medication errors and adverse effects, the simulation scenarios have used SPs and mannequins [35,43] and a board game [33]. The rest of the safety aspects discussed have used simulated patients or mannequins, except the study on patient deterioration, which has used VR technology [44].

High-fidelity simulation through HFSs, VR, and SPs predominated across studies. However, low-fidelity approaches, such as role-playing, were also used, either alone or combined with other modalities, particularly in interprofessional communication [50], medication administration [38], and incident or system failure reporting [33].

Safety Outcomes

In general, almost all studies link training through simulation scenarios with an improvement in safety-related skills and knowledge [32-47,50]. One study directly compared live (face-to-face) simulation with immersive VR simulation in scenario-based training [49]. The comparator was therefore 2 different simulation modalities rather than simulation versus traditional teaching. No statistically significant differences were found between live and VR simulation in the development of teamwork and communication skills. These findings suggest that VR-based simulation may represent a comparable alternative to live simulation for fostering these competencies.

Simulation training complemented by other techniques (lectures, presentations, debates, and demonstrations) achieves better results than simulation training alone for improving patient safety competency among nursing students [43]. However, the effects of a blended classroom plus clinical simulation versus clinical simulation alone on teamwork attitudes did not further improve teamwork attitudes, perceptions, and performance in medical and nursing students compared with clinical simulation alone [48]. A quasi-experimental study comparing traditional clinical practice, manikin-based high-fidelity simulation, and screen-based virtual simulation (n=193) found no significant differences in cognitive outcomes between groups [45]. However, students in the manikin-based simulation group achieved significantly higher scores than those in the traditional clinical group in several competency domains measured by the Creighton Competency Evaluation Instrument, including communication (effect size=0.52; P=.04), although no significant differences were observed in the Patient Safety

XSL•FO
RenderX

domain. These findings suggest that the type of simulation experience may influence specific clinical competency outcomes [45].

Only 2 studies followed up over time after training [46,49]. In one study, the results were better than before after 1 week of practice [46]. A RCT compared immersive VR simulation with conventional live simulation in 120 medical and nursing students. Both modalities significantly improved teamwork attitudes and interprofessional socialization immediately postintervention. At 2-month follow-up, only the VR group maintained a significant improvement in Interprofessional Socialization and Valuing Scale scores (P=.047), although no statistically significant differences were found between groups at any time point, suggesting comparable effectiveness [49].

Two studies evaluated competencies using Objective Structured Clinical Examination–based assessments. In one study, 47 second-year nursing students receiving simulation-based training achieved significantly higher scores in pressure ulcer risk assessment than those receiving standard instruction (mean 29.04, SD 6.00 vs mean 12.38, SD 4.15; P<.001) across 3 simulated scenarios using SPs [39]. In another randomized educational trial including 90 nursing and medical students, competency-based simulation combined with e-learning led to higher communication competence (60%) compared with e-learning alone (6.9%) and standard simulation (13%), using a clinical structured communication tool [47].

Discussion

Principal Findings

This systematic review aimed to evaluate the effect of simulation-based education on patient safety outcomes in undergraduate nursing students, to identify the patient safety domains addressed through simulation, and to describe the characteristics of simulation-based educational interventions. Overall, the findings suggest that simulation-based education is associated with improvements in several competencies related to patient safety, particularly communication, teamwork, medication safety, and the recognition and reporting of clinical incidents. High-fidelity simulation predominated among the included interventions, with 16 studies using this modality [32,34-43,45,47,48,50,51]. In addition, the incorporation of VR [35,37,40] expands the range of clinical scenarios that can be recreated in educational environments and may facilitate the development of knowledge and skills related to patient safety [43].

Among the patient safety domains addressed, communication and teamwork were the most frequently studied. Several studies reported improvements in these competencies following simulation-based training among nursing and medical students [38-42]. Simulation scenarios provide opportunities for students to practice interprofessional collaboration and structured information transfer in controlled learning environments. Structured communication tools, such as SBAR and its adaptations, were frequently incorporated into simulation scenarios and were associated with improved team coordination and information transfer [32,34,37,46,47,50]. Similarly, structured communication strategies have been shown to improve communication between students and patients [36]. Communication-focused simulations often involved interprofessional scenarios in which students practiced structured communication during clinical deterioration or emergencies, strengthening collaborative decision-making between professionals [47,49,50].

Medication safety was another frequently addressed domain. Three studies evaluated students' ability to safely administer medications, identify potential errors, and follow appropriate safety procedures [38,40,42]. Simulation provides a controlled environment in which students can practice medication administration while recognizing potential safety risks without endangering real patients. In addition, simulation-based education was associated with improvements in the reporting of adverse events and medication-related incidents [40,42]. Other 3 studies also explored the development of systems thinking and incident reporting competencies through simulation scenarios focused on adverse event reporting or clinical error disclosure [33,35,43]. These activities encouraged students to recognize systemic factors contributing to patient safety incidents and to adopt a nonpunitive perspective toward error management.

This review also highlights the diversity of simulation modalities used to address patient safety competencies. SPs were commonly used to train communication with patients [36], reporting incidents such as falls or medication errors [35,37], pressure ulcer prevention [39], teamwork and interprofessional communication [49,51], and medication administration [38,42]. High-fidelity simulation using mannequins was the most frequently used modality [32,34,39-41,43,45,47,48]. VR simulations were also used, particularly in scenarios focused on communication and teamwork [35,46,49].

Although high-fidelity simulation was widely used, lower-technology simulation approaches also demonstrated positive learning outcomes. Role-playing, tabletop simulations, and screen-based virtual simulations improved competencies related to communication, systems thinking, and incident reporting [33,38,50]. These findings suggest that simulation-based education can be implemented across a variety of educational settings, including institutions with limited technological resources. In addition to simulation modality, the instructional design of the intervention influenced outcomes. One study reported greater improvements when simulation was integrated with complementary educational strategies such as lectures, structured feedback, or online modules [43].

Comparison With Previous Literature

The findings of this review are consistent with previous literature highlighting the educational value of simulation in health professions education [5252]. Improvements in communication and teamwork competencies observed in the included studies align with recent meta-analyses emphasizing the importance of interprofessional education for improving role understanding and collaborative practice among health care students [53,54]. Improvements in these competencies are particularly relevant to patient safety, as failures in communication and teamwork are frequently associated with preventable adverse events in

XSL•FO
RenderX

clinical practice. Interprofessional simulation experiences have been shown to strengthen role clarity and collaborative practice, thereby promoting safety-oriented behaviors among future health professionals [54].

Similarly, the improvements observed in medication safety practices are consistent with previous reviews suggesting that simulation-based education can be an effective method for improving patient safety competencies [24]. Evidence from systematic reviews also indicates that the use of SPs may enhance communication skills, learning outcomes, and problem-solving abilities in health professions education [23]. Previous reviews have also reported improvements in knowledge, confidence, and clinical skills following simulation-based learning [21,22], although many of these studies did not specifically focus on patient safety outcomes. Compared with earlier literature, the present review provides a more focused synthesis of evidence regarding the use of simulation-based education to address patient safety competencies in undergraduate nursing students.

The findings of this review also suggest that the use of different simulation methods and varying levels of fidelity, as observed across the included studies, can be similarly effective for developing competencies related to patient safety. These results are consistent with previous literature, indicating that the effectiveness of simulation-based education does not depend solely on the level of technological fidelity. Studies comparing high-fidelity simulation with alternative teaching strategies, such as written case studies, have not demonstrated clear advantages of high-fidelity simulation alone for improving critical thinking skills in nursing students [55].

Sustainability of Training Effects

Despite the positive outcomes reported in most studies, the long-term sustainability of simulation-based training effects remains unclear. Only 2 studies included follow-up assessments after the intervention, with relatively short follow-up periods of 1 week and 2 months [46,49]. Consequently, it is difficult to determine whether the improvements observed are maintained over time or translate into sustained behavioral changes in clinical practice. This limitation is particularly relevant because previous research on medication safety suggests that safety-related competencies may decline if they are not reinforced through continued practice and training [56].

Practical Implications for Nursing Education

The findings of this systematic review have several important practical implications for undergraduate nursing education. Overall, the evidence indicates that simulation-based education is an effective strategy for developing key patient safety competencies, particularly in communication, teamwork, medication safety, and error recognition [32-47].

The results support the systematic integration of simulation-based learning into nursing curricula, rather than its use as an isolated or supplementary teaching activity. Simulation-based interventions embedded within educational programs improved students' patient safety competencies [40,42,43,45,46]. Simulation scenarios focused on real-world patient safety challenges enable students to apply theoretical knowledge in a safe and controlled environment [39,41,44].

The impact of structured communication tools, such as SBAR, ISBAR (Identification, Situation, Background, Assessment, Recommendation), and SEGUE, suggests that these frameworks should be explicitly incorporated into simulation-based training [32,34,36,37,47]. The use of these tools during simulation may reduce communication-related errors and enhance patient safety.

In terms of the characteristics of the simulation, the included studies used various methods, including manikin simulation, SPs, virtual simulation, role-playing, and tabletop simulation. The level of fidelity within these modalities varied between low, medium, and high, depending on the technological complexity and degree of clinical realism described by the authors. While experiences with a higher level of technological fidelity were common, several studies demonstrated that modalities with lower technological requirements, such as role-playing or tabletop simulation [33,38,50], as well as basic virtual environments [44-46,49], could also be effective in improving patient safety knowledge and skills. These findings suggest that meaningful educational outcomes can be achieved with modalities that have lower technological requirements, which is particularly relevant for institutions with limited resources.

Current Gaps in the Literature

Even with mounting evidence backing simulation-based training in nursing, this review identified several important gaps in existing literature. One of the most significant limitations is the lack of long-term follow-up. Only 2 of the studies included assessed outcomes beyond the immediate postintervention period, with follow-up periods of 1 week and 2 months, respectively [46,49].

Another gap relates to study design and methodological rigor. Although RCTs were included, 9 of studies used quasi-experimental designs with convenience sampling and relatively small sample sizes [24,25,29-31,33-36]. In addition, blinding of participants and assessors was rarely reported. Also, studies used a wide range of instruments to assess patient safety competencies, including self-developed questionnaires and modified scales [25,27,29,33,35].

Geographically, most studies were conducted in high-income countries [22-42], with limited representation from low- and middle-income settings, indicating a lack of global diversity in the evidence base.

Recommendations for Future Research

Future research should address the limitations identified in this review. Longitudinal studies are needed to assess retention and maintenance of competencies. The use of standardized tools and methodologically sound, multicenter clinical trials would improve external validity and avoid bias. On the other hand, studies comparing the efficiency and cost-effectiveness of different simulation modalities would help guide curricula and resource allocation. It would be necessary to include other educational contexts in low- and middle-income countries in order to apply the results and guide nursing training plans. The

XSL•FO
RenderX

limited reporting of facilitator preparation across studies highlights the need for future research to examine the role of educator training in simulation-based patient safety education.

Limitations

It is important to consider the limitations of this study when interpreting the results. First, there is a considerable degree of heterogeneity between the studies, not only in terms of the type of intervention and duration, but also in terms of the training of the participants, experience, and origin. Additionally, the variability of measurement instruments, convenience samples, and the use of self-reported data may introduce bias. Furthermore, the restriction of the search to studies published in English, Spanish, and Portuguese may have resulted in the exclusion of some relevant studies. The criteria used to establish the cutoff point in the methodological quality assessment may have influenced the final selection of evidence included.

The search strategy was limited to the terms “patient safety” and “safety,” which may have led to the exclusion of relevant studies addressing patient safety outcomes. Although additional manual screening of references was performed to identify potentially relevant studies, this restriction may have reduced the sensitivity of the search.

Although a systematic approach to study selection ensures the quality of the evidence, it may result in the omission of relevant research that, although less rigorous, could offer valuable insights.

Conclusions

Despite considerable discrepancies between individual studies, simulation is an important tool that empowers nursing students to identify, mitigate, or eliminate potential risks to patient safety. The use of simulation methodology facilitates the acquisition of essential competencies, including communication skills, teamwork, medication administration, and error detection and adverse effect recognition. These skills are crucial for ensuring patient safety upon entering the professional workforce. The use of high-fidelity simulation enables the recreation of clinical scenarios, facilitating the integration of theoretical and practical training, and thus the development of skills and a more comprehensive integration of knowledge.

Simulation-based education improves nursing students’ competence in key safety domains (communication, teamwork, medication safety, and error recognition), thus contributing to improved patient safety outcomes.

Acknowledgments

No artificial intelligence (AI) was used in the preparation of this manuscript.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of the studies.
[PDF File, 91 KB - nursing_v9i1e87898_app1.pdf]

Multimedia Appendix 2
Characteristics of the interventions.
[PDF File, 68 KB - nursing_v9i1e87898_app2.pdf]

Multimedia Appendix 3
Quality assessment 1.
[PDF File, 36 KB - nursing_v9i1e87898_app3.pdf]

Multimedia Appendix 4
Quality assessment 2.
[PDF File, 56 KB - nursing_v9i1e87898_app4.pdf]

Multimedia Appendix 5
Quality assessment 3.
[PDF File, 60 KB - nursing_v9i1e87898_app5.pdf]

Checklist 1

XSL•FO
RenderX

PRISMA checklist.
[PDF File, 82 KB - nursing_v9i1e87898_app6.pdf]

References

1. Patient safety strategy for the national health system, 2015-2020 period. Ministerio de Sanidad, Servicios Sociales e Igualdad. 2016. URL: https://seguridaddelpaciente.sanidad.gob.es/docs/Estrategia_Seguridad_del_Paciente_2015-2020.pdf [accessed 2024-06-25]
2. Patient safety. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/patient-safety [accessed 2024-07-14]
3. Slawomirski L, Klazinga N. The economics of patient safety: from analysis to action. : OECD; 2022 URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2022/08/the-economics-of-patient-safety_dda2a072/761f2da8-en.pdf [accessed 2026-05-17] [doi: 10.1787/761f2da8-en]
4. Kohn L, Corrigan J, Donaldson M. To Err Is Human: Building a Safer Health System: National Academies Press; 2000. URL: https://www.ncbi.nlm.nih.gov/books/NBK225182/ [accessed 2024-07-14]
5. Sammer CE, Lykens K, Singh KP, Mains DA, Lackan NA. What is patient safety culture? A review of the literature. J Nurs Scholarsh 2010 Jun;42(2):156-165. [doi: 10.1111/j.1547-5069.2009.01330.x] [Medline: 20618600]
6. Seventy-second world health assembly, WHA726, agenda item 125: global action on patient safety. : World Health Organization; 2019 May 28 URL: https://apps.who.int/gb/ebwha/pdf_files/WHA72/A72_R6-en.pdf [accessed 2026-05-17]
7. Global patient safety action plan 2021-2030: towards eliminating avoidable harm in health care. : World Health Organization; 2021 URL: https://iris.who.int/bitstream/handle/10665/343477/9789240032705-eng.pdf [accessed 2026-04-15]
8. Macrae C. Imitating incidents: how simulation can improve safety investigation and learning from adverse events. Simul Healthc 2018 Aug;13(4):227-232. [doi: 10.1097/SIH.0000000000000315] [Medline: 29771816]
9. Putz F, Kattan E, Maestre JM. The use of clinical simulation to train teams in conflict management within healthcare: an exploratory systematic review. Enfermería Clínica 2022 Jan;32(1):21-32. [doi: 10.1016/j.enfcli.2020.10.032]
10. Valencia Castro JLV, Tapia Vallejo S, Olivares Olivares SLO. Clinical simulation as a strategy for developing critical thinking in medical students. Inv Ed Med 2019;8(29):13-22. [doi: 10.1016/j.riem.2016.08.003]
11. Meakim C, Boese T, Decker S, et al. Standards of best practice: simulation standard I: terminology. Clin Simul Nurs 2013 Jun;9(6):S3-S11. [doi: 10.1016/j.ecns.2013.04.001]
12. Terminology and Concepts Working Group. In: Lioce L, Lopreiato J, Anderson M, editors. Healthcare Simulation Dictionary, 3rd edition: Agency for Healthcare Research and Quality; 2025. URL: https://www.ahrq.gov/patient-safety/resources/simulation/terms.html [accessed 2026-05-19]
13. Martínez Arce A, Araujo Blesa M, Tovar Reinoso A, Rodríguez Gómez P, Vélez Vélez E, García-Carpintero Blas E. Integration of clinical simulation teaching methods into the nursing degree curriculum. Enferm Docente 2021:17-22 [FREE Full text]
14. Zaheer S, Ginsburg L, Wong HJ, Thomson K, Bain L, Wulffhart Z. Acute care nurses' perceptions of leadership, teamwork, turnover intention and patient safety - a mixed methods study. BMC Nurs 2021 Jul 30;20(1):134. [doi: 10.1186/s12912-021-00652-w] [Medline: 34330272]
15. Altmiller G, Pepe LH. Influence of technology in supporting quality and safety in nursing education. Nurs Clin North Am 2022 Dec;57(4):551-562. [doi: 10.1016/j.cnur.2022.06.005] [Medline: 36280294]
16. Hémon B, Michinov E, Guy D, Mancheron P, Scipion A. Speaking up about errors in routine clinical practice: a simulation-based intervention with nursing students. Clin Simul Nurs 2020 Aug;45:32-41. [doi: 10.1016/j.ecns.2020.03.003]
17. Diaz-Navarro C, Armstrong R, Charnetski M, et al. Global consensus statement on simulation-based practice in healthcare. Clin Simul Nurs 2024 Aug;93:101552. [doi: 10.1016/j.ecns.2024.101552]
18. Serrat N, Camps A. Simulation as a teaching method in university classrooms: an introduction. Octadero. 2023. URL: https://octaedro.com/wp-content/uploads/2023/06/16547-CDU-47.pdf [accessed 2024-07-18]
19. Lee SE, Morse BL, Kim NW. Patient safety educational interventions: a systematic review with recommendations for nurse educators. Nurs Open 2022 Jul;9(4):1967-1979. [doi: 10.1002/nop2.955] [Medline: 34047058]
20. De Rezende H, Vitorio AMF, Morais AS, et al. Effectiveness of educational interventions to develop patient safety knowledge, skills, behaviours and attitudes in undergraduate nursing students: a systematic review protocol. BMJ Open 2022 Mar 8;12(3):e058888. [doi: 10.1136/bmjopen-2021-058888] [Medline: 35260461]
21. Alharbi A, Nurfianti A, Mullen RF, McClure JD, Miller WH. The effectiveness of simulation-based learning (SBL) on students' knowledge and skills in nursing programs: a systematic review. BMC Med Educ 2024 Oct 7;24(1):1099. [doi: 10.1186/s12909-024-06080-z] [Medline: 39375684]
22. Sim JJM, Rusli KDB, Seah B, Levett-Jones T, Lau Y, Liaw SY. Virtual simulation to enhance clinical reasoning in nursing: a systematic review and meta-analysis. Clin Simul Nurs 2022 Aug;69:26-39. [doi: 10.1016/j.ecns.2022.05.006] [Medline: 35754937]
23. Ma J, Lee Y, Kang J. Standardized patient simulation for more effective undergraduate nursing education: a systematic review and meta-analysis. Clin Simul Nurs 2023 Jan;74:19-37. [doi: 10.1016/j.ecns.2022.10.002]

XSL•FO
RenderX

24. El Hussein MT, Harvey G, Bell N. The influence of nursing simulation on patient outcomes and patient safety: a scoping review. Clin Simul Nurs 2022 Sep;70:37-46. [doi: 10.1016/j.ecns.2022.06.004]
25. Wallace S, Schuler MS, Kaulback M, Hunt K, Baker M. Nursing student experiences of remote learning during the COVID-19 pandemic. Nurs Forum 2021 Jul;56(3):612-618. [doi: 10.1111/nuf.12568] [Medline: 33728660]
26. Ramos-Morcillo AJ, Leal-Costa C, Moral-García JE, Ruzafa-Martínez M. Experiences of nursing students during the abrupt change from face-to-face to e-learning education during the first month of confinement due to COVID-19 in Spain. Int J Environ Res Public Health 2020 Jul 30;17(15):5519. [doi: 10.3390/ijerph17155519] [Medline: 32751660]
27. PRISMA 2020. PRISMA. URL: https://www.prisma-statement.org/prisma-2020 [accessed 2024-07-14]
28. Critical appraisal tools. JBI. URL: https://jbi.global/critical-appraisal-tools [accessed 2024-07-14]
29. Barker TH, Habibi N, Aromataris E, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for quasi-experimental studies. JBI Evid Synth 2024 Mar 1;22(3):378-388. [doi: 10.11124/JBIES-23-00268] [Medline: 38287725]
30. Barker TH, Stone JC, Sears K, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for randomized controlled trials. JBI Evid Synth 2023 Mar 1;21(3):494-506. [doi: 10.11124/JBIES-22-00430] [Medline: 36727247]
31. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R, et al. Chapter 7: systematic reviews of etiology and risk. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: Adelaide: JBI; 2020. URL: https://synthesismanual.jbi.global [accessed 2026-01-10] [doi: 10.46658/JBIRM-17-06]
32. Lee DH, Lim EJ. Effect of a simulation-based handover education program for nursing students: a quasi-experimental design. Int J Environ Res Public Health 2021 May 28;18(11):5821. [doi: 10.3390/ijerph18115821] [Medline: 34071549]
33. Sanko JS, Mckay M. Participation in a system-thinking simulation experience changes adverse event reporting. Sim Healthcare 2020;15(3):167-171. [doi: 10.1097/SIH.0000000000000473]
34. Lee KR, Kim EJ. Relationship between interprofessional communication and team task performance. Clin Simul Nurs 2020 Jun;43:44-50. [doi: 10.1016/j.ecns.2020.02.002]
35. Heier L, Schellenberger B, Schippers A, Nies S, Geiser F, Ernstmann N. Interprofessional communication skills training to improve medical students' and nursing trainees' error communication - quasi-experimental pilot study. BMC Med Educ 2024 Jan 3;24(1):10. [doi: 10.1186/s12909-023-04997-5] [Medline: 38172793]
36. Son HK, Hee Kim D. Effect of SEGUE-based communication education on nursing simulation practice: a quasi-experimental design. Contemp Nurse 2019;55(4-5):330-340. [doi: 10.1080/10376178.2019.1641421] [Medline: 31280701]
37. Jeong JH, Kim EJ. Development and evaluation of an SBAR-based Fall Simulation Program for nursing students. Asian Nurs Res (Korean Soc Nurs Sci) 2020 May;14(2):114-121. [doi: 10.1016/j.anr.2020.04.004] [Medline: 32335316]
38. Musharyanti L, Haryanti F, Claramita M. Improving nursing students' medication safety knowledge and skills on using the 4C/ID learning model. J Multidiscip Healthc 2021;14:287-295. [doi: 10.2147/JMDH.S293917] [Medline: 33574673]
39. Du YL, Ma CH, Liao YF, Wang L, Zhang Y, Niu G. Is clinical scenario simulation teaching effective in cultivating the competency of nursing students to recognize and assess the risk of pressure ulcers? Risk Manag Healthc Policy 2021;14:2887-2896. [doi: 10.2147/RMHP.S315138] [Medline: 34267566]
40. Craig SJ, Kastello JC, Cieslowski BJ, Rovnyak V. Simulation strategies to increase nursing student clinical competence in safe medication administration practices: a quasi-experimental study. Nurse Educ Today 2021 Jan;96:104605. [doi: 10.1016/j.nedt.2020.104605] [Medline: 33096362]
41. Park J, Kim KJ. Effects of patient deterioration simulation using inattentional blindness for final year nursing students: a randomized controlled trial. Nurse Educ Today 2021 Nov;106:105080. [doi: 10.1016/j.nedt.2021.105080] [Medline: 34340194]
42. Pol-Castañeda S, Carrero-Planells A, Moreno-Mulet C. Use of simulation to improve nursing students' medication administration competence: a mixed-method study. BMC Nurs 2022 May 16;21(1):117. [doi: 10.1186/s12912-022-00897-z] [Medline: 35578199]
43. Li H, Zhao W, Li B, Li Y. Effects of the small private online course combined with simulation-based training in a patient safety education program among nursing students: a quasi-experimental study. Int J Nurs Sci 2023 Oct;10(4):555-561. [doi: 10.1016/j.ijnss.2023.09.014] [Medline: 38020829]
44. Goldsworthy S, Muir N, Baron S, et al. The impact of virtual simulation on the recognition and response to the rapidly deteriorating patient among undergraduate nursing students. Nurse Educ Today 2022 Mar;110:105264. [doi: 10.1016/j.nedt.2021.105264] [Medline: 35030352]
45. Haerling K, Kmail Z, Buckingham A. Contributing to evidence-based regulatory decisions: a comparison of traditional clinical experience, mannequin-based simulation, and screen-based virtual simulation. J Nurs Regul 2023 Jan;13(4):33-43. [doi: 10.1016/S2155-8256(23)00029-7]
46. Chou CH, Tai HC, Chen SL. The effects of introducing virtual reality communication simulation in students' learning in a fundamentals of nursing practicum: a pragmatic randomized control trials. Nurse Educ Pract 2024 Jan;74:103837. [doi: 10.1016/j.nepr.2023.103837] [Medline: 38006647]
47. Breen D, O'Brien S, McCarthy N, Gallagher A, Walshe N. Effect of a proficiency-based progression simulation programme on clinical communication for the deteriorating patient: a randomised controlled trial. BMJ Open 2019 Jul 9;9(7):e025992. [doi: 10.1136/bmjopen-2018-025992] [Medline: 31289064]

XSL•FO
RenderX

48. Wai AK, Lam VS, Ng ZL, et al. Exploring the role of simulation to foster interprofessional teamwork among medical and nursing students: a mixed-method pilot investigation in Hong Kong. J Interprof Care 2021;35(6):890-898. [doi: 10.1080/13561820.2020.1831451] [Medline: 33290116]
49. Liaw SY, Ooi SW, Rusli KDB, Lau TC, Tam WWS, Chua WL. Nurse-physician communication team training in virtual reality versus live simulations: randomized controlled trial on team communication and teamwork attitudes. J Med Internet Res 2020 Apr 8;22(4):e17279. [doi: 10.2196/17279] [Medline: 32267235]
50. Raurell-Torredà M, Rascón-Hernán C, Malagón-Aguilera C, et al. Effectiveness of a training intervention to improve communication between/awareness of team roles: a randomized clinical trial. J Prof Nurs 2021;37(2):479-487. [doi: 10.1016/j.profnurs.2020.11.003] [Medline: 33867108]
51. Chen HW, O'Donnell JM, Chiu YJ, et al. Comparison of learning outcomes of interprofessional education simulation with traditional single-profession education simulation: a mixed-methods study. BMC Med Educ 2022 Aug 30;22(1):651. [doi: 10.1186/s12909-022-03640-z] [Medline: 36042449]
52. Tonapa SI, Mulyadi M, Ho KHM, Efendi F. Effectiveness of using high-fidelity simulation on learning outcomes in undergraduate nursing education: systematic review and meta-analysis. Eur Rev Med Pharmacol Sci 2023 Jan;27(2):444-458. [doi: 10.26355/eurrev_202301_31040] [Medline: 36734697]
53. Sezgin MG, Bektas H. Effectiveness of interprofessional simulation-based education programs to improve teamwork and communication for students in the healthcare profession: a systematic review and meta-analysis of randomized controlled trials. Nurse Educ Today 2023 Jan;120:105619. [doi: 10.1016/j.nedt.2022.105619] [Medline: 36343420]
54. Li YY, Au ML, Tong LK, Ng WI, Wang SC. High-fidelity simulation in undergraduate nursing education: a meta-analysis. Nurse Educ Today 2022 Apr;111:105291. [doi: 10.1016/j.nedt.2022.105291] [Medline: 35158134]
55. Blakeslee JR. Effects of high-fidelity simulation on the critical thinking skills of baccalaureate nursing students: a causal-comparative research study. Nurse Educ Today 2020 Sep;92:104494. [doi: 10.1016/j.nedt.2020.104494] [Medline: 32544764]
56. Fusco LA, Alfes CM, Weaver A, Zimmermann E. Medication safety competence of undergraduate nursing students. Clin Simul Nurs 2021 Mar;52:1-7. [doi: 10.1016/j.ecns.2020.12.003]

Abbreviations

HFS: high-fidelity simulator
ISBAR: Identification, Situation, Background, Assessment, Recommendation
JBI: Joanna Briggs Institute
MeSH: Medical Subject Headings
PIO: Patient, Intervention, Outcome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SBAR: Situation, Background, Assessment, Recommendation
SEGUE: Set the stage, Elicit information, Give information, Understand the patient's perspective, End the encounter
SP: standardized patient
VR: virtual reality
WHO: World Health Organization

Edited by M Kleib; submitted 16.Nov.2025; peer-reviewed by F Alshehri, S Jones; revised version received 10.Mar.2026; accepted 23.Mar.2026; published 26.May.2026.

Please cite as:
González-Sanz A, López-García JC, Sutil-Rodríguez E, Juárez-Vela R, Santos-Sánchez JÁ, Navas-Echazarreta N, Martínez-Sabater A, Sancho-Sanchez C
Simulation-Based Training for Nursing Students to Improve Patient Safety: Systematic Review
JMIR Nursing 2026;9:e87898
URL: https://nursing.jmir.org/2026/1/e87898
doi:10.2196/87898

© Azucena González-Sanz, José Carlos López-García, Elena Sutil-Rodríguez, Raúl Juárez-Vela, José Ángel Santos-Sánchez, Noelia Navas-Echazarreta, Antonio Martínez-Sabater, Consuelo Sancho-Sanchez. Originally published in JMIR Nursing (https://nursing.jmir.org), 26.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Nursing Informaticians in Spain: Scoping Review and Expert-Validated Gap Analysis

Adrián Marco-Moyano[1], RN, BSN, MSN; Patricia Verdú Rodríguez[1], BSBE, MSBI; Manuel Lillo-Crespo[2], RN, BSN, MBAExec, MSN, MSCAnthro, PhD

[1]University of Alicante, San Vicente del Raspeig, Alicante, Spain

[2]Faculty of Health Sciences, University of Alicante, Carretera de San Vicente del Raspeig, s/n, San Vicente del Raspeig, Alicante, Spain

Corresponding Author:
Manuel Lillo-Crespo, RN, BSN, MBAExec, MSN, MSCAnthro, PhD
Faculty of Health Sciences, University of Alicante, Carretera de San Vicente del Raspeig, s/n, San Vicente del Raspeig, Alicante, Spain

Related Article:

This is a corrected version. See correction statement: https://nursing.jmir.org/2026/1/e100481

Abstract

Background: The role of nursing informaticians is well-established in countries like the United States, Canada, and Australia, supported by competency frameworks and educational programs that enable nurses to lead technological integration in health care. However, in Spain, this role is not formally recognized, and specialized university training is scarce, creating a significant gap in digital health leadership among nurses.

Objective: The study aimed to analyze the international landscape of the nursing informatician role, comparatively focusing on the situation in Spain, to subsequently identify the specific gaps for its implementation through experts' views and insights.

Methods: First, a scoping review following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines was conducted in English and Spanish using scientific evidence searched in PubMed, Scopus, CINAHL, and Web of Science from 2018 to the present, as well as gray literature on the topic. A total of 55 published studies were included after screening 1356 records and 10 gray literature documents. Subsequently, findings were validated through a gap analysis comprising a panel of 10 experts selected according to their experience in digital literacy.

Results: The review identified 6 core competencies for nursing informaticians: information management, cybersecurity, patient safety, evaluation and development of clinical information systems, leadership and coordination of digital tools, implementation of new technologies and specialized applications, and education and digitalization in health. Internationally, training is delivered via postgraduate programs, certifications, and leadership initiatives. Experts validated the relevance of these competencies for Spain (rated 5/5) and the applicability and desirability of implementing training programs (rated 4.8/5). Key barriers identified were the lack of official recognition, scarce training, and organizational resistance to change.

Conclusions: There is a contrast between the established role of nursing informaticians internationally and its absence in Spain. The lack of a formal framework and specific training programs is the primary barrier. Implementing validated competencies and tailored educational strategies is crucial for Spain to advance its digital health transformation and empower nursing leadership in technology.

(JMIR Nursing 2026;9:e83373) doi:10.2196/83373

KEYWORDS

nursing informatics; nursing informaticians; enfermeras especialistas en sistemas de información; digital health; competency-based education

Introduction

Global Perspective

Nursing informatics (NI) is a discipline that has gained relevance in recent years due to the increasing integration of information and communication technologies. The definition of NI includes the integration of information technologies in patient care and the clinical setting. Digital transformation facilitates and enhances diagnosis, data management, information, and optimization of processes. This can lead to improved quality of care and efficiency of the health care system [1]. Within this transformation, NI emerges as a crucial discipline that integrates

XSL•FO
RenderX

nursing science, information technology, innovative technologies, and data analytics to manage and communicate data, information, technology integration, and nursing practice [1,2]. The professionals who lead this charge are nursing informaticians (in Spanish "Enfermeras Especialistas en Sistemas de Información" or "Enfermeras Informáticas" [EESIs]), who serve as the vital connection between clinical practice, information systems, and technical teams [3].

Literature Review

In countries like the United States, Canada, and Australia, this advanced role is well-established and robustly supported. The presence of professional bodies such as the American Medical Informatics Association, the Healthcare Information and Management Systems Society, the Canadian Nursing Informatics Association, and the Health Informatics Society of Australia provides a structured framework for professional development. These frameworks help define the role of nursing informaticians as a link between clinical practice and information systems that analyze, design, implement, and evaluate health information technologies. "Nursing informatician" is a generic term that includes variations depending on the formative level and leadership. For example, "nurse informaticist" or "nursing informatics specialist" is a common and specific job title for a professional with advanced knowledge [4]. Other alternatives are "chief nursing informatics officer," an executive role that provides an informatics strategy focused on nursing practice, and "chief X informatics officer," which includes clinical informatics leaders from various disciplines [5]. Furthermore, universities in these nations have prioritized the integration of informatics competencies into nursing curricula, empowering graduates to lead the use of technology in clinical practice [6]. This combination of professional communities, academic programs, and defined competency frameworks enables nursing informaticians to lead high-impact projects in the digital transformation of health care systems.

This leadership is present in several key areas. For instance, nursing informaticians can contribute to the development of interactive digital platforms to supervise and ensure the quality of clinical placements. In this context, a user-centered design has proven fundamental to successful adoption by supervisors and students [7,8]. Their expertise becomes essential in guiding institutions through the creation or development of e-learning platforms, where success depends on a broad and systematic analysis of needs rather than on technical specifications alone [9]. Moreover, they command the "co-creation" of educational resources, mostly virtual teaching packages. This methodology involves end users in the design process, creating training tools perceived as relevant and useful by them [10].

Emerging technologies, such as artificial intelligence, language processing models, big data, and 3D printing, can also be part of these nurses' competencies. However, without appropriate training and a legislative framework, the integration of these advances into clinical practice can be challenging. Their training enables them to stay updated on technological advancements [11].

These examples of NI competencies that are commonly assumed in the international context by their role underscore a profound functional gap in Spain. There, the nursing informatician's role is not formally recognized, and relevant university training is scarce. This absence raises a key question: "Why is the concept of nursing informaticians developed and implemented globally while absent in Spain?" As experts point out, digital health in the country faces significant challenges due to a lack of specialized training and a defined competency framework for nursing in this field [8]. This requires defining specific competencies that go beyond just using clinical information systems [12]. Furthermore, nursing students' basic digital literacy in their current studies is insufficient [13], so specialized training becomes essential to prepare future nurses to use and integrate technology in health care [14].

Research Objectives

Given the differences between international and national paradigms, a comprehensive analysis was required. This led to a 2-phase study design comprising 2 objectives: the first was to analyze the international scientific and gray literature on the nursing informaticians' role, comparatively focusing on the situation in Spain. The second objective consisted of validating these findings by identifying the specific barriers and opportunities for implementing the role in Spain according to experts' views and insights.

Methods

In line with the objectives, the design was divided into 2 phases: the first was related to the main objective for the scoping review, and the second phase focused on the second aim related to a gap analysis.

Phase 1

A scoping review was chosen to conduct a comprehensive analysis of existing literature, identify research gaps in publications, and examine the diverse evidence surrounding the nursing informaticians' role, as well as gray literature. Analysis is required before drawing specific conclusions [15]. To guide this review, a research question was framed using the population, concept, context framework:

- Population: the nursing profession and, specifically, the role of nursing informaticians
- Concept: professional roles, competencies, training models, barriers, and facilitators for the implementation of the nursing informaticians' role
- Context: the international health care landscape, considering experienced countries in NI, is used as a benchmark to analyze the existing gap in the Spanish health care system

The scoping review was conducted following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines (Checklist 1) [16]. A comprehensive search strategy was developed to reflect the international scope and the specific focus on Spain. Key terms in both languages, English and Spanish, included "nursing informatics," "nursing informatics," "nursing informaticians," "informática aplicada a la enfermería," "enfermeras especialistas en sistemas de información," and "enfermería informática." The search was performed for articles published from 2018 to the present across high-impact

XSL•FO
RenderX

databases: PubMed, Scopus, CINAHL, and Web of Science. Filters for language, document type, and full-text accessibility were applied. To ensure a comprehensive overview, the search was supplemented with gray literature from international NI associations, relevant institutional documents, and university curricula in Spain related to the discipline.

Inclusion and Exclusion Criteria

Articles selected during the methodological analysis process contained one or more of the keywords corresponding to NI and nursing informaticians. In the next step, the content of the screened articles was studied to determine their relevance to the current study.

During the screening process, the following exclusion criteria were applied to scientific articles:

- Nonspecific article content despite mentioning the search terms
- Low academic quality: studies with inconsistent methodology or insufficient bibliographic support
- Full text unavailable
- Article irretrievable from repositories

After screening, a review of all literature was conducted following the steps recommended by the Joanna Briggs Institute (Table 1) for conducting scoping reviews. This process included identifying the research question, searching for relevant evidence, selecting studies, extracting data, and presenting the results [15].

XSL•FO
RenderX

Table . Joanna Briggs Institute appraisal tools.

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Reid L et al [1]. Defining nursing informatics: a narrative review.	Narrative Review	JBI[a] Critical Appraisal Checklist for Textual Evidence: Narrative	1[b]	1	1	1	1	3[c]	3	1	3	1	1	1
Sensmeier J [2]. The value and impact of the alliance for nursing informatics.	Program Evaluation	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	__[d]	—	—	—	—	1
Kirchner RB [3]. Presentando a la enfermera especialista en sistemas de información.	Narrative Article	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	—	—	—	—	—	1
Benavente-Rubio A [6]. El rol de enfermería en la salud digital.	Systematic Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	3	3	1	1	1	1	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Konstantinidis ST [7]. HANDAL digital interactive platform for European and national placements audit for medicine and allied health professions following a user-centered design.	Book chapter	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	—	—	—	—	—	1
Papamalis F [8]. Ensuring quality health care practice for doctors and medical allied professionals through a digital interactive audit platform.	Conference paper	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	—	—	—	—	—	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Baggia A [9]. Ecosystem of organizations in the digital age: Conference proceedings.	Conference paper	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	—	—	—	—	—	1
Konstantinidis S [10]. Co-creation of a virtual interactive teaching package for auditors of health care placements-towards assurance of quality of health care traineeships.	Conference paper	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	—	—	—	—	—	1
Ball MJ. et al [11]. The health informatics series: evolving with a new discipline.	Narrative Review	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	—	—	—	—	—	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
León Molina J [12]. Papel de enfermería en el sistema de información hospitalario.	Narrative Article	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	—	—	—	—	—	1
Chipps J et al [13]. Nursing informatics skills relevance and competence.	Interview-based descriptive study	JBI Critical Appraisal Checklist for Qualitative Research	1	1	1	1	3	3	1	1	—	—	—	1
Backonja U [14]. How to support the nursing informatics leadership pipeline.	Descriptive cross-sectional study	JBI Critical Appraisal Checklist for Analytical Cross Sectional Studies	1	1	1	1	3	3	1	1	—	—	—	1
Peters MDJ [15]. Updated methodological guidance for the conduct of scoping reviews.	Publication guideline	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	1	1	—	—	—	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Page MJ et al [16]. Declaración PRISMA 2020.	Publication guideline	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	1	1	—	—	—	1
Luppa N and Suresh S [4]. Physician and nurse informatics collaboration.	White Paper	JBI Critical Appraisal Checklist for Textual Evidence: Policy	1	1	1	3	1	2[e]	1	—	—	—	—	1
Conte G et al [17]. Embracing digital and technological solutions in nursing.	Scoping Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	3	1	1	1
Farzandipur M et al [18]. Designing a national model for assessment of nursing informatics competency.	Quantitative study	JBI Critical Appraisal Checklist for Analytical Cross Sectional Studies	1	1	1	1	3	3	1	1	—	—	—	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Farokhzadian J et al [19]. Necessary prerequisites for evidence-based practice.	Cross-sectional study	JBI Critical Appraisal Checklist for Analytical Cross Sectional Studies	1	1	1	1	1	1	1	1	—	—	—	1
Kulju E et al [20]. Educational interventions and their effects on health care professionals' digital competence development.	Narrative Review	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	1	2	3	1	1	1
Saco P [21]. Informática aplicada a la enfermería.	University Subject Guide	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	2	1	1	—	—	—	—	—	1
Nes AAG et al [22]. Technological literacy in nursing education.	Systematic Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	3	1	1	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Yildiz M et al [23]. Hemşirelerde tiger temelli hemşirelik bilişimi yetkinlikleri.	Descriptive-correlational study	JBI Critical Appraisal Checklist for Analytical Cross-Sectional Studies	1	1	1	1	3	2	1	1	—	—	—	1
Canadian Nurses Association [24]. Nursing informatics.	Policy Brief	JBI Critical Appraisal Checklist for Textual Evidence: Policy	1	1	1	3	1	2	1	—	—	—	—	1
Ramos Rodríguez JM [25]. Las TICs en enfermería de práctica avanzada.	Narrative Article	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	—	—	—	—	—	1
González-Pardo Maza E [26]. Tecnología Big Data y su misión en el campo de la enfermería.	Narrative Review	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	—	—	—	—	—	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Reid L et al [27]. Nursing informatics: competency challenges for nursing faculty.	Scoping review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	3	1	1	1
AMIA[f] Policy Invitational [28]. Redefining our picture of health.	White Paper	JBI Critical Appraisal Checklist for Textual Evidence: Policy	1	1	1	1	1	1	1	—	—	—	—	1
Toffoletto MC and Ahumada Tello JD [29]. Telenursing in care, education, and management.	Integrative Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	3	3	1	3	1	1	1
Kaihlanen A et al [30]. Nursing informatics competence profiles and perceptions.	Cross-sectional study	JBI Critical Appraisal Checklist for Analytical Cross Sectional Studies	1	1	1	1	1	1	1	1	—	—	—	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Forman TM et al [31]. A review of clinical informatics competencies in nursing.	Systematic Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	3	3	1	3	1	1	1
Kleib M et al [32]. Approaches for defining and assessing nursing informatics competencies.	Scoping Review (protocol)	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	3	1	2	2	2	1	1
Kleib M et al [33]. Approaches for defining and assessing nursing informatics competencies.	Scoping Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	1	1	1	1
Davies A et al [34]. Core competencies for clinical informaticians.	Systematic Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	1	1	1	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Domingos CS et al [35]. A aplicação do processo de enfermagem informatizado.	Narrative Review	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	1	—	—	—	—	—	1
Nahm ES et al [36]. Cybersecurity essentials for nursing informaticists.	Narrative Review	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	3	1	3	3	1	1	1
Taylor-Pearson K et al [37]. The role of nurse informaticians in advancing 3D printing use in health care.	Systematic Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	3	1	1	1
Australia's Digital Health Community [5]. Leadership in clinical informatics.	White Paper	JBI Critical Appraisal Checklist for Textual Evidence: Policy	1	1	1	1	1	3	1	—	—	—	—	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Brown J et al [38]. Issues affecting nurses' capability to use digital technology.	Integrative Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	1	1	1	1
Gonzalo de Diego B et al [39]. Competencies in the robotics of care for nursing robotics.	Scoping Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	3	1	2	3	2	1	1
Nazeha N et al [40]. A digitally competent health workforce.	Scoping review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	1	1	1	1
Le Y et al [41]. A bibliometric and visualized analysis of nursing informatics competencies in China.	Bibliometric Review	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	3	3	1	1	—	—	—	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
von Gerich H et al [42]. Artificial intelligence-based technologies in nursing.	Scoping review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	3	1	1	1
Harerimana A et al [43]. Nursing informatics in undergraduate nursing education.	Scoping Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	3	1	1	1	1	1	1
Wynn M et al [44]. Digital nursing practice theory.	Scoping review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	3	1	1	3	1	1	1
Oh J et al [45]. Evaluation of the effects of flipped learning of a nursing informatics course.	Quasi-experimental study	JBI Critical Appraisal Checklist for Quasi-Experimental Studies	1	1	1	1	1	1	3	1	1	3	2	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Lokmic-Tomkins Z et al [46]. Perspectives on the implementation of health informatics curricula frameworks.	Qualitative study	JBI Critical Appraisal Checklist for Qualitative Research	1	1	1	1	3	2	1	1	—	—	—	1
Haupeltshofer A et al [47]. Promoting health literacy.	Scoping review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	3	1	1	1
Lee KH et al [48]. Empowering healthcare through comprehensive informatics education.	Narrative Review	JBI Critical Appraisal Checklist for Textual Evidence: Narrative	1	1	1	1	1	3	3	1	3	1	1	1

XSL•FO
RenderX

Main author [reference] and title	Article type	Critical appraisal tools	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Include
Stunden A et al [49]. Nursing students' preparedness for the digitalized clinical environment.	Integrative Review	JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses	1	1	1	1	1	1	1	1	1	1	1	1

[a]JBI: Joanna Briggs Institute.

[b]1: yes (valid).

[c]3: No (not valid).

[d]Not applicable.

[e]2: unclear.

[f]AMIA: American Medical Informatics Association.

Data Extraction and Analysis

This review was conducted through manual data extraction by the researchers to maximize the quality of information selection. The search strategy was supplemented by reviewing gray literature, obtaining documents from relevant associations, and white papers on NI around the world. Some were deemed eligible, and others were excluded after a screening process. The data were then processed in an ordered format with the help of the Microsoft Office suite. This helped create a much more sensible and accurate starting point for the analysis using that software. This method likely ensures a high-quality review of the literature and, in turn, provides a rigorous summary of the evidence.

Artificial intelligence was used on a very limited basis to make only editorial corrections within the final manuscript in order to enhance writing quality and correct grammar, spelling, and structure in different article files. These tools were not used to extract or analyze results. This approach ensures the relevance and currency of the extracted data within the study context.

Phase 2

This phase aimed to validate the scoping review findings through a gap analysis, consisting of a cross-sectional expert panel consultation survey with quantitative items and related qualitative open-ended questions (this survey or questionnaire can be found in Multimedia Appendix 1). For this purpose, a panel of 10 national experts in NI was assembled. All participants volunteered to participate and were recruited through an open call issued by the authors. Experts were selected based on the following criteria: (1) 10 or more years of experience related to health and digital literacy environments in Spain, (2) extensive training in digital literacy, and (3) peer-recognized expertise within their professional field. Their sociodemographic characteristics are detailed in Table 2.

XSL•FO
RenderX

Table . Experts' sociodemographic characteristics.

ID	Gender	Age (y)	Years of experience	Highest education level	Current position	Main area of expertise	Autonomous community
E1	Male	46	22	PhD	University professor	Digital health literacy	Madrid
E2	Male	53	29	PhD	Director of Digital Strategy	Hospital information systems	Catalonia
E3	Female	49	25	Master's	Clinical Informatics Specialist	Clinical workflows and EHR[a]	Andalusia
E4	Female	57	31	PhD	Senior eHealth Researcher	Telemedicine and patient care	Madrid
E5	Female	40	16	Master's	Informatics Unit Supervisor	EHR implementation	Valencian Community
E6	Male	32	15	PhD	Digital Health Consultant	Data governance and security	Valencian Community
E7	Female	43	19	PhD	Health care Innovation Manager	Semantic interoperability	Galicia
E8	Female	56	33	PhD	Dean of Health Sciences	Ethics and legislation in eHealth	Castile and León
E9	Male	45	20	Master's	Advanced Practice Nurse	Big data and health analytics	Valencian Community
E10	Female	53	27	PhD	Director of Nursing Services	Leadership and digital change	Andalusia

[a]EHR: electronic health record.

Sample size was determined based on the principles of purposive sampling and data saturation. This approach is consistent with the methodological literature on Delphi studies and expert panels, where a sample of 10 to 20 highly specialized participants is considered sufficient to achieve a comprehensive depth of information and reach thematic saturation [50]. Qualitative comments from the experts were analyzed through a simplified thematic analysis to identify relevant information not captured by quantitative items. The quality and richness of the data provided by these experts were prioritized over a larger, less specialized sample. To ensure objectivity, anonymity among experts was maintained so that none of the experts knew the identity of the rest of the panel.

Ethical Considerations

The scoping review conducted in this study exclusively used publicly available data and literature. Therefore, it did not require ethical approval. For the gap analysis, all participating experts were informed of the study's objectives and provided written consent to participate before answering the questionnaire. Anonymity and confidentiality of their responses were guaranteed throughout the process. No compensation was provided to the participants.

Trial Registration

The results of a health care intervention on human subjects are not reported in this study; therefore, registration in a clinical trials registry was not applicable.

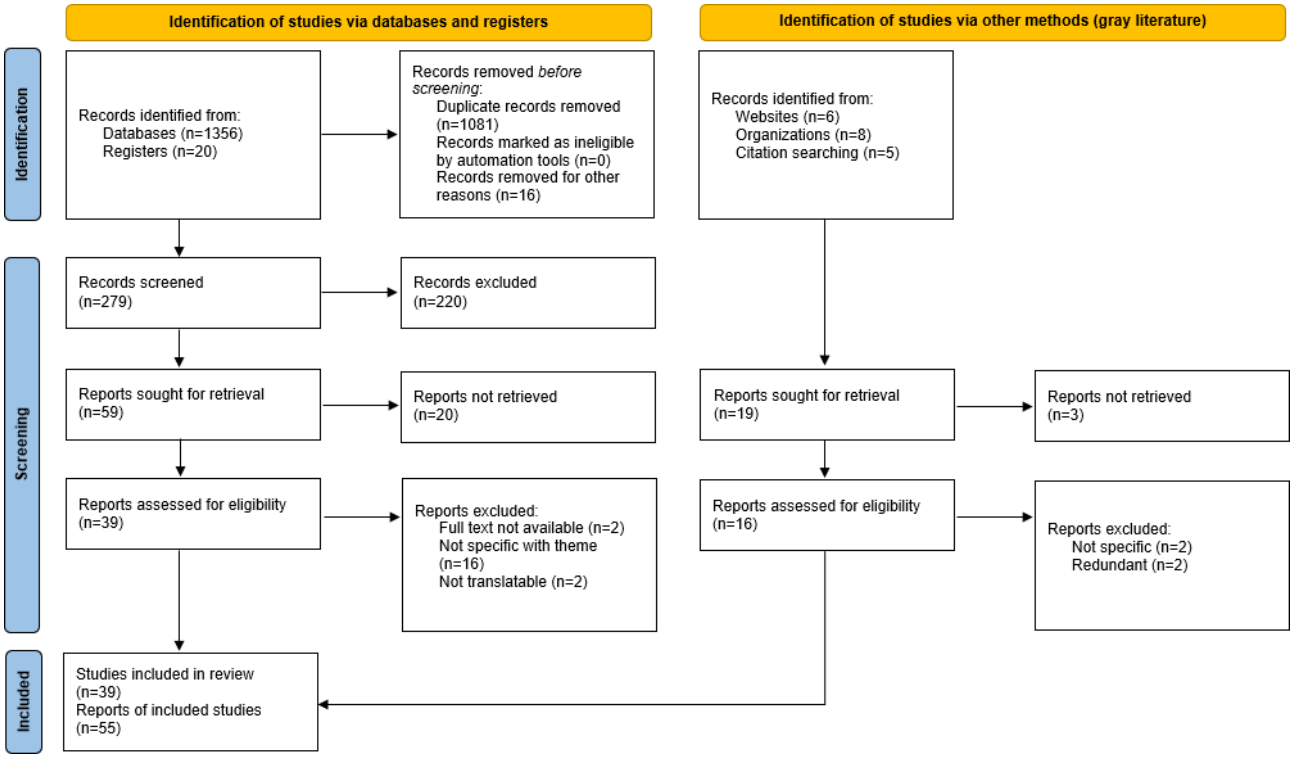
Results

Scientific Literature Review

The scoping review included 55 published studies (the full list is available in Multimedia Appendix 2). The literature review identified an increasing scientific output internationally on NI and the role of nursing informaticians (see the PRISMA [Preferred Reporting Items for Systematic Reviews and Meta-Analyses] flow diagram [Figure 1] for the process of the complete search [16]). However, production in Spanish is significantly lower.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



Academic Training

During the study and after an exhaustive literature analysis, several training methods were identified that are used to impart the necessary knowledge and skills to nursing informaticians in clinical practice (Table 3).

Table . Academic training.

Academic training	Description	References
Undergraduate and postgraduate programs	Degrees incorporating health information systems, data analysis, and emerging technologies.	[1,17-19]
Research projects and theses	Advanced training (Master's/PhD) focused on investigating specific NI[a] problems to develop analytical skills.	[18]
Practical training and rotations	Supervised clinical placements allowing the application of theoretical knowledge in real-world settings.	[2,17,19]
Simulations and labs	Use of EHR[b] simulators and informatics laboratories for safe, interactive learning.	[17,26,27]
Continuing education	Professional certifications, seminars, and workshops for updating skills (eg, cybersecurity, data management).	[2,20,23,24,29]
Leadership programs	Mentorship initiatives for emerging leaders (eg, Alliance for Nursing Informatics Emerging Leaders Program).	[2,25]

[a]NI: nursing informatics.

[b]EHR: electronic health record.

In summary, training in NI is conducted through a combination of academic programs, competency models, continuing education, clinical placements, simulations, online courses, seminars, workshops, research projects, and leadership training. These training methods must be developed to increase knowledge about NI [30] and evaluated to ensure that nurses acquire the necessary competencies to effectively implement and use information systems in clinical practice, as well as to integrate new technologies into the clinical environment [31].

Competencies

As a result of the consulted literature and its exhaustive analysis, essential competencies for nursing informaticians have been identified [32,33]. These competencies are fundamental to

XSL•FO
RenderX

understanding the functions they would fulfill in this nursing leadership role within the Spanish health care system. They are structured into 6 general competencies (Table 4).

Table . Competencies.

Competencies	Description	References
Information management	Includes the collection, storage, organization, and analysis of health data. Nursing informaticians must ensure patient privacy (applying current regulations) regarding their private information at all times. It also implies establishing standardized nursing documentation models for efficient and accurate information (NI[a] process), as well as the use of standardized taxonomies. Measures nursing care complexity and ensures privacy.	[35,36,51,52]
Cybersecurity and patient safety	The integration of new technologies must prioritize patient safety, minimizing risks. In turn, the accuracy and reliability of data must be guaranteed, as well as their confidentiality.	[36,37]
Evaluation and development of clinical information systems	Analysis and management, creation, evaluation, and modification of hospital information systems to improve the quality of care. They must be practical and relevant to nursing practice. They must also integrate connectivity with new technologies.	[36]
Leadership and coordination of digital tools	Leadership in digital communication and patient care management using advanced digital tools are important competencies.	[5,14,36]
Implementation of new technologies and specialized applications	Integration of artificial intelligence, 3D printing, metaverse, big data, and robotics. Nursing informaticians as technological experts integrating and promoting these technologies.	[37-40]
Education and digitalization in health	Continuing education in new technologies is crucial to prepare health professionals. Nursing informaticians must be trained to transmit knowledge in information technologies and the use of emerging technologies to health care professionals in their environment.	[41,42]

[a]NI: nursing informatics.

Contextual Framework

The successful implementation of nursing informaticians requires the definition, training, and aptitude of nurses in digital competencies within a specific contextual framework. This framework must include both generalist nurses and nursing informaticians [17,38,47].

All nurses must have the opportunity to receive training and qualifications as expert users to fully leverage the possibilities offered by technology while also knowing how to mitigate information overload [5,38]. Generally, younger nurses tend to possess more positive training, education, and attitudes toward technology; however, all professionals must be provided with opportunities adapted to their existing knowledge to undergo training in digital transformation [17,38]. Strategies must be established to combat computer illiteracy, structural problems (such as lack of staff and financial resources), and limited access to new technologies, which constitute the most common barriers. Likewise, these conditions must be met while ensuring the privacy and security of obtained data, servers, and devices to guarantee the safety of both patients and professionals [28,38].

Specific leadership roles in digital transformation will be assumed by nursing informaticians [5,17]. Their objective must be to increase system efficiency and improve the quality of care through evidence-based data in an interdisciplinary, coordinated, and patient-centered manner [4,5,17,24].

Benefits

Based on the literature, the implementation of the nursing informaticians' role in Spain presents several potential benefits, inspired by the experiences of other countries (Table 5).

XSL•FO
RenderX

Table . Benefits.

Benefits	Description	References
Improved management and use of health information systems	Nursing informaticians combine clinical nursing knowledge with an understanding of information sciences, acting as a connection between nurses and technical teams	[1,3,4,23,28,32,33,47]
Leadership in the adoption of new technologies	Responsible for developing and implementing ICTs[a] in the clinical setting	[3,8,14,22,24,38,47]
Optimization of the development and implementation of the electronic health record	Nursing informaticians ensure that EHR[b] systems are designed to facilitate compatibility with care and traceability of clinical practice	[3,18,30,38,42]
Contribution to nursing research	Participating in the research and implementation of digitalization solutions, as well as the collection, processing, and interpretation of large datasets	[6,14,15,28,29,40,46]
Effective application of the NP[c], care complexity, and resource allocation	Strengthening evidence-based practice and consolidating nursing as a science. Nursing informaticians can ensure the effective application of the NP and promote the use of standardized taxonomies. Furthermore, the use of structured data allows quantifying care complexity. Consequently, nursing informaticians become essential for designing systems that move from operational documentation to predictive decision-making	[8,19,23,35,42,51-53]

[a]ICT: information and communication technology.

[b]EHR: electronic health record.

[c]NP: nursing process.

Barriers

Potential barriers to its implementation have been identified (Table 6).

Table . Barriers.

Barriers	Description	References
Lack of formal recognition and specific training programs in Spain.	Currently, training in NI[a] is limited.	[13,19,25,32,33,35,42,48]
Need to improve digital literacy in nursing.	Current competencies in applied informatics are insufficient.	[8,13,19,26,35,38,42]
Challenges in the integration of technologies and systems.	Includes a plurality of terminologies and a lack of experience among nurses.	[6,43,49]
Absence of a national digital health competency framework specific to nursing.	Essential for defining and acquiring the competencies of nursing informaticians.	[6,17,43]
Resistance to change and a lack of understanding of the value of structured data.	On the part of some professionals.	[17,26,42]
Limitations in accessing and using clinical nursing data.	Depend on electronic records.	[2,6,13,42,43]
Need for investment and resources.	For the training and integration of these professionals.	[5,6,14,20,38,46]

[a]NI: nursing informatics.

Expert-Validated Gap Analysis

Following the scoping review, the opinions of a panel of 10 experts in the field of NI were gathered. Using a form, they evaluated the results retrieved from the scoping review based on their extensive experience. They rated items on a scale from 1 (strongly disagree) to 5 (strongly agree) and wrote down their perspectives on each aspect from the scoping review.

Academic Training

The applicability and desirability of implementing specific training programs for nursing informaticians were

XSL•FO
RenderX

overwhelmingly rated very positively by the experts (4.5/5). The proposals for their implementation are diverse:

- Continuing education: the importance of "continuing education from professional associations" is emphasized, including "practical workshops" and accessible courses. A "hybrid training model, in small teams, by knowledge level and very real needs applied to the specific practice of the work area" is proposed.
- Methodology: the development of a "modular and progressive curriculum, with training in digital competencies from the degree level, interdisciplinary master's programs, and accredited continuing education, integrating mixed methodologies (in-person and online), digital clinical simulations, and practicums in real health care settings" is valued.
- Collaboration: the need for universities, in collaboration with professional associations and technology companies, to work in symbiosis to develop the most relevant content is emphasized.
- Implementation phases: some experts suggest "starting with pilot experiences in hospitals that are already digitized" or "initially through postgraduate training, possibly due to low demand," to later integrate it into the undergraduate degree.
- Culture and recognition: the importance of "creating the culture and the social or professional symbolic image of this role before its future implementation" and the need for a "prior change in the culture of digital literacy among nursing professionals" are highlighted.

Competencies

The competencies described in the scoping review received the highest rating (4.7/5), along with the benefits. The experts consider that the competencies identified for nursing informaticians are "highly relevant and applicable within the Spanish health care system, as they directly respond to current and emerging needs":

- Key benefits: it is highlighted that they are crucial for "patient safety" and the "improvement in quality of care, potential for research, and analysis of results."
- Adaptation to the environment: it is acknowledged that they must be "adapted to our reality: limited resources, differences between autonomous communities, and the need for more practical training."
- Transdisciplinary role: it is suggested that nursing informaticians "should be integrated into transdisciplinary teams in research, as well as in management or decision-making regarding the design, development, validation, and implementation of new technologies."
- European trend: it is pointed out that "they correlate with basic European digital competencies, therefore showing traceability with our educational system."

Benefits

There was broad recognition among experts (4.7/5) of the multiple benefits that implementing the EESI role would bring:

- Improved quality and efficiency: an "optimization in the management and use of health information systems" is expected, improving "care continuity, interoperability […] and the traceability of care."
- Leadership in digital transformation: EESIs would provide "key leadership in the adoption and implementation of information and communication technologies." One expert believes that "nursing is the most qualified profession to enhance its training in digital competencies and lead technological projects."
- Boost to research: "nursing research would be promoted through the analysis of clinical data, allowing evidence to be generated from practice."
- Strengthening the profession: the EESI role would help to "revalue the nursing profession as a key agent in the digital transformation of health." The aim is for nurses to lead technological change, "(that we) do not just follow it."
- Patient benefits: "care plans and health education, monitoring of care plans," and the promotion of "active patient participation in their care process" are mentioned.

Barriers

The experts' ratings in this section were the most varied (4.4/5). Although the majority continue to identify significant obstacles, one expert is optimistic, noting that "all the classically identified barriers are being progressively addressed, which creates a more favorable and realistic scenario for the effective implementation of nursing informaticians' "role," while the others refer to these previously identified barriers:

- Lack of recognition and regulatory framework: the "absence of official recognition of the role" and the lack of a "regulatory framework and defined policies" are crucial.
- Training and digital literacy: the "scarce specific training at the national level (and a) low level of digital literacy in the nursing community" are mentioned. One expert indicates that "the main barrier is the nursing professionals themselves, due to a lack of time and a culture of digital competencies."
- Resistance to change and organizational culture: "resistance to change from health care professionals accustomed to more traditional work models," and the "lack of clinical leadership from nurses who are experts in information systems" are relevant.
- Resources and infrastructure: "scarce economic, technological, and human resources" and "difficulties in systems integration and terminological standardization" are impediments.

Main Findings

This study reveals a significant dichotomy between the internationally recognized role of the nursing informaticians and their unrecognized role within the Spanish health care system. Through a scoping review, the international evidence has been synthesized, delineating a clear framework of competencies, training models, and benefits associated with this specialized role. Second, the expert-validated gap analysis confirmed that this framework is not only relevant because of its benefits but also essential for the Spanish context nowadays. Barriers that impede its implementation were also identified. The consensus among experts (4.7/5 for benefits and

XSL•FO
RenderX

competencies) underscores a professional and academic mandate to bridge this gap.

Findings articulate 6 core competencies for the nursing informaticians: information management, cybersecurity and patient safety, evaluation and development of clinical information systems, leadership and coordination of digital tools, implementation of new technologies and specialized applications, and education and digitalization in health. These are more than technical skills, as they represent functions such as empowering nurses to act as the central professionals between clinical practice, information technology, and organizational strategy [3,29]. The strong validation by Spanish experts (4.7/5) confirms their applicability and designates a direction for future educational and professional development programs regarding this role.

Discussion

Comparison With Spanish Legal Framework

The status of NI in Spain has proven to be distanced from that of countries such as the United States, Canada, and Australia, where professional associations like the American Medical Informatics Association and the Healthcare Information and Management Systems Society have led the integration of nursing informaticians as those responsible for driving digital transformation [2,8,12,23,25]. A critical analysis of the Spanish legal framework reveals the structural nature of this gap. While international nursing informaticians possess advanced competencies in systems design, cybersecurity, and data leadership, the current regulation of the nursing profession in Spain—specifically Order CIN/2134/2008, which regulates the nursing degree [54], and Royal Decree 954/2015 on nursing prescription [55]—limits digital competencies to the user level.

Both regulations solely and explicitly define the digital competency of nurses in Spain as "applying technologies and information systems" [54,55]. This creates a dichotomy: the legislation exclusively restricts nurses to using existing tools, whereas the health care system increasingly demands qualified specialists to evaluate, design, and lead the implementation of new tools. This restriction within the legal framework illustrates why Spanish nursing staff often feel insufficiently prepared for the digitized clinical environment, echoing findings from Australia [13,49].

Frameworks for Transformation and Cultural Change

A proposed conceptual framework as a driver for change could be the Technology Informatics Guiding Education Reform, with the intent of moving from basic computer literacy toward information management and innovation leadership. The progression of these competencies would reflect the Data-Information-Knowledge-Wisdom model, in which nursing informaticians in Spain would guide digital transformation by using health information and resources to create solutions adapted to the Spanish context.

Overcoming these barriers requires a deliberate and structured strategy, similar to that already achieved by other international pioneers. The emphasis placed by the experts consulted in this study on the "lack of a digital literacy culture" and "resistance from the professionals themselves" points to the need for a cultural shift. Technical training must be accompanied by an evolution of the organizational culture that values data-driven care and nursing leadership in technology.

Implications for Practice, Policy, and Future Research

The findings of this study have significant implications: hospital and health care managers can use the validated competency framework from this study to formally introduce the role requirements. Providing nurses with specialized training and empowering them to lead improvement projects related to information systems or the implementation of new digital tools could potentially create momentum and demonstrate the value of the role so that they are not only technical but also strategic for making nursing care visible [17,26,42,51].

Results suggest a call to action for policymakers and regulatory bodies in Spain, such as the Ministry of Health and National Agency for Quality Assessment and Accreditation, to integrate nursing informaticians' competencies into nursing curricula. This could begin with undergraduate introductions and extend to interdisciplinary master's programs, as suggested by the experts. Collaboration among academia, professional associations, and technology companies is essential to ensure that this training is relevant, practical, and aligned with the needs of the health care system.

Longitudinal studies are needed to evaluate the impact of pilot training programs on nursing competencies. Further research should also focus on developing and validating a standardized instrument to assess NI competencies specifically tailored to the Spanish health care context.

Limitations

This study has several limitations that should be acknowledged. First, the search was limited to studies published in English, Spanish, and Portuguese, which may have excluded relevant literature in other languages. The manuscript has been carefully reviewed by a native English speaker, although some phrasing or constructs may still reflect the linguistic influences of the researchers' native language. Second, the search of gray literature may not have captured all relevant institutional documents or unpublished curricula. Third, the expert panel was selected through purposive sampling to ensure richness of data. Although thematic saturation was reached, the sample was limited to 10 participants. Furthermore, a potential selection bias exists due to the voluntary nature of the recruitment (open call); the experts may possess a baseline "pro-technology" attitude, which could have influenced the highly positive ratings compared to what might be observed in a generalist nursing population. Fourth, the protocol for this scoping review was not prospectively registered in a public repository such as PROSPERO (International Prospective Register of Systematic Reviews). However, the study was conducted following a detailed internal protocol developed by the authors following the PRISMA-ScR and Joanna Briggs Institute guidelines. Finally, a potential publication bias in the literature was spotted, where successful implementations of NI are more likely to be reported than failures. This could skew the international perspective. Future research should address the operational and

XSL•FO
RenderX

budgetary requirements needed to change the organizational governance model for implementing the nursing informaticians' role.

Conclusions

This 2-phase study provides a comprehensive and critical diagnosis of the situation of NI and their role in Spain. It revealed a gap between the nation's current reality and the established international standard. The lack of a conceptual framework, specialized university training, and official recognition for the nursing informaticians' role are the principal barriers preventing nursing informaticians from leading the digital transformation of health care.

The international literature findings were validated by national experts, creating an actionable roadmap. The identified competencies offer a robust foundation for designing educational programs. The successful training models from other countries provide a blueprint for their implementation. The experts' high level of agreement was unequivocal: the nursing informaticians' role is imperative for improving patient safety, optimizing system efficiency, and furthering the development of NI as a scientific discipline in Spain.

Challenges are significant, although the pathway forward is clear. This research serves as a foundational call to action for policymakers, academic institutions, and health care leaders: building regulatory, educational, and organizational infrastructure is needed to empower a new generation of nursing leaders equipped to lead the future of digital health in Spain.

Acknowledgments

The authors would like to acknowledge the experts who participated in the gap analysis for their valuable time and insights. This study is part of the doctoral thesis of Adrián Marco-Moyano, led by Dr Manuel Lillo-Crespo, within the research line "Innovation and Improvement in Health Care" (University of Alicante). The authors used Gemini exclusively for English language editing and proofreading. No artificial intelligence tools were used to create content, data, or references. The authors created, drafted, and verified all aspects of this manuscript and take full accountability for its accuracy and integrity.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The data extracted and analyzed during the scoping review are available from the corresponding author upon reasonable request. The individual responses from the experts are not publicly available to protect participant privacy.

Authors' Contributions

AMM conceived the study and its design. MLC and AMM conducted the literature search, screening, and data extraction. PVR participated in the data extraction and adequacy due to her technological background. MLC and AMM developed the questionnaire and the gap analysis with the experts. All authors drafted the manuscript, critically revised it for intellectual content, and approved the final version for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Gap analysis questionnaire.
[PDF File, 209 KB - nursing_v9i1e83373_app1.pdf]

Multimedia Appendix 2
Scoping review chart 1.
[DOCX File, 1723 KB - nursing_v9i1e83373_app2.docx]

Checklist 1
PRISMA checklist.
[PDF File, 509 KB - nursing_v9i1e83373_app3.pdf]

References

1. Reid L, Maeder A, Button D, Breaden K, Brommeyer M. Defining nursing informatics: a narrative review. Stud Health Technol Inform 2021 Dec 15;284:108-112. [doi: 10.3233/SHTI210680] [Medline: 34920485]

XSL•FO
RenderX

2. Sensmeier J. The value and impact of the alliance for nursing informatics emerging leaders program. Comput Inform Nurs 2019 Dec;37(12):612-614. [doi: 10.1097/CIN.0000000000000603] [Medline: 31809306]
3. Kirchner RB. Introducing nursing informatics. Nursing (Auckl) 2014 Sep;44(9):22-23. [doi: 10.1097/01.NURSE.0000453006.79653.33] [Medline: 25140937]
4. Luppa N, Suresh S. Physician and nurse informatics collaboration boosts clinical practice, engagement and overall digital experience. : Healthcare Information and Management Systems Society (HIMSS); 2022 URL: https://keystone.himss.org/sites/hde/files/media/file/2022/12/01/himss_nurseinfocollaboration_wp.pdf
5. Leadership in clinical informatics. : HISA; 2018 URL: https://digitalhealth.org.au/wp-content/uploads/2020/01/HISA-_Leadership-Clinical-Informatics_FINAL.pdf [accessed 2026-03-18]
6. Benavente-Rubio A. The role of nursing in digital health: opportunities and challenges for the science of caring. Rev Med Clin Condes 2022;33(6):598-603. [doi: 10.1016/j.rmclc.2022.11.004]
7. Konstantinidis ST, Poultourtzidis I, Papamalis F, et al. HEALINT4ALL digital interactive platform for european and national placements audit for medicine and allied health professions following a user-centered design. In: Auer ME, Tsiatsos T, editors. : Springer; 2023 Presented at: International Conference on Interactive Mobile Communication, Technologies and Learning (IMCL 2023); Nov 9-10, 2023; Thessaloniki, Greece p. 208-215. [doi: 10.1007/978-3-031-54327-2_21]
8. Papamalis FE, Cunningham S, Hodges P, et al. ENSURING quality healthcare practice for doctors and medical allied professionals through a digital interactive audit platform. 2023 Presented at: 17th International Technology, Education and Development Conference; Mar 6-8, 2023; Valencia, Spain p. 5825-5831. [doi: 10.21125/inted.2023.1529]
9. Baggia A, Leskovar R, Rajkovič U, et al. Selecting the e-learning platform for nursing education. Presented at: 38 mednarodna konferenca o razvoju organizacijskih znanosti; Mar 20-22, 2019; Portorož p. 49-63. [doi: 10.18690/978-961-286-250-3.5]
10. Konstantinidis S, Tilley Z, Ko S, et al. CO-creation of a virtual interactive teaching package for auditors of healthcare placements – towards assurance of quality of health care traineeships. Presented at: 15th International Conference on Education and New Learning Technologies; Jul 3-5, 2023; Palma, Spain p. 3189-3195. [doi: 10.21125/edulearn.2023.0886]
11. Ball MJ, Hannah KJ, Cortes-Comerer N, Douglas JV. Editorial: the health informatics series: evolving with a new discipline. Int J Med Inform 2023 May;173:105008. [doi: 10.1016/j.ijmedinf.2023.105008] [Medline: 36868101]
12. León Molina J. Role of nursing in the hospital information system [Article in Spanish]. Enferm Glob 2004;3(2):1-8 [FREE Full text] [doi: 10.6018/eglobal.3.1.587]
13. Chipps J, Le Roux L, Agabus J, Bimerew M. Nursing informatics skills relevance and competence for final year nursing students. Curationis 2022 Nov 21;45(1):e1-e8. [doi: 10.4102/curationis.v45i1.2277] [Medline: 36453814]
14. Backonja U, Langford LH, Mook PJ. How to support the nursing informatics leadership pipeline. Comput Inform Nurs 2022 Jan;40(1):8-20. [doi: 10.1097/CIN.0000000000000827] [Medline: 34996883]
15. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
16. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
17. Conte G, Arrigoni C, Magon A, Stievano A, Caruso R. Embracing digital and technological solutions in nursing: a scoping review and conceptual framework. Int J Med Inform 2023 Sep;177:105148. [doi: 10.1016/j.ijmedinf.2023.105148] [Medline: 37453178]
18. Farzandipour M, Mohamadian H, Akbari H, Safari S, Sharif R. Designing a national model for assessment of nursing informatics competency. BMC Med Inform Decis Mak 2021 Feb 2;21(1):35. [doi: 10.1186/s12911-021-01405-0] [Medline: 33531023]
19. Farokhzadian J, Jouparinejad S, Montazeri M, Bakhshipour E, Sabzi A, Falahati-Marvast F. Necessary prerequisites for evidence-based practice: results of investigating nurses' informatics competency and information literacy skills. Front Nurs 2024 Mar 1;11(1):67-76. [doi: 10.2478/fon-2024-0007]
20. Kulju E, Jarva E, Oikarinen A, Hammarén M, Kanste O, Mikkonen K. Educational interventions and their effects on healthcare professionals' digital competence development: a systematic review. Int J Med Inform 2024 May;185:105396. [doi: 10.1016/j.ijmedinf.2024.105396] [Medline: 38503251]
21. Saco P. Informática aplicada a la enfermería [Web page in Spanish]. Universidade de Santiago de Compostela. 2021. URL: https://www.usc.gal/es/estudios/grados/ciencias-salud/grado-enfermeria/20202021/informatica-aplicada-enfermeria-15970-15393-11-80407 [accessed 2025-07-05]
22. Nes AAG, Steindal SA, Larsen MH, Heer HC, Lærum-Onsager E, Gjevjon ER. Technological literacy in nursing education: a scoping review. J Prof Nurs 2021;37(2):320-334. [doi: 10.1016/j.profnurs.2021.01.008] [Medline: 33867086]
23. Yıldız M, Yıldız M, Yıldırım MS, et al. Hemşirelerde tiger temelli hemşirelik bilişimi yetkinlikleri ile dijital okuryazarlık düzeyleri arasındaki i˙lişkinin i˙ncelenmesi: Makine öğrenmesi yaklaşımı ile analizi [Article in Turkish]. Fenerbahçe Univ J Health Sci 2023;3(3):372-385. [doi: 10.56061/fbujohs.1252026]
24. Nursing informatics: joint position statement. : Canadian Nurses Association & Canadian Nursing Informatics Association; 2012 URL: https://cnia.ca/resources/Documents/Nursing-informatics-joint-position-statement.pdf [accessed 2026-03-18]
25. Ramos Rodríguez JM. Las tics en enfermería de práctica avanzada [Report in Spanish]. : Universidad de Cádiz, RODIN Repository; 2017 URL: https://rodin.uca.es/bitstream/handle/10498/19361/

XSL•FO
RenderX

Las%20TICs%20en%20Enfermer%c3%ada%20de%20Pr%c3%a1ctica%20Avanzada.pdf?sequence=1&isAllowed=y [accessed 2026-03-18]
26. González-Pardo Maza E. Tecnología big data y su misión en el campo de la enfermería [Web page in Spanish]. Repositorio Institucional de la Universidad de Cantabria. 2021. URL: https://repositorio.unican.es/xmlui/handle/10902/22498 [accessed 2026-07-05]
27. Reid L, Button D, Breaden K, Brommeyer M. Nursing informatics: competency challenges for nursing faculty. Stud Health Technol Inform 2024 Jan 25;310(41-45):1196-1200. [doi: 10.3233/SHTI231154] [Medline: 38270004]
28. Redefining our picture of health: towards a person-centered integrated care, research, wellness and community ecosystem. : AMIA; 2017 URL: https://brand.amia.org/m/4b538b840ece6a23/original/Redefining-Our-Picture-of-Health.pdf [accessed 2026-03-18]
29. Toffoletto MC, Tello JDA. Telenursing in care, education and management in Latin America and the Caribbean: an integrative review. Rev Bras Enferm 2020;73(Suppl 5):e20190317. [doi: 10.1590/0034-7167-2019-0317] [Medline: 33027492]
30. Kaihlanen AM, Elovainio M, Virtanen L, et al. Nursing informatics competence profiles and perceptions of health information system usefulness among registered nurses: a latent profile analysis. J Adv Nurs 2023 Oct;79(10):4022-4033. [doi: 10.1111/jan.15718] [Medline: 37243421]
31. Forman TM, Armor DA, Miller AS. A review of clinical informatics competencies in nursing to inform best practices in education and nurse faculty development. Nurs Educ Perspect 2020;41(1):E3-E7. [doi: 10.1097/01.NEP.0000000000000588] [Medline: 31860501]
32. Kleib M, Chauvette A, Furlong KE, Nagle LM, Slater L, McCloskey R. Approaches for defining and assessing nursing informatics competencies: a scoping review protocol. JBI Database System Rev Implement Rep 2019 Jun;17(6):1071-1078. [doi: 10.11124/JBISRIR-2017-003889] [Medline: 31192897]
33. Kleib M, Chauvette A, Furlong K, Nagle L, Slater L, McCloskey R. Approaches for defining and assessing nursing informatics competencies: a scoping review. JBI Evid Synth 2021 Apr;19(4):794-841. [doi: 10.11124/JBIES-20-00100] [Medline: 33625068]
34. Davies A, Mueller J, Moulton G. Core competencies for clinical informaticians: a systematic review. Int J Med Inform 2020 Sep;141:104237. [doi: 10.1016/j.ijmedinf.2020.104237] [Medline: 32771960]
35. Domingos CS, Boscarol GT, Brinati LM, Dias AC, Souza CC, Salgado PDO. The application of the computerized nursing process: an integrative review [Article in Spanish]. Enferm Glob 2017;16(4):603-652. [doi: 10.6018/eglobal.16.4.278061]
36. Nahm ES, Poe S, Lacey D, Lardner M, Van De Castle B, Powell K. Cybersecurity essentials for nursing informaticists. Comput Inform Nurs 2019 Aug;37(8):389-393. [doi: 10.1097/CIN.0000000000000570] [Medline: 31393328]
37. Taylor-Pearson K, Pearson D, Hoelscher SH. The role of nurse informaticians in advancing 3D printing use in health care: a promising frontier for improved patient outcomes. J Inform Nurs 2024;9(1):6. [doi: 10.63123/JIN.2024.9.1.6]
38. Brown J, Pope N, Bosco AM, Mason J, Morgan A. Issues affecting nurses' capability to use digital technology at work: an integrative review. J Clin Nurs 2020 Aug;29(15-16):2801-2819. [doi: 10.1111/jocn.15321] [Medline: 32416029]
39. Gonzalo de Diego B, González Aguña A, Fernández Batalla M, et al. Competencies in the robotics of care for nursing robotics: a scoping review. Healthcare (Basel) 2024 Mar 8;12(6):617. [doi: 10.3390/healthcare12060617] [Medline: 38540581]
40. Nazeha N, Pavagadhi D, Kyaw BM, Car J, Jimenez G, Tudor Car L. A digitally competent health workforce: scoping review of educational frameworks. J Med Internet Res 2020 Nov 5;22(11):e22706. [doi: 10.2196/22706] [Medline: 33151152]
41. Le Y, Cao S, Wang M, Lin X, Qian B. A bibliometric and visualized analysis of nursing informatics competencies in China (2000-2020). Stud Health Technol Inform 2021 Dec 15;284(147-151):403-407. [doi: 10.3233/SHTI210759] [Medline: 34920559]
42. von Gerich H, Moen H, Block LJ, et al. Artificial Intelligence -based technologies in nursing: a scoping literature review of the evidence. Int J Nurs Stud 2022 Mar;127:104153. [doi: 10.1016/j.ijnurstu.2021.104153] [Medline: 35092870]
43. Harerimana A, Wicking K, Biedermann N, Yates K. Nursing informatics in undergraduate nursing education in Australia before COVID-19: a scoping review. Collegian 2022 Aug;29(4):527-539. [doi: 10.1016/j.colegn.2021.11.004] [Medline: 34867065]
44. Wynn M, Garwood-Cross L, Vasilica C, Davis D. Digital nursing practice theory: a scoping review and thematic analysis. J Adv Nurs 2023 Nov;79(11):4137-4148. [doi: 10.1111/jan.15660] [Medline: 36971282]
45. Oh J, Kim SJ, Kim S, Vasuki R. Evaluation of the effects of flipped learning of a nursing informatics course. J Nurs Educ 2017 Aug 1;56(8):477-483. [doi: 10.3928/01484834-20170712-06] [Medline: 28787070]
46. Lokmic-Tomkins Z, Raghunathan K, Almond H, et al. Perspectives on the implementation of health informatics curricula frameworks. Contemp Nurse 2024 Apr;60(2):178-191. [doi: 10.1080/10376178.2024.2343010] [Medline: 38662767]
47. Haupeltshofer A, Egerer V, Seeling S. Promoting health literacy: what potential does nursing informatics offer to support older adults in the use of technology? A scoping review. Health Informatics J 2020 Dec;26(4):2707-2721. [doi: 10.1177/1460458220933417] [Medline: 32627639]

XSL•FO
RenderX

48. Lee KH, Kim MG, Lee JH, et al. Empowering healthcare through comprehensive informatics education: the status and future of biomedical and health informatics education. Healthc Inform Res 2024 Apr;30(2):113-126. [doi: 10.4258/hir.2024.30.2.113] [Medline: 38755102]
49. Stunden A, Ginige A, O'Reilly R, Sanagavarapu P, Heaton L, Jefferies D. Nursing students' preparedness for the digitalised clinical environment in Australia: an integrative review. Nurse Educ Pract 2024 Feb;75:103908. [doi: 10.1016/j.nepr.2024.103908] [Medline: 38335697]
50. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment in data saturation and variability. Field methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
51. Cesare M, D' Agostino F, Damiani G, et al. Exploring the impact of medical complexity on nursing complexity of care in paediatric patients: a retrospective observational study. J Clin Nurs 2025 Jul;34(7):2748-2765. [doi: 10.1111/jocn.17460] [Medline: 39381892]
52. Cesare M, Cocchieri A. Can an increase in nursing care complexity raise the risk of intra-hospital and intensive care unit transfers in children? A retrospective observational study. J Pediatr Nurs 2025;80:91-99. [doi: 10.1016/j.pedn.2024.11.015] [Medline: 39602875]
53. Cesare M, D'Agostino F, Cocchieri A. Exploring the association between complexity of care, medical complexity, and length of stay in the paediatric setting using a nursing minimum data set: a study protocol. Nurs Rep 2024 Oct 10;14(4):2923-2934. [doi: 10.3390/nursrep14040213] [Medline: 39449450]
54. Ministerio de Ciencia e Innovación. Orden CIN/2134/2008, de 3 de julio, por la que se establecen los requisitos para la verificación de los títulos universitarios oficiales que habiliten para el ejercicio de la profesión de enfermero [Article in Spanish]. Boletín Oficial del Estado 2008 Jul 19;174:31680-31683 [FREE Full text]
55. Ministerio de Sanidad, Consumo y Bienestar Social (Spain). Real Decreto 1302/2018, de 22 de octubre, por el que se modifica el real decreto 954/2015, de 23 de octubre, por el que se regula la indicación, uso y autorización de dispensación de medicamentos y productos sanitarios de uso humano por parte de los enfermeros. Boletín Oficial del Estado 2018 Oct 23;256:102609-102615 [FREE Full text]

Abbreviations

EESI: Enfermeras Especialistas en Sistemas de Información
NI: nursing informatics
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
PROSPERO: International Prospective Register of Systematic Reviews

Edited by E Borycki; submitted 01.Sep.2025; peer-reviewed by M Kleib, M Cesare, P Lacroix, S Yokota; revised version received 08.Jan.2026; accepted 03.Feb.2026; published 27.Apr.2026.

Please cite as:
Marco-Moyano A, Verdú Rodríguez P, Lillo-Crespo M
Nursing Informaticians in Spain: Scoping Review and Expert-Validated Gap Analysis
JMIR Nursing 2026;9:e83373
URL: https://nursing.jmir.org/2026/1/e83373
doi:10.2196/83373

© Adrián Marco-Moyano, Patricia Verdú Rodríguez, Manuel Lillo-Crespo. Originally published in JMIR Nursing (https://nursing.jmir.org), 27.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Shift From Efficacy to Implementation Science (2020-2026) in Nursing Practice for Digital Mental Health: Scoping Review and Bibliometric Analysis

Shoukai Yu[1,2], Dr; Lingmei Qian[2,3]; Hao Wang[1,2], DMed

[1]School of Medicine, Shanghai Jiao Tong University, 720 Xianxia Road, Room 104, Shanghai, Shanghai, China

[2]Hongqiao International Institute of Medicine, Shanghai Tongren Hospital and Clinical Research Institute, Shanghai Jiao Tong University School of Medicine, Shanghai, Shanghai, China

[3]Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, Shanghai, China

Corresponding Author:
Shoukai Yu, Dr
School of Medicine, Shanghai Jiao Tong University, 720 Xianxia Road, Room 104, Shanghai, Shanghai, China

Abstract

Background: Nurses are pivotal as end users and implementers of digital mental health interventions (DMHIs). However, the successful translation of efficacious DMHIs into sustainable nursing practice is hindered by multifaceted implementation challenges.

Objective: This study aimed to systematically map and analyze the evolving research landscape of nurse-involved DMHIs to determine if a paradigm shift from efficacy testing to implementation science is occurring and characterize the methodological and thematic trends associated with this shift.

Methods: We conducted a scoping review integrated with a bibliometric analysis following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines. A comprehensive search of the Web of Science Core Collection, Scopus, and PubMed databases (January 2020 to April 2026) yielded 1014 eligible primary research studies. Studies were classified as “efficacy/effectiveness” or “implementation” research using a validated framework. We analyzed temporal trends and methodological designs and performed co-word thematic mapping.

Results: Implementation research constituted the largest proportion of clearly classified studies (327/470, 69.6%), exceeding efficacy research (143/470, 30.4%). Implementation research demonstrated a significant growth trajectory from 2020 to 2025 (β=8.63 articles per year; *P*=.003), surpassing efficacy research in 2022. Implementation studies used more mixed methods (105/327, 32.1% vs 7/143, 4.9%; *P*<.001) and qualitative designs (92/327, 28.1% vs 2/143, 1.4%; *P*<.001). Thematic analysis revealed “nursing workflow integration” and “nurse-led implementation” as emerging motor themes.

Conclusions: This review provides robust evidence of an accelerating turn to implementation in digital mental health research in nursing. The field is moving beyond internal validity toward real-world implementation. Future efforts should focus on implementation-ready digital tools and embedding implementation science competencies into nursing education.

(JMIR Nursing 2026;9:e91498) doi:10.2196/91498

KEYWORDS

digital mental health; mobile health; mHealth; eHealth; implementation science; nursing informatics; bibliometric analysis; scoping review; workflow integration; mobile phone

Introduction

The integration of digital mental health interventions (DMHIs) into routine nursing practice represents both an opportunity and a challenge for contemporary health care systems [1-3]. Nurses, as the largest group of health care professionals worldwide, are uniquely positioned to serve as frontline implementers, coordinators, and sustainers of DMHIs across diverse clinical settings [4,5]. However, despite growing evidence of DMHI efficacy from over a decade of randomized controlled trials [6,7], their real-world implementation in nursing practice remains limited, creating what has been termed the “implementation-to-practice gap” in digital mental health [8,9].

Recent systematic reviews have consistently identified nurse-specific barriers to DMHI adoption, including workflow disruption, inadequate training, and concerns about increased workload [10,11]. A 2020 mixed methods study by Bourla et al [12] found that although over 70% of student nurses reported willingness to adopt digital mental health technologies in clinical practice, more than 60% voiced worries that such tools could damage the therapeutic relationship, indicating that the readiness

XSL•FO
RenderX

of future nursing practitioners to implement DMHIs remains insufficient. Concurrently, the global nursing shortage, exacerbated by the COVID-19 pandemic, has intensified interest in digital solutions that can extend the reach of mental health services [13,14].

Implementation science offers conceptual frameworks and methodological approaches specifically designed to address these implementation challenges [15,16]. Frameworks such as the reach, effectiveness, adoption, implementation, and maintenance framework [17] and the Consolidated Framework for Implementation Research [18] shift the research focus from questions of efficacy ("Does it work under ideal conditions?") to questions of implementation ("How can nurses successfully integrate this into their workflow?"). While calls for greater attention to implementation science in nursing digital health research have increased [19,20], the extent to which the published literature reflects this methodological paradigm shift remains empirically unexamined.

To inform the development of more implementable digital tools and effective support strategies for nurses, this scoping review and bibliometric analysis systematically mapped the published literature from 2020 to 2026. We sought to determine whether nurse-involved DMHI research is indeed undergoing this critical implementation turn. Our specific objectives were to (1) quantify the prevalence of implementation-focused research vs efficacy research, (2) analyze temporal trends in this paradigm adoption, (3) compare the methodological approaches characterizing each paradigm, and (4) identify the key thematic foci and their evolution within the nursing DMHI literature. A conservative keyword-based classification strategy was used to prioritize specificity, providing a lower-bound estimate of implementation research prevalence.

Methods

Study Design

This study was conducted as a scoping review incorporating a bibliometric analysis in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) statement [21]. The review protocol was not preregistered. Our primary objective was to systematically map and characterize the published literature on nurse-involved DMHIs from 2020 to 2026. The completed PRISMA-ScR checklist is provided in Checklist 1.

Data Source and Search Strategy

Data were retrieved from 3 electronic databases on April 15, 2026: Web of Science Core Collection, Scopus, and PubMed. The search was limited to peer-reviewed journal articles published in English between January 1, 2020, and April 15, 2026. The search strategy combined terms related to three conceptual domains: (1) digital interventions [22,23], (2) mental health conditions [3], and (3) nursing involvement and research paradigms [4,8]. The complete search strategy for each database is provided in Multimedia Appendix 1.

Inclusion and Exclusion Criteria

Articles were included if they met all of the following criteria: (1) focused primarily on a digital intervention targeting mental health conditions; (2) involved nursing roles (nurse-led, nurse-delivered, or nurse-involved implementation); (3) reported original empirical research with human participants; (4) were published in English in peer-reviewed journals; and (5) were published between January 1, 2020, and April 15, 2026. We excluded review articles, study protocols, commentaries, editorials, theoretical papers, and studies in which nursing involvement was peripheral.

Article Classification and Validation

Publications were classified using a multistage process informed by implementation science frameworks [18,24]. Implementation research was defined by a focus on real-world adoption and implementation outcomes in nursing contexts [25], and efficacy and effectiveness research was defined by a focus on clinical outcomes under controlled conditions. A conservative keyword-based classification strategy was used to prioritize specificity, providing a lower-bound estimate of implementation research prevalence. Classification followed 3 stages: automated keyword screening, manual verification, and validation against a consensus standard (κ=0.73, 95% CI 0.60-0.86).

Data Extraction and Quality Assessment

Two reviewers independently extracted data from each included study using a standardized form. Extracted items included first author, publication year, country, study design, sample size, research paradigm (implementation vs efficacy or effectiveness), theoretical framework used, implementation outcomes measured (if applicable), and key themes. Disagreements were resolved through discussion or consultation with a third reviewer. Quality assessment was not performed as this was a scoping review focused on mapping the literature rather than critically appraising evidence quality [21].

Statistical and Bibliometric Analysis

Descriptive statistics summarized publication characteristics. Proportions with 95% CIs were calculated using the Wilson score method. Temporal trends were analyzed using linear regression. Methodological differences between research paradigms were examined using chi-square tests. Co-word analysis and thematic mapping were conducted using the *bibliometrix* package in R (version 4.2; R Foundation for Statistical Computing) [26].

Ethical Considerations

This study involved analysis of publicly available published literature and did not require ethics approval.

Results

Study Selection and Characteristics

Our search identified 1737 records across 3 databases (Scopus: n=1183, 68.1%; Web of Science: n=431, 24.8%; PubMed: n=123, 7.1%). To ensure a conservative estimate of unique publications, we performed deduplication using both digital object identifier (DOI) and title. Of the 1737 identified records,

XSL•FO
RenderX

DOI-based deduplication removed 337 (19.4%) duplicates (resulting in n=1400, 80.6% of records), and subsequent title-based deduplication identified an additional 386 (22.2%) duplicate records with inconsistent or missing DOI assignments across databases. This 2-stage approach yielded 1014 unique records (Figure 1). All 1014 records met the inclusion criteria for original empirical research on nurse-involved DMHIs published in English between 2020 and 2026.

The final sample included studies from 44 countries. The largest contributions came from the United States (373/1014, 36.8%), the United Kingdom (161/1014, 15.9%), Australia (110/1014, 10.8%), and China (83/1014, 8.2%). The mean sample size of primary studies was 162 (SD 241) participants.

Across all 1014 included studies, randomized controlled trials represented the most common study design (n=303, 29.9%). Observational designs followed as the second most frequent category (n=129, 12.7%), followed by mixed methods research (n=106, 10.5%). Feasibility and pilot studies and qualitative studies each comprised 7.1% (n=72) of the studies, whereas process evaluation designs were the least prevalent (n=44, 4.3%).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. DOI: digital object identifier.

XSL•FO
RenderX

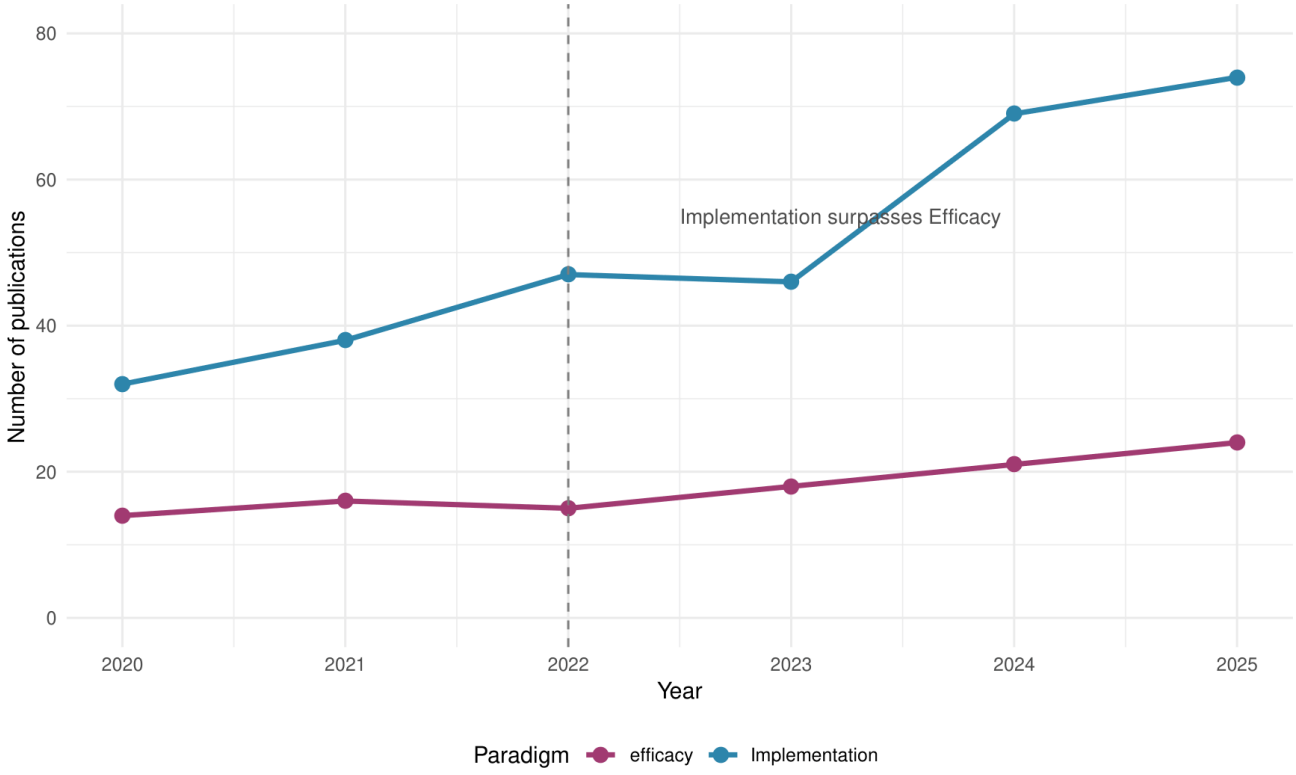
Prevalence of Research Paradigms

For research paradigm classification, a conservative keyword coding approach left 53.6% (544/1014) of the total studies with an ambiguous categorization that could not be clearly assigned to either implementation- or efficacy-focused work. Among the remaining 470 clearly classified studies, implementation-focused research constituted the dominant paradigm at 69.6% (n=327), whereas efficacy and effectiveness research made up the remaining 30.4% (n=143). All mixed or hybrid methodological studies were subsumed within the implementation research grouping in line with standard implementation science classification frameworks [27]. The implementation-to-efficacy study ratio was calculated at 2.29:1, demonstrating a preponderance of implementation research within the clearly classified sample.

Temporal Trends in Paradigm Adoption

Publication volume demonstrated substantial growth from 2020 to 2025 (Figure 2), increasing from 102 articles in 2020 to 247 in 2025 (142% increase). Linear regression analysis revealed a significant positive trend in implementation research from 2020 to 2025 (β=8.63 articles per year, 95% CI 5.21-12.05; *P*=.003; R^2=0.91). Implementation research studies outnumbered efficacy research studies in 2022 (47 implementation-related articles vs 15 efficacy or effectiveness articles) and retained this predominance in subsequent years, peaking at 74 articles in 2025. Efficacy and effectiveness research demonstrated no statistically significant growth across the study period (β=1.8 articles per year; *P*=.48). The lower count in 2026 (21 implementation articles) reflects incomplete year data and was excluded from trend analysis.

Figure 2. Annual publication trends by research paradigm (2020-2025). Implementation research demonstrated a significant growth trajectory, increasing from 32 articles in 2020 to 74 in 2025, surpassing efficacy research in 2022 (47 vs 15). Linear regression analysis (2020-2025) revealed a β value of 8.63 articles per year (95% CI 5.21-12.05; *P*=.003; R^2=0.91). Efficacy and effectiveness research showed no significant growth (β=1.8; *P*=.48). The year 2026 was excluded due to incomplete data.



Methodological Divergence Between Paradigms

Analysis of studies with unambiguous classification revealed substantial methodological divergence (Table 1 and Figure 3). Implementation studies used significantly more mixed methods designs (105/327, 32.1% vs 7/143, 4.9%; χ^2_5=32.4; *P*<.001) and qualitative approaches (92/327, 28.1% vs 2/143, 1.4%; χ^2_5=41.2; *P*<.001) compared to efficacy and effectiveness studies. In contrast, efficacy research remained predominantly reliant on randomized controlled trials compared to implementation studies (97/143, 67.8% vs 82/327, 25.1%, respectively; χ^2_5=24.8; *P*<.001).

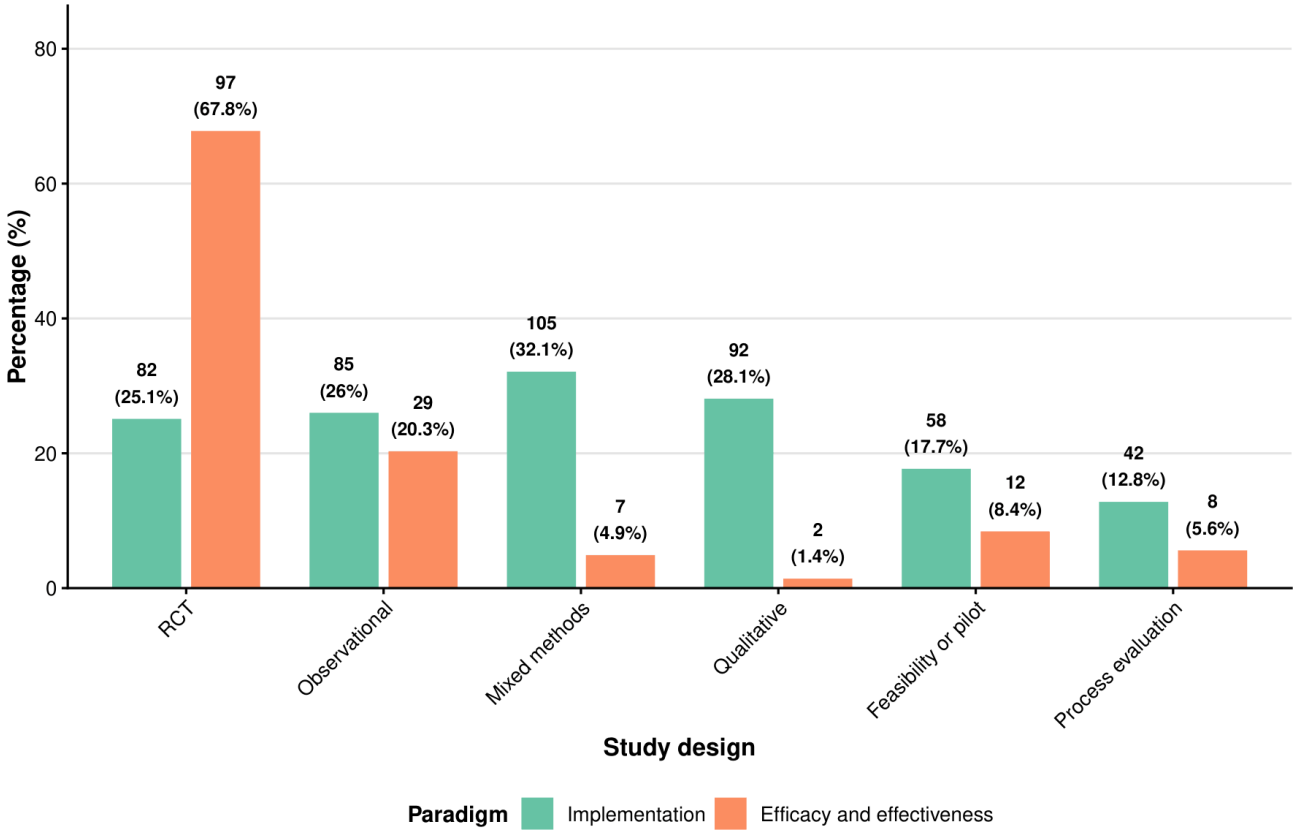
XSL•FO
RenderX

Table . Methodological approaches across research paradigms (n=470 studies with unambiguous classification).[a]

Study design	Implementation research (n=327), n (%)	Efficacy and effectiveness research (n=143), n (%)	Chi-square (*df*)	*P* value
RCT[b]	82 (25.1)	97 (67.8)	24.8 (5)	<.001
Observational	85 (26.0)	29 (20.3)	1.8 (5)	.41
Mixed methods	105 (32.1)	7 (4.9)	32.4 (5)	<.001
Qualitative	92 (28.1)	2 (1.4)	41.2 (5)	<.001
Feasibility or pilot	58 (17.7)	12 (8.4)	2.3 (5)	.32
Process evaluation	42 (12.8)	8 (5.6)	3.9 (5)	.14

[a]Percentages exceed 100% because studies may use multiple designs. Chi-square tests compare implementation vs efficacy and effectiveness research.
[b]RCT: randomized controlled trial.

Figure 3. Methodological approaches across research paradigms (n=470 studies with unambiguous classification). Implementation research used significantly more mixed methods (105/327, 32.1% vs 7/143, 4.9%; *P*<.001) and qualitative designs (92/327, 28.1% vs 2/143, 1.4%; *P*<.001) compared to efficacy and effectiveness research, which remained predominantly reliant on randomized controlled trials compared to implementation studies (97/143, 67.8% vs 82/327, 25.1%, respectively; *P*<.001). RCT: randomized controlled trial.

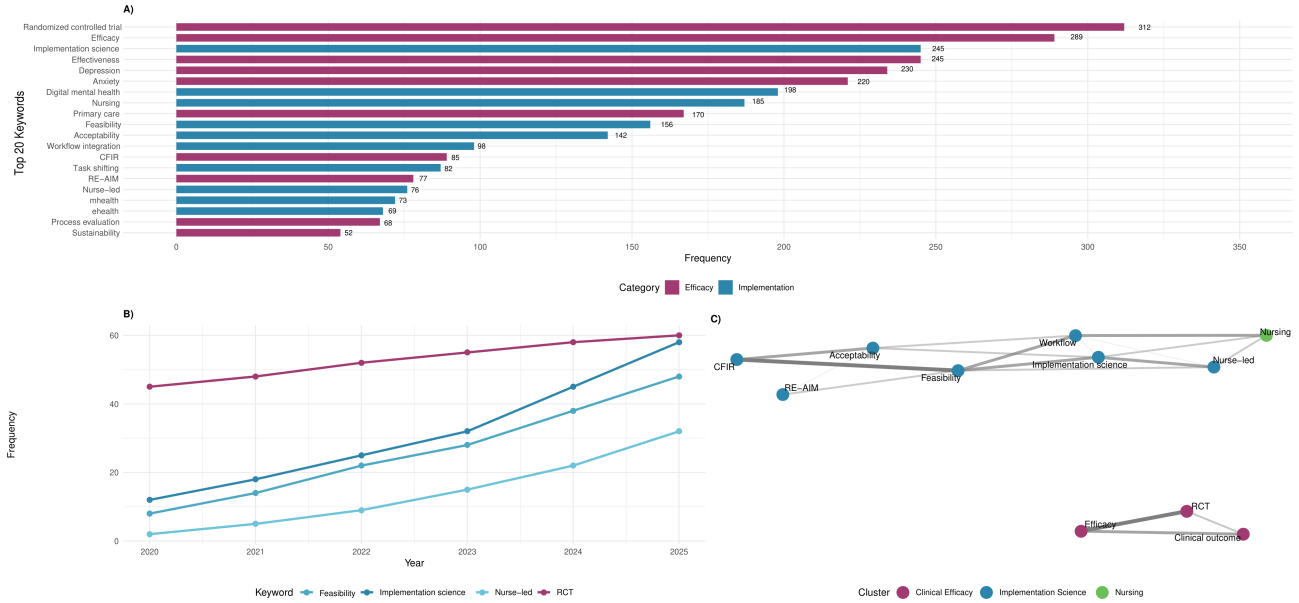


Thematic Evolution and Nursing-Specific Patterns

Co-word analysis of keywords from all 1014 publications revealed a clear thematic evolution toward implementation science (Figure 4). Implementation-focused constructs including "implementation science," "nursing workflow," "task-shifting," and "nurse-led implementation" emerged as motor themes characterized by high centrality and density. Nursing-specific implementation barriers such as "time constraints," "training needs," and "workflow integration" showed increasing prominence in recent publications (2024 - 2025). Notably, implementation-related keywords showed the highest annual growth rate among all thematic categories (β=2.1 appearances per year; *P*=.008).

XSL•FO
RenderX

Figure 4. Thematic evolution in nurse-involved digital mental health research (N=1014). (A) Bar chart of the top 20 most frequently occurring keywords across all the included studies. Implementation-related keywords (blue) dominate the top ranks, with "randomized controlled trial" (n=312) and "efficacy" (n=289) being the most frequent efficacy-related terms. (B) Annual trends of the top 10 keywords (2020-2025), demonstrating increasing prominence of implementation-related terms ("implementation science," "feasibility," and "nurse-led"), whereas efficacy-related terms ("RCT") remained stable. The year 2026 was excluded due to incomplete data. (C) Keyword co-occurrence network showing three distinct clusters: (1) implementation science (blue; centrality=0.85), (2) clinical efficacy research (purple; centrality=0.32), and (3) nursing workflow integration (green; centrality=0.68). Implementation science emerged as a motor theme with high centrality and density.



Discussion

Principal Findings

Our bibliometric analysis of 1014 nurse-involved DMHI studies provides robust empirical evidence of a decisive paradigm shift toward implementation science. Three convergent lines of evidence support this conclusion. First, implementation research constituted the largest proportion of clearly classified studies (327/470, 69.6%), exceeding traditional efficacy research (143/470, 30.4%). Second, implementation research demonstrated a significant annual growth rate of 8.63 articles per year from 2020 to 2025 (P=.003; R^2=0.91), surpassing efficacy research in 2022. Third, fundamental methodological differences characterize these paradigms, with implementation studies using substantially more mixed methods (105/327, 32.1% vs 7/143, 4.9%) and qualitative designs (92/327, 28.1% vs 2/143, 1.4%).

Comparison With Existing Literature

Our findings align with broader trends in implementation science but extend them to the specific context of nursing digital health [28,29]. Previous research has documented increasing attention to implementation science in health care generally [30], but our study is the first to quantify this shift specifically in nurse-involved DMHI research using a multi-database bibliometric approach. The methodological divergence we observed—with implementation studies favoring mixed methods approaches—reflects the complex, context-dependent nature of implementation challenges in nursing practice [27].

Recent studies have begun to identify effective nurse-specific implementation strategies. Targeted, blended training combining online learning and in-person coaching markedly elevates nurses' adoption of digital mental health tools, with intervention groups demonstrating more than double the uptake relative to untrained control staff [31]. Similarly, research confirms that "implementation champions" drive long-term sustainability; grassroots nurse-led digital interventions maintain stable clinical use far longer than institution-mandated top-down rollouts [32].

Implications for Nursing Practice

The documented shift has direct implications for creating more nurse-friendly digital health ecosystems.

First, for digital health developers, our findings underscore the necessity of "implementation by design." DMHI development must move beyond user experience for patients to include clinician experience for nurses, explicitly considering workflow integration, time burden, and interoperability with clinical systems from the earliest stages [22,23].

Second, for nursing informatics leadership, chief nursing informatics officers and other leaders must advocate for and allocate resources toward implementation infrastructure. This includes dedicated time for nurse training and adaptation, the creation of role-specific implementation playbooks, and the evaluation of DMHIs based on implementation outcomes (eg, adoption, feasibility, and sustainability) alongside clinical outcomes [33].

Third, for education and training, nursing curricula and continuing professional development must evolve to build digital health implementation competency. This includes skills in evaluating digital tools for implementability, leading practice change, and using data from DMHIs to inform care—a core component of modern nursing informatics literacy [34,35].

XSL•FO
RenderX

Future Directions for Nursing Research

Our findings suggest several directions for future nursing research. First, there is a need for more hybrid effectiveness-implementation designs that simultaneously examine clinical outcomes and implementation processes [27]. Second, research should explore nurse-specific implementation strategies such as tailored training programs and workflow integration tools [36]. Third, future studies should examine how implementation outcomes vary across nursing roles and settings [37]. Finally, given the geographic concentration of research in high-income countries, there is an urgent need for implementation research in low- and middle-income settings, where nurses often serve as the primary mental health care providers [38].

Strengths and Limitations

Strengths of this study include its comprehensive multi-database search strategy (Web of Science, Scopus, and PubMed), rigorous classification system with validation (κ=0.73), updated search up to April 2026, and integration of multiple bibliometric techniques, analyzing 1014 eligible studies.

Several limitations warrant acknowledgment. First, reliance on title and abstract text for initial classification may have misclassified some studies, although manual verification and consensus validation mitigated this risk. Our conservative classification strategy prioritized specificity, resulting in a higher proportion of unclear classifications (544/1014, 53.6%); this provides a lower-bound estimate of implementation research prevalence. Second, the exclusion of non–English-language publications introduces language bias, potentially missing relevant research from non–English-speaking countries. Third, we did not include CINAHL due to institutional access limitations, which may have underrepresented nursing-specific literature. Future reviews should include this database. Fourth, the rapid evolution of digital health technologies means that findings from older studies (2020-2022) may not fully reflect current capabilities, although our temporal trend analysis addresses changing patterns over time. Fifth, the lack of quality assessment, consistent with scoping review methodology, means that we cannot comment on the methodological rigor of the included studies.

Conclusions

This scoping review and bibliometric analysis documents a decisive and accelerating paradigm shift toward implementation science in nurse-involved DMHI research. Implementation research now constitutes the dominant paradigm (327/1014, 32.2% vs 143/1014, 14.1% for efficacy research) and demonstrated a significant annual growth rate of 8.63 articles per year from 2020 to 2025 (P=.003; R^2=0.91), surpassing efficacy research in 2022. The field is expanding beyond a primary focus on internal validity to systematically investigate the complex determinants and processes of real-world implementation in nursing contexts.

Acknowledgments

The authors thank the librarians at Shanghai Jiao Tong University for assistance with database searching.

SY conceived the study, drafted the manuscript and collated table data. SY, LQ, and HW completed all statistical and bibliometric analyses. All authors revised the manuscript and approved the final version.

Funding

This work was supported by the Shanghai Municipal Human Resources Bureau and Shanghai Science and Technology Commission (Pujiang Talents Program; grant 19PJC085 to SY).

Data Availability

The data that support the findings of this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Complete search strategy.
[DOCX File, 24 KB - nursing_v9i1e91498_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 27 KB - nursing_v9i1e91498_app2.docx]

References

1. Naslund JA, Bondre A, Torous J, Aschbrenner KA. Social media and mental health: benefits, risks, and opportunities for research and practice. J Technol Behav Sci 2020 Sep;5(3):245-257. [doi: 10.1007/s41347-020-00134-x] [Medline: 33415185]

XSL•FO
RenderX

2. Torous J, Andersson G, Bertagnoli A, et al. Towards a consensus around standards for smartphone apps and digital mental health. World Psychiatry 2019 Feb;18(1):97-98. [doi: 10.1002/wps.20592] [Medline: 30600619]
3. Lattie EG, Stiles-Shields C, Graham AK. An overview of and recommendations for more accessible digital mental health services. Nat Rev Psychol 2022 Feb;1(2):87-100. [doi: 10.1038/s44159-021-00003-1] [Medline: 38515434]
4. Janes G, Chesterton L, Heaslip V, et al. Current nursing and midwifery contribution to leading digital health policy and practice: an integrative review. J Adv Nurs 2025 Jan;81(1):116-139. [doi: 10.1111/jan.16265] [Medline: 38946119]
5. Booth RG, Strudwick G, McBride S, O'Connor S, Solano López AL. How the nursing profession should adapt for a digital future. BMJ 2021;373:n1190. [doi: 10.1136/bmj.n1190]
6. Andrews G, Basu A, Cuijpers P, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78. [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
7. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health-a systematic review and meta-analysis. NPJ Digit Med 2019;2:118. [doi: 10.1038/s41746-019-0188-8] [Medline: 31815193]
8. Graham AK, Lattie EG, Powell BJ, et al. Implementation strategies for digital mental health interventions in health care settings. Am Psychol 2020 Nov;75(8):1080-1092. [doi: 10.1037/amp0000686] [Medline: 33252946]
9. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of eHealth interventions: systematic review of the literature. J Med Internet Res 2018 May 1;20(5):e10235. [doi: 10.2196/10235] [Medline: 29716883]
10. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387. [doi: 10.2196/24387] [Medline: 33759801]
11. Li Y, Ji W, Chen H, Xie X, Yang J, Gao J. Psychometric properties of instruments used to measure the informatics competence of nurses: a systematic review. Nurse Educ Pract 2024 Aug;79:104070. [doi: 10.1016/j.nepr.2024.104070] [Medline: 39003784]
12. Bourla A, Mouchabac S, Ogorzelec L, Guinchard C, Ferreri F. Are student nurses ready for new technologies in mental health? Mixed-methods study. Nurse Educ Today 2020 Jan;84:104240. [doi: 10.1016/j.nedt.2019.104240] [Medline: 31715472]
13. State of the world's nursing 2020: investing in education, jobs and leadership. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/9789240003279 [accessed 2026-06-10]
14. Drissi N, Ouhbi S, Marques G, de la Torre Díez I, Ghogho M, Janati Idrissi MA. A systematic literature review on e-mental health solutions to assist health care workers during COVID-19. Telemed J E Health 2021 Jun;27(6):594-602. [doi: 10.1089/tmj.2020.0287] [Medline: 32970532]
15. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. BMC Psychol 2015 Sep 16;3(1):32. [doi: 10.1186/s40359-015-0089-9] [Medline: 26376626]
16. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10:53. [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]
17. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
18. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 7;4:50. [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
19. Al-Moteri M, Aljuaid J. Development of an implementation science higher diploma for registered nurses: phase III of the EQUIP initiative. Nurs Health Sci 2025 Mar;27(1):e70062. [doi: 10.1111/nhs.70062] [Medline: 39928966]
20. Fontaine G, Vinette B, Weight C, et al. Effects of implementation strategies on nursing practice and patient outcomes: a comprehensive systematic review and meta-analysis. Implement Sci 2024 Sep 30;19(1):68. [doi: 10.1186/s13012-024-01398-0] [Medline: 39350295]
21. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
22. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
23. Mair FS, May C, O'Donnell C, Finch T, Sullivan F, Murray E. Factors that promote or inhibit the implementation of e-health systems: an explanatory systematic review. Bull World Health Organ 2012 May 1;90(5):357-364. [doi: 10.2471/BLT.11.099424] [Medline: 22589569]
24. Kirk MA, Kelley C, Yankey N, Birken SA, Abadie B, Damschroder L. A systematic review of the use of the Consolidated Framework for Implementation Research. Implement Sci 2016 May 17;11:72. [doi: 10.1186/s13012-016-0437-z] [Medline: 27189233]
25. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]

XSL•FO
RenderX

26. Aria M, Cuccurullo C. bibliometrix: an R-tool for comprehensive science mapping analysis. J Informetr 2017 Nov;11(4):959-975. [doi: 10.1016/j.joi.2017.08.007]
27. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226. [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
28. Jha R, Sachdeva G. Digital mapping of health-related quality of life research in non-communicable diseases: a topic modeling and bibliometric synthesis for policy and practice. Inform Health Soc Care 2026 Jan 7;51(2):1-21. [doi: 10.1080/17538157.2025.2610690] [Medline: 41503941]
29. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. FASEB J 2008 Feb;22(2):338-342. [doi: 10.1096/fj.07-9492LSF] [Medline: 17884971]
30. Palinkas LA, Aarons GA, Horwitz S, Chamberlain P, Hurlburt M, Landsverk J. Mixed method designs in implementation research. Adm Policy Ment Health 2011 Jan;38(1):44-53. [doi: 10.1007/s10488-010-0314-z] [Medline: 20967495]
31. Boß L, Ross J, Reis D, et al. Effectiveness of an integrated platform-based intervention for promoting psychosocial safety climate and mental health in nursing staff: a pragmatic cluster randomised controlled trial. Int J Nurs Stud 2025 Jul;167:105076. [doi: 10.1016/j.ijnurstu.2025.105076] [Medline: 40215602]
32. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161. [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
33. Laakkonen N, Jarva E, Hammarén M, et al. Digital competence among healthcare leaders: a mixed-methods systematic review. J Nurs Manag 2024;2024:8435248. [doi: 10.1155/2024/8435248] [Medline: 40224897]
34. Kleib M, Nagle LM, Furlong KE, Paul P, Duarte Wisnesky U, Ali S. Are future nurses ready for digital health?: informatics competency baseline assessment. Nurse Educ 2022;47(5):E98-E104. [doi: 10.1097/NNE.0000000000001199] [Medline: 35324499]
35. Belchez CA, Kunkel D, Kulhanek B, Tietze MF, Pinekenstein B. Integrating digital health and informatics competencies in nursing curricula: a framework for nurse educators. Nurse Educ 2026;51(3):158-162. [doi: 10.1097/NNE.0000000000002117] [Medline: 41557522]
36. Powell BJ, Waltz TJ, Chinman MJ, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci 2015 Feb 12;10:21. [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
37. Lynch EA, Mudge A, Knowles S, Kitson AL, Hunter SC, Harvey G. "There is nothing so practical as a good theory": a pragmatic guide for selecting theoretical approaches for implementation projects. BMC Health Serv Res 2018 Nov 14;18(1):857. [doi: 10.1186/s12913-018-3671-z] [Medline: 30428882]
38. Ibeneme S, Karamagi H, Muneene D, Goswami K, Chisaka N, Okeibunor J. Strengthening health systems using innovative digital health technologies in Africa. Front Digit Health 2022;4:854339. [doi: 10.3389/fdgth.2022.854339] [Medline: 35434700]

Abbreviations

DMHI: digital mental health intervention
DOI: digital object identifier
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by UM Kinnunen; submitted 15.Jan.2026; peer-reviewed by R Gil-Gutiérrez, X Wu; revised version received 31.May.2026; accepted 03.Jun.2026; published 08.Jul.2026.

Please cite as:
Yu S, Qian L, Wang H
The Shift From Efficacy to Implementation Science (2020-2026) in Nursing Practice for Digital Mental Health: Scoping Review and Bibliometric Analysis
JMIR Nursing 2026;9:e91498
URL: https://nursing.jmir.org/2026/1/e91498
doi:10.2196/91498

© Shoukai Yu, Lingmei Qian, Hao Wang. Originally published in JMIR Nursing (https://nursing.jmir.org), 8.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Artificial Intelligence–Based Nursing Interventions for Chronic Illness Care: Umbrella Review

Jee Young Joo[1], RN, PhD; Megan Liu[2], RN, PhD; Youngwoo Cho[1], RN, BSN; Hyungbin Cho[1]

[1]College of Nursing, Gachon University, 191 Hambangmoe-ro, Yeonsu-gu, Medical Science Building, Room 902, Incheon, Republic of Korea

[2]School of Gerontology and Long-Term Care, College of Nursing, Taipei Medical University, Taipei, Taiwan

Corresponding Author:
Jee Young Joo, RN, PhD
College of Nursing, Gachon University, 191 Hambangmoe-ro, Yeonsu-gu, Medical Science Building, Room 902, Incheon, Republic of Korea

Abstract

Background: Artificial intelligence (AI)–based nursing interventions are increasingly being used to manage chronic illnesses; however, their definitive impact on clinical outcomes remains inconclusive, necessitating a comprehensive evidence synthesis.

Objective: This umbrella review aimed to synthesize the evidence regarding the effectiveness of AI-based nursing interventions for chronic illness care and their subsequent impact on health care outcomes in clinical settings.

Methods: We conducted an umbrella review and prospectively registered the protocol. A systematic search of 5 electronic databases (PubMed, CINAHL, Cochrane Library, Scopus, and Web of Science) was performed to identify systematic reviews and meta-analyses published in English between 2021 and 2025. The methodological quality of the included studies was evaluated using the Joanna Briggs Institute (JBI) Critical Appraisal Checklist.

Results: Eight high-quality systematic reviews were included, with machine learning identified as the predominant technology. Three primary outcome domains emerged: predictive, psychosocial, and hospital utilization. Due to measurement heterogeneity, the results were synthesized narratively. Our findings demonstrated that AI-based nursing interventions are effective in predicting adverse clinical events, unplanned hospital utilization, and health care costs. However, evidence regarding psychosocial outcomes remains insufficient.

Conclusions: This review provides systematic evidence supporting the utility of AI in chronic illness management, particularly for improving predictive and utilization outcomes. These findings offer actionable insights for nursing leaders to integrate AI into clinical practice and education. Future research should prioritize rigorous empirical designs to further strengthen the evidence base for AI-driven nursing care.

(JMIR Nursing 2026;9:e97905) doi:10.2196/97905

KEYWORDS

artificial intelligence; chronic illness; nursing intervention; systematic review; umbrella review

Introduction

Background

In recent years, the evolution of artificial intelligence (AI) has led to major transformations in the nursing domain. The World Health Organization (WHO) announced that health care systems should use AI technologies for patient-centered and continuous care [1]. The International Council of Nurses (ICN) has highlighted that the adoption of AI is an innovative approach to health care and a potentially beneficial tool for providing nursing care to patients to enhance global health and wellness [2]. Ultimately, AI technologies adopted in health care systems have been integrated into nursing research, practice, and health care systems.

Globally, chronic illnesses have been a serious public health problem for decades. Referred to as “the Silent Pandemic,” chronic illnesses account for approximately 74% of global deaths and pose challenges to global health systems and the economy [1]. Individuals with chronic illnesses experience multiple comorbidities, require continuous care from nurses, and experience recurrent planned or unplanned hospital visits. Evidence from nursing research has shown that interventions by nursing professionals improve chronic illness care. Nurses are key health care professionals who provide patient-centered and effective coordinated care to improve chronic illness outcomes [3,4].

With the evolution of AI in nursing, extensive efforts have been made to apply AI technologies and algorithms. Specific AI techniques in nursing include machine learning (ML), deep learning (DL), robotics, decision support systems, and natural

XSL•FO
RenderX

language processing [5,6]. AI algorithms are computational procedures that enable computers to execute specific tasks by following predefined rules and processing data. Common AI algorithms include decision trees, logistic regression, random forests, neural networks, and Naïve Bayes [7]. AI techniques and algorithms have been used to predict patient conditions and risks, support disease diagnosis, and provide personalized care [8,9].

Several reviews have examined AI-based health care interventions and their impact on patient care [10-12]. Previous evidence has also explored the application of AI in nursing research, practice, and education [5,13,14]. In addition, reviews focusing on AI-based nursing interventions for chronic illness care have been conducted [15-17]. However, the findings regarding the effectiveness of AI-based nursing interventions remain inconsistent because of varied outcome measures and a lack of consensus regarding their clinical impact. This evidence gap may hinder the advancement and implementation of AI-supported nursing care for chronic illnesses. Therefore, a comprehensive synthesis of current evidence is needed. Accordingly, this umbrella review aimed to identify and evaluate the effectiveness of AI-based nursing interventions for chronic illnesses and their impact on health care outcomes in clinical settings.

Objective

This study addressed two research aims: (1) to identify nursing interventions in health care settings that apply AI in caring for chronic illnesses to influence health care outcomes and (2) to synthesize and evaluate the impacts of AI-based nursing interventions on chronic illnesses.

Methods

Study Design

This study followed the umbrella review methodology developed using the Joanna Briggs Institute (JBI) guidelines [18]. An umbrella review is an approach that compiles the best evidence from existing published systematic reviews (SRs) to obtain a single comprehensive review [18]. This umbrella review methodology is useful for aggregating and expanding recent issues and for highlighting synthesized evidence for future directions in health care research [18]. This umbrella review was conducted using the PICO (Population, Intervention, Comparison, and Outcome) framework to define the eligibility criteria and guide paper selection [19]. The review process adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (see Checklist 1). The PICO framework was considered appropriate because the review sought to capture a broad range of evidence regarding AI-based nursing interventions across diverse intervention designs, health care settings, and patient outcomes. This review has been registered with the PROSPERO (International Prospective Register of Systematic Reviews) database. No institutional review board approval was required because no humans were included in the study.

Search Strategy

Five electronic bibliographic databases (CINAHL, Cochrane Library, PubMed, Scopus, and Web of Science) were searched to identify relevant studies. The complete search strategy for each database, including Boolean operators and keyword combinations, was developed in consultation with the research team and is provided in Table S1 of Multimedia Appendix 1 to enhance reproducibility.

The following keywords and Medical Subject Headings were used with Boolean operators (AND and OR): (“artificial intelligence” OR “AI” OR “machine learning” OR “technology”), (“systematic review” OR “meta-analysis”), and (“chronic” OR “chronic disease” OR “chronic illness”) or (“nurse” OR “nursing”). Several additional search terms were used: “prediction” or “effectiveness” or “health outcome” or “nursing research.” To capture the most current evidence regarding the rapid evolution of AI in nursing, we limited our selection to SRs published within the past 5 years. Therefore, SRs with and without meta-analyses published between January 2021 and December 2025 were included in this study. This time frame was selected because AI technologies in nursing and health care have rapidly evolved in recent years, particularly following the increased implementation of ML, conversational agents, and predictive analytics in clinical practice after 2020. Restricting this review to recent SRs enabled the inclusion of the most contemporary evidence reflecting current AI capabilities, health care infrastructure, and nursing practices. However, we acknowledge that this restriction may have excluded earlier relevant evidence and potentially introduced selection bias. Articles available ahead of print publications or as online-first publications were also included. The initial search was conducted on December 1, 2025, and a confirmatory search was conducted on February 28, 2026, to identify potentially eligible studies.

Inclusion and Exclusion Criteria

Textbox 1 shows the inclusion and exclusion criteria. Based on the PICO framework, the Population (P) included patients with chronic illnesses (eg, diabetes, hypertension, and chronic kidney disease) who received nursing care services; the Intervention (I) was defined as the integration of AI-based technologies into nursing interventions in health care settings. The Comparison (C) included the usual care, control groups, or non–AI-based interventions. The Outcome (O) included quantitative patient outcomes. The eligible studies were SRs with or without meta-analysis that included primary empirical studies such as cross-sectional studies, quasi-experimental studies, and randomized controlled trials (RCTs).

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Quantitative systematic review or systematic review with meta-analysis focused on nursing interventions (nurse-led or nurses involved and nursing scholars conducted) that explicitly applied artificial intelligence (AI)–based technologies to chronic ill patients
- Systematic reviews that targeted chronically ill patients who diagnosed diseases or admitted to hospitals or health care facilities with nursing care services
- Systematic reviews with nursing interventions (nurse-led interventions) that explicitly and strictly applied AI-based technologies to patients who may benefit
- Systematic reviews published in peer-reviewed academic journals
- Systematic reviews synthesized empirical studies as their included primary studies
- Systematic reviews published from January 2021 to December 2025 (included ahead of print)

Exclusion criteria

- Editorials, policy papers, commentaries, or qualitative synthesis or qualitative reviews
- Reviews that did not include patients as participants and did not report patient outcomes
- Reviews that did not use AI-powered health technologies and not mainly delivered by nurses (eg, papers that used clinical decision systems that are not based on AI are excluded)
- Reviews that were published in non–peer-reviewed academic journals
- Reviews that did not focus on nursing or nurses or were not conducted by nursing scholars
- Studies in languages other than English

In this umbrella review, AI-based nursing interventions were operationally defined as nursing-related interventions that incorporated AI technologies to support assessment, clinical decision-making, patient monitoring, communication, education, or care coordination for patients with chronic illnesses. In this review, AI-based nursing interventions refer to those that incorporate AI technologies, such as ML models, predictive analytics, or clinical decision support systems, to support nursing assessment, decision-making, monitoring, or care coordination. These interventions included nurse-led applications implemented directly by nurses or nurse practitioners (NPs), as well as AI-supported systems integrated into nursing workflows within health care settings. Studies were included if AI technologies were integrated into nursing care delivery, nursing assessment, nurse-led chronic disease management, or other nursing-related clinical workflows. However, studies focusing solely on institution-level administrative AI systems without direct relevance to nursing care delivery were excluded. This umbrella review also excluded publications that were not published in English or that indicated the use of generative AI, such as ChatGPT, in manuscript writing.

In this umbrella review, “adverse events” were operationally defined as undesirable clinical outcomes associated with chronic illness progression or health care utilization, including mortality, unplanned hospital readmission, intensive care unit (ICU) admission, and disease-related deterioration. “Complications” referred to disease-related secondary conditions such as neuropathy, nephropathy, retinopathy, or other chronic illness–related sequelae. “Hospital utilization outcomes” included hospital readmissions, emergency department visits, length of hospital stay, and health care–related costs. To improve consistency and interpretability, these terms were used consistently throughout this review.

Study Selection

Two reviewers independently screened the titles, abstracts, and full texts based on the predefined eligibility criteria. Any disagreements regarding study inclusion were resolved through discussion and consensus among all research team members. When consensus could not be reached initially, the final decision was made after a repeated review of the full-text article and methodological criteria. To assess potential double-counting bias, overlap among the included SRs was examined using a citation matrix. The primary studies included in each SR were extracted and compared across reviews to identify overlapping studies and duplicate publications.

Quality Appraisal of Selected SRs

For umbrella reviews, it is recommended to assess the professional and methodological quality of the included SRs using a validated checklist [18,20]. This study used the JBI Critical Appraisal Checklist for SRs [21]. This checklist is recommended for umbrella reviews and provides a structured assessment of the methodological quality of SRs. It was developed to evaluate research methodologies and address the risks of bias in included studies when conducting an umbrella review [20]. Therefore, it was considered appropriate for the methodological approach in this umbrella review. The checklist consists of 11 questions in the following domains: study aim, search strategy, inclusion and exclusion criteria, data selection and extraction, publication bias, and appropriate results. These questions have 4 response options: “yes,” “no,” “unclear,” and “not applicable.” The total sum of responses ranged from 0 to 11; if the answer was “yes,” it was scored as 1; otherwise, it was scored as 0. The JBI checklist does not provide predefined cutoff scores to classify reviews as high, moderate, or low quality. However, high scores indicate high quality and low risk

XSL•FO
RenderX

of methodological bias. The first and second researchers independently evaluated the quality of the 8 selected studies using the checklist.

Data Extraction and Synthesis

In this study, data were extracted by the primary researcher using an Excel data sheet. A standardized Excel-based extraction form was developed prior to data extraction to improve consistency and reproducibility. The form included (1) bibliographic details (authors, types and designs of reviews, authors' nationality, and publication years), (2) aims, (3) number of primary studies used in an SR, (4) details of databases used, (5) quality assessment tools, and (6) study participants. A summary of the types of AI reported in the selected studies, the algorithms used, and the findings of common health care outcomes was also extracted onto another data sheet. The primary researcher extracted the data, and the third researcher verified the extraction.

Due to substantial heterogeneity across the included SRs in terms of AI techniques, intervention characteristics, patient populations, chronic illness types, outcome measurements, and study designs, a quantitative meta-analysis was considered inappropriate. In particular, the included reviews used diverse outcome measures and varying definitions of health care outcomes, which limited the comparability of pooled effect sizes. Therefore, a narrative synthesis approach was adopted to comprehensively summarize and interpret the findings.

Results

General Characteristics of Included Reviews

Figure 1 presents a summary of the search process and study selection outcomes using a PRISMA flow diagram. A total of 541 records were initially identified from the 5 electronic databases. After removing duplicate records (n=190) using EndNote X9 (Clarivate Analytics), 351 unique records remained for screening. Of these, 284 records were excluded during title screening, and 44 records were excluded after abstract screening based on the predefined eligibility criteria. Full-text eligibility assessment was conducted for the remaining 23 reports, and 15 reports were excluded because they did not meet the inclusion and exclusion criteria. Finally, 8 SRs were included in this umbrella review.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram.

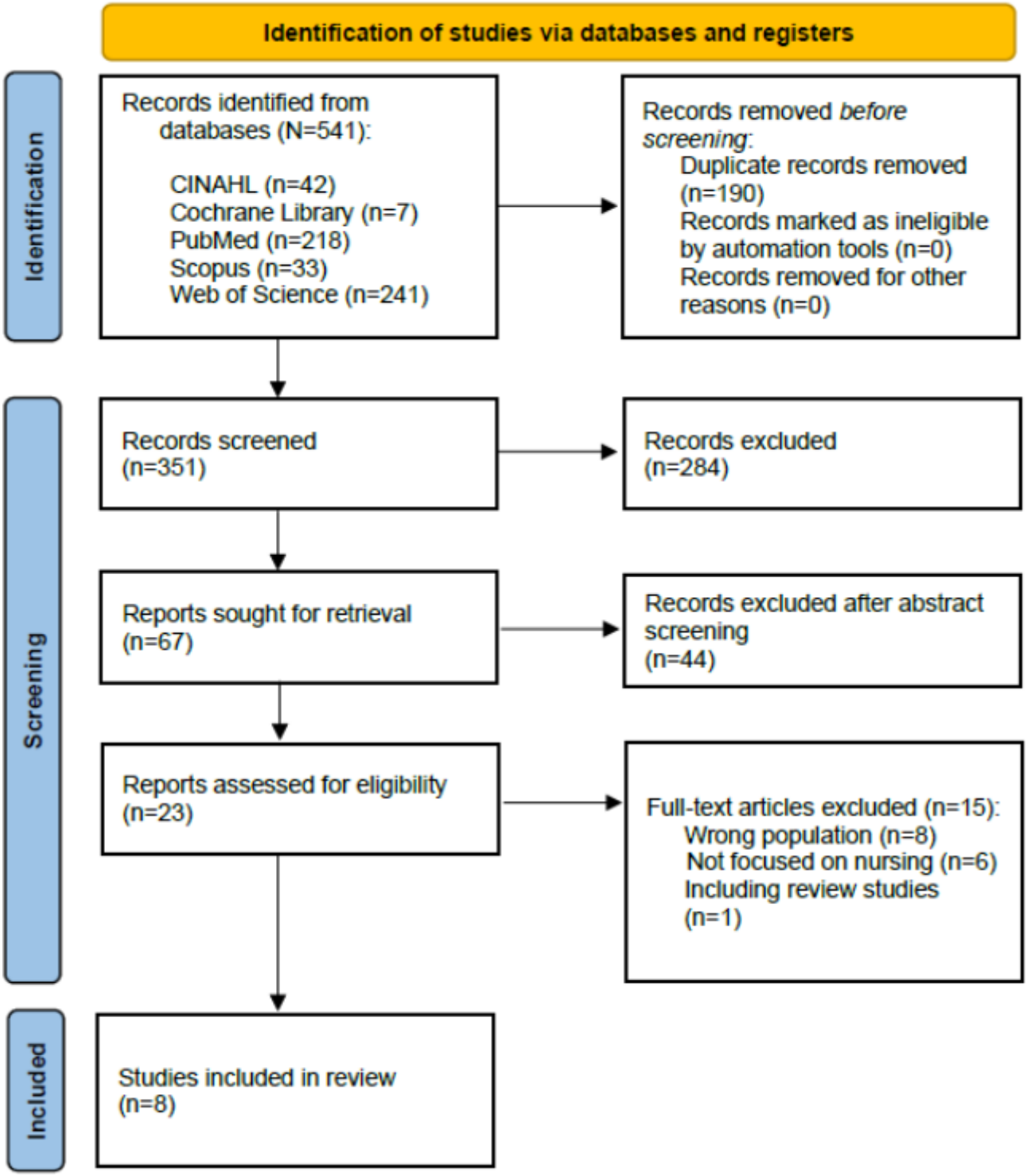


Table 1 presents the general characteristics of the 8 SRs, including the study aim, number of primary studies included, review methods, sample sizes, and year of publication. The 8 SRs were published between 2021 and 2025 [4,15-17,22-25]. These studies investigated the impact of AI-based nursing interventions on patients with chronic illnesses and analyzed their health care–related outcomes resulting from the interventions. The 8 included reviews were conducted in China (n=2), the United Kingdom, the United States, Canada, South Korea, Indonesia, and Slovenia (n=1 each), and they all followed the PRISMA guidelines . Three SRs had registered their review protocols with PROSPERO, while the rest did not report any registration. Three SRs included only RCTs [16,22,25]. Among the included SRs, only 1 SR conducted a meta-analysis [17].

The 8 SRs included a total of 150 primary study records. A citation matrix identified 7 overlapping primary studies across the reviews, corresponding to 8 duplicate publications. After accounting for overlap, 142 unique primary studies were represented in this umbrella review (Table S4 in Multimedia Appendix 1). Given the relatively small number of overlapping studies, the risk of double-counting bias was considered limited. Among these, 41 studies were RCTs, and the rest were

XSL•FO
RenderX

quasi-experimental, retrospective cohort, or cross-sectional studies. Each SR included 5 to 65 primary studies published across various continents, such as North and South America, Europe, and Asia, from 2008 to 2024. Although all the SRs were published in English, the primary studies included in the reviews were published in the English, Chinese [17], or Korean [25] languages. These 2 SRs, which included primary studies in languages other than English, used translation and back-translation methods.

Based on the inclusion criteria, the participants were patients aged at least 16 years or older who had at least 1 chronic illness, such as type 2 diabetes mellitus or cancer, had been hospitalized in ICUs, or had visited the emergency department due to their chronic illness. The number of participants in primary studies ranged from 28 to 469,976.

XSL•FO
RenderX

Table . Characteristics of the included reviews (N=8).

Review (year) country	Title	Aim	Review characteristics (number of primary studies: primary study designs/review methods/PROSPERO[a] registration/review type)	Search characteristics (DBs[b]/time frame/search strategy/number of countries of the primary studies)	Study characteristics (patients; datasets/sample size range or total number of samples/quality assessment tool of review/year of primary studies)
Bressler et al [15], (2025) United States	Leveraging AI/ML models to identify potential palliative care beneficiaries: a systematic review	To examine how ML[c] has been used in palliative care, specifically focusing on models used to identify potential beneficiaries of palliative services among individuals with chronic and terminal illnesses	• 5: 1 RCT[d], 1 quasi-experimental study, and 3 retrospective cohort studies • PRISMA[e] • N/A[f] • SR[g]	• 4: PubMed, CINAHL, Cochrane, and Ovid • No restrictions • Search terms used English only publications • 2 (United States 4 and Germany 1)	• Patients comprising adults (aged >19 y) living with advanced illnesses • 53,836 adult patients • Revised Cochrane Risk-of-Bias tool for randomized studies (RoB 2.0; Sterne et al, 2019) and the Newcastle Ottawa Scale (NOS; Wells, 2021) • 2018-2023
Gosak et al [16], (2024) Slovenia	AI-based prediction models for individuals at risk of multiple diabetic complications: a systematic review of the literature	To determine the effectiveness of predicting multimorbid diabetes-related complications with AI[h]-based models and determine which methods provide the best results in terms of prediction performance	• 11: 11 RCTs • PRISMA • N/A • SR	• 4: PubMed, CINAHL, MEDLINE, and Scopus • No restrictions • Search terms used English language publications • N/A	• T2DM[i] patients; datasets from electronic health record • 72 to 135,492 • NR[j] specific tool (based on inclusion and exclusion criteria) • 2015-2021
Hu et al [17], (2025) China	Prognostic models for unplanned intensive care unit readmission risk prediction: a systematic review and meta-analysis based on HSROC model	To identify prognostic models for unplanned ICU[k] readmission and compare the performance of ML models with scoring systems	• 65: Cohort studies • PRISMA • N/A • SR and meta-analysis	• 10: CINAHL, PubMed, EMBASE, Cochrane, ProQuest, Web of Science, CNKI (Chinese), VIP (Chinese), Wan fang (Chinese), ClinicalTrial.gov • No restrictions • Keywords and search terms used. • 17 (N/A, but English and Chinese publications)	• Patients aged 16 years or older who were hospitalized in the ICU for more than 24 hours • 421 to 469,976 • Prediction Model Risk of Bias Assessment Tool (PROBAST) • 2008-2024
Kurniawan et al [22], (2024) Indonesia	A systematic review of AI-powered chatbot intervention for managing chronic illness	To assess user satisfaction, intervention efficacy, and the specific characteristics and AI architectures of chatbot systems designed for chronic diseases	• 10: RCTs (10) • PRISMA • Yes • SR	• 6: PubMed, CINAHL, Embase, PsycINFO, ACM Digital Library, and Scopus • 2013-2021 • Keywords and search terms used. English publications • 7 (United States 3, France, Spain, Australia, Germany, Switzerland, Egypt)	• Patients with various chronic conditions • 37 to 187 • RoB 2.0 tool • 2013-2021

XSL•FO
RenderX

Review (year) country	Title	Aim	Review characteristics (number of primary studies: primary study designs/review methods/PROSPERO[a] registration/review type)	Search characteristics (DBs[b]/time frame/search strategy/number of countries of the primary studies)	Study characteristics (patients; datasets/sample size range or total number of samples/quality assessment tool of review/year of primary studies)
Li et al [23], (2023) Hong Kong, China	Feasibility and effectiveness of AI-driven conversational agents in health care interventions: a systematic review of randomized controlled trials	To examine the feasibility and effectiveness of conversational agent-based interventions evaluated by RCTs in the health care context, as well as to evaluate the information quality of AI-driven conversational agents	• 21: RCTs (21) • PRISMA • Yes • SR	• 7: Scopus, PubMed, Embase, PsycINFO, Cochrane Library, Information Science & Technology, and Web of Science • No restrictions • Keywords and search terms used. English publications • 9 (United States 11, South Korea 2, United Kingdom, France, Spain, Australia, Switzerland, Argentina, Japan)	• Adult population with various chronic conditions (at least 18 y) • 28 to 927 (mean=160), with a median of 82 participants • RoB 2.0 tool • 2013-2021
O'Connor et al [24], (2024) United Kingdom	The application and use of AI in cancer nursing: a systematic review	To (1) examine the areas of cancer nursing AI has been applied in, (2) determine how involved cancer nurses were in AI research, and (3) understand the limitations and risks of AI in cancer nursing	• 20: observational, cross-sectional, case-control, etc • PRISMA • NR • SR	• 4: CINAHL (EBSCOhost), MEDLINE (Ovid), PubMed (Central), and PsycINFO (Ovid) • 2010-2022 • Keywords and search terms used. English publications • 10 (China 7, United States 3, South Korea 3, Denmark, Iran, Italy, Japan, Switzerland, Taiwan, Netherlands)	• Patients with cancer • 33 to 12,358 • NR • 2011-2022
Raymond et al [4], (2022) Canada	Nurse practitioners' involvement and experience with AI-based health technologies: a systematic review	To characterize NPs'[l] involvement with AI-based health technologies in terms of (1) its expected impact on the clinical activities and performance of NPs and (2) its potential outcomes for NPs' patients and for the general population	• 11: retrospective studies, cross-sectional, prospective, etc • PRISMA • NR • SR	• 4: CINAHL, Medline, PubMed, and Scopus • NR • Keywords and search terms used. English publications	• ED[m], primary, or hospitals (tertiary) visited patients • NR • Joanna Briggs • 2018-2022
Yi et al [25], (2025) South Korea	The effects of applying AI to triage in the emergency department: a systematic review of prospective studies	To identify the intervention contents of AI-based emergency patient triage systems and investigate the intervention effect of AI-based emergency patient triage systems	• 7: 5 RCTs and 2 quasi-experimental designs • PRISMA • Yes • SR		• ED visited patients • 146 to 17,072 • STROBE[n] protocol • 2017-2023

XSL•FO
RenderX

Review (year) country	Title	Aim	Review characteristics (number of primary studies: primary study designs/review methods/PROSPERO[a] registration/review type)	Search characteristics (DBs[b]/time frame/search strategy/number of countries of the primary studies)	Study characteristics (patients; datasets/sample size range or total number of samples/quality assessment tool of review/year of primary studies)
				• 7: CINAHL, PubMed, EMBASE, Cochrane, ProQuest, RISS, and KISS • No restrictions • Search terms used English and Korean publications • 6 (Greece 2, South Korea 1, Germany 1, Iran 1, Taiwan 1, China 1)	

[a]PROSPERO: International Prospective Register of Systematic Reviews.

[b]DB: database.

[c]ML: machine learning.

[d]RCT: randomized controlled trial.

[e]PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

[f]N/A: not applicable.

[g]SR: systematic review.

[h]AI: artificial intelligence.

[i]T2DM: type 2 diabetes mellitus.

[j]NR: not reported.

[k]ICU: intensive care unit.

[l]NP: nurse practitioners.

[m]ED: emergency department.

[n]STROBE: Strengthening the Reporting of Observational Studies in Epidemiology.

Summary of AI-Based Nursing Interventions in the SRs

Table 2 summarizes the characteristics of AI-based nursing interventions used in the studies. Additionally, Tables S2 and S3 in Multimedia Appendix 1 summarize the characteristics of AI-based nursing interventions used in the studies. Table S2 in Multimedia Appendix 1 provides the number of studies that use different types of AI interventions. All the primary studies clearly reported the utilization of AI techniques and AI algorithms in their nursing interventions. In this study, the AI-based nursing interventions in the reviews were classified as AI techniques [26]. The main AI interventions retrieved from the studies included ML, which was used in all 8 studies, followed by DL, clinical decision support, and natural language processing. The types of AI algorithms used were the decision tree model, logistic regression, neural networks, and Naïve Bayes. Several AI algorithms, such as regressions, Naïve Bayes, and neural networks, were used as prediction models. The most frequently used AI algorithm was logistic regression (used in all 8 SRs). Two SRs reported that chatbots and conversational agents [22,23] were used in their nursing interventions, and 1 SR used real-time voice AI support [25].

XSL•FO
RenderX

Table . Summary of artificial intelligence (AI)–based nursing interventions and outcomes in the selected systematic reviews.

Review (year)	AI-powered interventions		Major findings
	Classification of AI techniques and types of AI algorithms	Comparison	
Bressler et al [15], 2025	• ML[a] • DL[b] • Decision support system • Natural language processing • Decision tree model • Logistic regression • Neural network	• Usual care or NR[c]	Predictions • Predict mortality risk (+)[d]: accurate and high performance in predicting 6-month (0.978), 1-year (0.956), and 2-year (0.943) mortality (2 studies) Hospital utilization outcomes • Hospital readmissions (+): significant decreased hospitalizations at 60 days and 90 days (2 studies) • Health care–related cost (+): cost saving (2 studies) Psychosocial outcomes (+) • Quality of care (+): increased care quality (1 study)

Gosak et al [16], 2022	• ML
	  • DL
	• Random forest
	• Naïve Bayes
	• Decision tree model
	• Logistic regression
	• Neural network
		• NR	Predictions
			• Predict T2DM[e] complications (+):
			  • Risk of diabetic neuropathy (7 studies)
			  • Risk of diabetic nephropathy (6 studies)
			  • Risk of diabetic retinopathy (2 studies)

Hu et al [17], 2025	• ML
	• DL
	• Logistic regression
		• Usual care with scoring systems	Predictions
			• Predict unplanned ICU[f] readmission (+):
			• Scoring systems: pooled sensitivities of 0.607, with pooled specificity of 0.699
			  • ML models: pooled sensitivities of 0.711, with pooled specificity of 0.899
			  • DL models: pooled sensitivity of 0.745, with pooled specificity of 0.709 (indicated a higher rate of false positives)
			  • Models predicting ICU readmission within 7 days: sensitivity=0.462, specificity=0.788
			  • Models predicting ICU readmission within 30 days: sensitivity=0.789, specificity=0.835

XSL•FO
RenderX

Review (year)	AI-powered interventions Classification of AI techniques and types of AI algorithms	Comparison	Major findings
Kurniawan et al [22], 2024	• AI-powered chatbots • ML • Natural language processing (3 of the included studies did not specify AI models)	• Usual care and education by nurse or traditional writing or counseling	Psychosocial outcomes: • Anxiety (+): positive results in the IGs[g] (3 studies; anxiety) • Depression, pain: no differences between groups (3 studies; depression, pain) • Participant satisfaction (+): significantly favorable and user-friendly to use (7 studies) • Patient safety (–)[h]: no transparent reporting (false alarm) of patients' adverse events (1 study)
Li et al [23], 2023	• AI-driven conversational agents–based intervention (add dosage) • ML • Natural language processing	• Usual care	Psychosocial outcomes (+): • Physical activity and function (+): significant improvement in participants' physical activity and function in the IGs (3 studies) • Self-efficacy (+): significantly increased self-efficacy in healthy lifestyle maintenance (weight loss; 1 study) • Illness-specific knowledge and understanding (+): significant increases in illness-specific knowledge and understanding of disease information (3 studies) • QoL[i] and well-being (+): improvement of QoL and well-being (4 studies) • Depression (+): significant improvements in reducing depression (5 studies)
O'Connor et al [24], 2024	• AI-based decision support system • ML • Natural language processing • Random forest • Naïve Bayes • Decision tree model • Logistic regression • Neural network	• Cancer nursing care	Predictions • Predict clinical issues and complications (+): 17 studies • Predictions: Predict hospital length of stay (+): significant improvements (2 studies) Psychosocial outcomes (+) • Quality of care (+): increased care quality (1 study)
Raymond et al [4], 2022	• ML • Clinical decision support system • Natural language processing		Psychosocial outcomes (±)[j] • Quality of care (±): improve the quality of health care provided by NPs (8 studies), no improvement of quality of care (1 study) • Patient safety (±): improve patient safety (3 studies) Hospital utilization outcomes • Rehospitalization (+): significantly reduce rehospitalization (1 study) • Health care–related cost (+): cost saving in hospital care (2 studies)

XSL•FO
RenderX

Review (year)	AI-powered interventions		Major findings
	Classification of AI techniques and types of AI algorithms	Comparison	
Yi et al [25], 2025	• Natural language processing • ML • DL • Random forest • Naïve Bayes • Decision tree • Logistic regression • Neural network	• Manual care by nurses or physicians	Predictions: Triage prediction (+): • By using ESI (emergency severe index), accuracy of the AI-powered model's triage prediction ranged from 80.5% to 99.1% (4 studies) • IG shows 0.9% higher accuracy with a low miss-triage rate than CG[k] (4 studies)

[a]ML: machine learning.
[b]DL: deep learning.
[c]NR: not reported.
[d](+): significant effectiveness.
[e]T2DM: type 2 diabetes mellitus.
[f]ICU: intensive care unit.
[g]IG: intervention group.
[h](–): no effect.
[i]QoL: quality of life.
[j](±): inconclusive and mixed results.
[k]CG: control group.

Most AI-based interventions were led by nurses or NPs [22,24,25] and delivered by nurses or NPs within multidisciplinary health care settings [4,15-17]. These interventions were directly integrated into nursing workflows, including patient monitoring and education, symptom assessment, discharge planning, follow-up, and chronic illness management [4,15-17,22,24]. Some SRs incorporated AI decision support systems and predictive algorithms [16,17,25]; however, these AI technologies were implemented within supporting nursing interventions such as assessment, clinical decision-making, and care coordination in nursing practices. The primary studies were conducted in clinical care settings, such as long-term care settings, intensive care, or emergency departments in hospitals or primary care settings.

Seven SRs reported the usual or manual care group as the control groups; however, in the remaining SR [16], such information was not clearly reported. Li et al [23] specifically described the dosage of nursing interventions, which were assisted by AI-driven conversational agents, as reported in the primary studies.

Methodological Evaluation Results

Table 3 presents the quality assessment results of the 8 SRs. The JBI Critical Appraisal Checklist tool does not have cutoff scores; however, study scores ranging from 8 to 11 are referred to as high-quality studies, indicating a low risk of bias [27]. The review by O'Connor et al [24] scored 9 because they did not clearly and adequately report the resources that were used for the study searches and the methods used to minimize data extraction errors. The other 7 SRs scored between 10 and 11, clearly indicating the methodological directions in their reviews. Thus, all the SRs indicated a low risk of methodological bias.

XSL•FO
RenderX

Table . Joanna Briggs Institute (JBI) Critical Appraisal Checklist for systematic reviews.

JBI Critical Appraisal Criteria	Bressler et al [15], 2025	Gosak et al [16], 2024	Hu et al [17], 2025	Kurniawan et al [22], 2024	Li et al [23], 2023	O'Connor et al [24], 2024	Raymond et al [4], 2022	Yi et al [25], 2025
Is the review question clearly and explicitly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the inclusion criteria appropriate for the review question?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the search strategy appropriate?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the sources and resources used to search for studies adequate?	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes
Were the criteria for appraising studies appropriate?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was critical appraisal conducted by 2 or more reviewers independently?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were there methods to minimize errors in data extraction?	Yes	Unclear	Yes	Yes	Unclear	Unclear	Unclear	Yes
Were the methods used to combine studies appropriate?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Was the likelihood of publication bias assessed?	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Were recommendations for policy and/or practice supported by the reported data?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Were the specific directives for new research appropriate?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Total score (quality score)	11	10	11	10	10	9	10	11

XSL•FO
RenderX

Main Patient Outcomes

Overview

Among the 8 included SRs, the primary patient outcomes of AI-based nursing interventions were categorized into 3 common domains: (1) predictive outcomes (n=5), (2) psychosocial outcomes (n=5), and (3) hospital utilization outcomes (n=2; Tables S2 and S3 in Multimedia Appendix 1). Table 2 summarizes the common outcomes of these reviews. Due to the heterogeneity of the results, the findings are synthesized narratively.

Predictive Outcomes

In the 5 SRs that reported predictive outcomes, the predictions of adverse events, such as disease-related complication risks, mortality risks, and possible hospital utilization, were commonly reported using predictive algorithms [15-17,24,25].

Three SRs reported predicted possible complications related to diseases using AI-based nursing interventions [15,16,24]. In a high-quality review, Gosak et al [16] reported that nursing interventions applying AI techniques could predict the possible complications of type 2 diabetes mellitus. In this review, the risks of neuropathy, nephropathy, and retinopathy associated with diabetes were precisely predicted using AI prediction algorithms, such as penalized regression, random forest, and neural networks. Bressler et al [15] predicted the mortality risks of patients at 6 months, 1 year, and 2 years. These predictions enabled the early and timely detection of nurses' care interventions. O'Connor et al [24] reported the predicted risk of complications in patients with cancer in 17 primary studies. Yi et al [25] reviewed emergency triage predictions and found that intervention groups for which AI models were applied showed 0.9% more accurate predictions in emergency triage than those shown by control groups. In addition, the accuracy of emergency triage prediction rates ranged from 80.5% to 99.1% in the 4 primary studies. Thus, the results indicated that AI-based interventions had an impact on predicting adverse outcomes in patients.

Two SRs reported predictions of unplanned hospital utilization [17,24]. In the study by O'Connor et al [24], the length of hospital stay was predicted using 2 primary studies, with a conclusion of a significant improvement in the reduction of length of hospital stay predictions. Hu et al [17] predicted ICU readmissions using pooled results from a meta-analysis of 56 primary studies. They found that ICU readmissions within 30 days showed higher sensitivity (0.789, 95% CI 0.751-0.824) and specificity (0.835, 95% CI 0.805-0.860); sensitivity and specificity within 7 days were 0.462 (95% CI 0.118-0.846) and 0.788 (95% CI 0.448-0.945), respectively. Hu et al [17] reported that using short-term expectations to predict ICU readmissions is more complex.

Psychosocial Outcomes

Five SRs reported patients' psychosocial outcomes as results of AI-based nursing interventions [4,15,22-24]. In the SRs, patients' self-reported outcomes, such as quality of care, patient safety, patient satisfaction, and depressive symptoms, were measured.

Regarding the quality of care, 2 SRs demonstrated that the quality of care increased [15,24]. However, Raymond et al [4] reported mixed results in a high-quality SR. There was an improvement in the quality of care provided by NPs in 8 primary studies; however, no improvement was reported in 1 study. The results of patient safety outcomes were mixed. In the study by Kurniawan et al [22], AI-based nursing interventions could not support patient safety. Raymond et al [4] reported improvements in patient safety in 3 primary studies. Regarding patient satisfaction outcomes, Kurniawan et al [22] reported significant differences between groups in 7 primary studies after applying AI-based nursing interventions. Two SRs that reported depression as the primary outcome of AI-based nursing interventions for chronic illnesses [4,23] reported a significant reduction in depression in their 5 included primary RCTs. However, Kurniawan et al [22] reported no differences between the intervention and control groups in 3 studies. Therefore, the impact of AI-based nursing interventions on psychosocial outcomes remains inconclusive.

Hospital Utilization Outcomes

Two SRs reported total hospital utilization outcomes resulting from AI-based nursing interventions for chronic illnesses [4,15]. Both SRs investigated hospital readmissions and health care–related costs, and the results were conclusive. In the review by Bressler et al [15], 60- and 90-day hospital readmissions were significantly decreased in 1 primary study. Similarly, Raymond et al [4] reported that 2 primary studies showed significant reduction in rehospitalizations. Both SRs reported cost savings in intervention groups compared with control groups in hospital-based settings [15] and hospital care, especially in emergency care settings [4]. Both SRs, which were high-quality reviews, concluded the cost-saving capabilities of AI-based nursing interventions.

Discussion

Principal Findings

The integration of AI techniques into patient care has been used by nurses in various health care settings and has been reported to improve patient health care outcomes. This umbrella review expands on the current inconclusive results of AI-based nursing interventions to provide robust evidence on AI-based nursing interventions for chronic illnesses. The first aim of this umbrella review was to summarize nursing interventions that applied AI in caring for patients with chronic illnesses in health care settings, which may have affected patients' health care outcomes by analyzing individual SRs conducted in 7 countries. In this study, we synthesized the findings from 8 SRs focusing on AI-based nursing interventions for chronic illnesses that were conducted using 143 primary studies. All 8 SRs were published recently within the past 5 years, from 2021 to 2025. This study identified the types of AI techniques and algorithms used in recent evidence-based research. Regarding the AI techniques applied in nursing interventions, ML, decision-support systems, and DL, which is a subset of ML [28], were identified in the included SRs. These findings should be interpreted in the context of the specific AI modality rather than as evidence for a single homogeneous category of AI-based nursing interventions.

XSL•FO
RenderX

Among the AI algorithms, logistic regression models were most frequently used in the included reviews, as 5 of the SRs focused on predicting adverse outcomes.

Based on the secondary aim of this review, we synthesized 3 primary and common patient outcomes—predictive, psychosocial, and hospital utilization outcomes. This review identified the advantages of AI-based nursing interventions in predicting adverse events and hospital utilization outcomes. According to the SRs, applying AI algorithm models to intervention groups enabled better prediction of potential risks in patients with chronic illnesses, such as complications and unplanned hospital utilizations, compared with control groups without AI assistance. Chronic illnesses are associated with multiple comorbidities and complications [29]; thus, early-detection systems using AI are important [8,30]. Consistent with previous reviews, AI-based predictive models impact disease-related risk predictions [31,32]. In this umbrella review, we found that nurses could identify the possible risks of adverse events during AI-assisted nursing interventions. Subsequently, the nurses perceived patients' care needs with AI algorithms and proactively provided appropriate care for patients with chronic illnesses to prevent possible adverse events [16,25]. From the included SRs, we pooled data to predict unplanned ICU readmissions using the AI technique for critical care nurses [17]. The results showed that applying AI techniques affected nurses' practice, ensuring the prioritization of high-risk patient care. However, there could be accuracy issues and an algorithm risk of bias in AI-assisted predictions [17,25]. Thus, more rigorous AI models that can support nurses' interventions and larger datasets to critically predict chronic illness risks are required.

This umbrella review also revealed reductions in hospital readmissions and health care–related costs. Two SRs reported that AI-based nursing interventions significantly decreased hospitalization and health care costs for chronic illnesses [4,15]. The results regarding hospital utilization outcomes are supported by multidisciplinary studies on patients at high risk or receiving postoperative care [9,33]. In this review, we concluded that hospital utilization prediction was effective; however, 2 of the included SRs (4 primary studies) reported an impact on chronic illnesses. Furthermore, none of the SRs specifically compared the full economic analysis related to hospital use. More robust evidence that includes a large dataset and an elaborate analysis of the impact of AI-based nursing interventions on hospital utilization and cost analysis for chronic illness in health care settings is recommended.

The results regarding psychosocial outcomes were inconclusive. This umbrella review found mixed results on the safety, quality of care, satisfaction, or depression status among patients with chronic illness. Owing to the complex and wide range of psychological outcome measurements, SRs reported varied outcomes. One study demonstrated a false alarm regarding patients' adverse outcomes; thus, it reported no effectiveness of AI-based interventions on patient safety [22]. Recently, some nursing reviews have explored patient safety outcomes using AI [30,34]. However, the results have been inconclusive. Therefore, more studies on clear psychosocial measures of psychological outcomes after AI-based nursing interventions are required to provide evidence.

In the included SRs, evidence regarding the effectiveness of AI-based theoretical frameworks, specifically which AI type or algorithm effectively affects patient outcomes, is limited. In evidence-based practice, it is imperative to have theoretical frameworks that can be accepted as part of AI technology in nursing. One SR by Hu et al [17] concluded that ML outperformed other algorithms in predicting chronic illness outcomes. This finding is congruent with previous reviews, indicating that ML has been the most useful AI technology for nursing interventions [5]. Recent studies have demonstrated that AI models with AI algorithms are highly accurate in predicting patient outcomes for specific diseases in nursing research [5,10]. However, in this umbrella review, none of the SRs indicated that AI algorithms outperformed others in patient outcome predictions, nor did any show that AI algorithms were mostly used to predict patient outcomes. Further empirical studies are required to identify successful implementations of AI algorithm-optimized nursing interventions that are most useful for chronic illness care provision. Furthermore, theoretical frameworks for AI, targeting chronic illness management, should be established to implement AI-based nursing interventions.

When analyzing such AI-based nursing interventions regarding patient outcomes, it is necessary to clearly and robustly include complex datasets. Moreover, it is imperative to combine and refine all nurses' clinical notes, including texts from other health care professionals, to identify the most effective AI algorithms or types [24]. It is important to include nursing notes or unstructured texts when analyzing patient outcomes in AI-based nursing interventions for robust data analysis [4,15,17,24]. Several SRs have reported that nurses' free-text documentation and clinicians' notes are useful for AI-based model development and for predicting patient outcomes [4,15]. Hu et al [17] suggested that ML outperformed other methods in predicting patient outcomes; however, they used only structured datasets and ignored clinical notes. Thus, it may be insufficient to conclude that ML is a promising AI technique for nursing interventions in care provision for chronic illness. It is recommended to use both structured and unstructured data, such as nursing notes, to predict outcome measurements, along with AI-based techniques to eliminate the risk of data bias.

Strengths and Limitations

One strength of this umbrella review is that the 8 SR results included 41 RCTs and experimental design studies, which is a little less than half of the 143 primary studies. A recent review reported that there is a dearth of experimental studies on AI-based intervention in nursing, and it may be possible to prove the effectiveness of AI on patient outcomes in nursing research [5]. In addition, this review provides evidence that elucidates the impact of AI-based nursing interventions on related patient outcomes due to the high methodological quality of SRs.

This study has a few limitations. First, we identified and synthesized 3 common patient outcomes from 8 SRs. This review identified 3 dominant patient outcomes as the impact of AI-based nursing interventions. There could be more positive outcomes and benefits of AI-based technologies. Moreover,

XSL•FO
RenderX

substantial heterogeneity in AI intervention types, outcome measures, patient populations, chronic illness types, and study designs limited the feasibility of conducting a quantitative meta-analysis. Therefore, this review could only aggregate and summarize the findings narratively. More frequently, pooled results reporting patient outcomes with meta-analytical methods are desirable. Given the heterogeneity of AI technologies, intervention characteristics, and outcome measures, these findings should not be interpreted as evidence supporting the effectiveness of AI-based nursing interventions as a homogeneous entity. Second, the 8 included SRs focused only on the effectiveness of AI technologies in participants with chronic illnesses. There was limited information about the dosage of the interventions, indicating a dearth of information about the frequency or duration of AI-based interventions. Only the review by Li et al [23] reported the dosage of interventions in their primary RCTs. Thus, this umbrella review had limited information from the included SRs to perform a dosage analysis of AI techniques and nursing interventions and its impact on patient outcomes. Third, this umbrella review included SRs published in peer-reviewed journals from 5 electronic databases in the past 5 years and excluded gray literature to meet the aim of this review. Thus, some data may have been missing. Furthermore, this study may have a potential overlap of primary studies across the included SRs, which may introduce double-counting bias and affect the validity of the findings. Fourth, participants with chronic illnesses included in the SRs were recruited within health care settings, such as emergency departments or long-term care. There was limited information about their age, number of chronic diseases, or any complications. Fifth, potential publication bias could have been present in the 8 SRs, as they had the tendency to report positive results. Finally, this review included only SRs published in English.

Several challenges should also be considered when integrating AI-based technologies into nursing practice. Algorithmic bias, data quality limitations, and the limited explainability of some AI models may affect the reliability and transparency of AI-assisted clinical decision-making. In addition, interoperability with existing electronic health record systems remains a significant barrier in many health care settings. Clinician trust, ethical considerations, data privacy, and regulatory compliance must also be addressed to support the safe and effective implementation of AI-enabled nursing interventions. Future research should evaluate not only effectiveness but also implementation feasibility, acceptability, and sustainability in clinical practice.

Implications for Nursing Practice and Policy

For frontline nurses caring for patients with chronic illnesses, the findings of this umbrella review will enable them to identify patient risks and effectively integrate AI-based techniques into clinical nursing practice. Nurses and NPs can use the results of this umbrella review as evidence for practice in chronic illness care provision. The ICN emphasizes that nurses should participate in and be educated on the development of AI-based interventions in health care systems and implement these AI technologies in their clinical practice to achieve better patient outcomes and patient safety [2]. Evidence exists on risk predictions, hospital utilization, and health care costs in AI-based nursing interventions for chronic illnesses. Thus, nurses need to explore more patient outcomes, including the investigation of inconclusive psychological outcomes of chronic illnesses. Furthermore, it is recommended that nurses should implement and lead the application of AI in clinical practice to improve patient-centered care [2].

Nursing leaders in global nursing organizations prioritize AI in nursing research, education, and practice [1,2]. The ICN has announced the importance of digital technology education in nursing and nurses' competency in using AI systems [2]. They must continuously leverage the empowerment of the nursing workforce within an AI health care system. They require stable support in nursing research, practice, and policy development domains to continuously realize the value of AI in clinical practice.

Conclusion

This umbrella review is the first to synthesize recent SRs examining the impact of AI-based nursing interventions on chronic illnesses in health care settings. The findings suggest that AI-based nursing interventions may have potential benefits for predictive outcomes and for reducing unplanned hospital utilization. However, evidence regarding psychosocial outcomes remains conflicting and inconclusive. These results should be interpreted cautiously due to the limited number of included reviews and the heterogeneity across studies. Therefore, further high-quality studies are needed to strengthen the identified evidence and optimize nurse-led AI-based interventions across different outcome domains in clinical practice.

Funding

This work was funded by the National Research Foundation of Korea and a grant (RS-2023-00209718129820768700001) from the Korean Government (Ministry of Science, Information, and Communication Technology).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies, artificial intelligence algorithms, review characteristics, outcomes, and study overlap across included systematic reviews.

XSL•FO
RenderX

[DOCX File, 33 KB - nursing_v9i1e97905_app1.docx]

Supplementary data

References

1. Ethics and governance of artificial intelligence for health: guidance for large multi-modal models. : World Health Organization; 2024 URL: https://iris.who.int/server/api/core/bitstreams/e9e62c65-6045-481e-bd04-20e206bc5039/content [accessed 2026-04-11]
2. Nurses together: a force for global health: 2023 ICN congress report. : International Council of Nurses; 2023 URL: https://www.icn.ch/sites/default/files/2024-05/WEB-HD-ICN-RA-2023-ANGLAIS_FINAL.pdf [accessed 2026-04-11]
3. Joo JY, Liu MF, Ho MH. Nurses' perceptions of artificial intelligence adoption in healthcare: a qualitative systematic review. Nurse Educ Pract 2025 Oct;88:104542. [doi: 10.1016/j.nepr.2025.104542] [Medline: 40925173]
4. Raymond L, Castonguay A, Doyon O, Paré G. Nurse practitioners' involvement and experience with AI-based health technologies: a systematic review. Appl Nurs Res 2022 Aug;66:151604. [doi: 10.1016/j.apnr.2022.151604] [Medline: 35840270]
5. Choi J, Lee H, Kim-Godwin Y. Decoding machine learning in nursing research: a scoping review of effective algorithms. J Nurs Scholarsh 2025 Jan;57(1):119-129. [doi: 10.1111/jnu.13026] [Medline: 39294553]
6. Guo Y, Hao Z, Zhao S, Gong J, Yang F. Artificial intelligence in health care: bibliometric analysis. J Med Internet Res 2020 Jul 29;22(7):e18228. [doi: 10.2196/18228] [Medline: 32723713]
7. Sarker IH. Machine learning: algorithms, real-world applications and research directions. SN Comput Sci 2021 May;2(3):160. [doi: 10.1007/s42979-021-00592-x] [Medline: 33778771]
8. Manzhalii E, Dekhtiar Y, Bannikov V, Girnyk G, Bavykin I. Artificial intelligence in clinical diagnostics for early detection of chronic diseases: a systematic review. Georgian Med News 2026 Jan(370):61-73. [Medline: 41804133]
9. Romero-Brufau S, Wyatt KD, Boyum P, Mickelson M, Moore M, Cognetta-Rieke C. Implementation of artificial intelligence-based clinical decision support to reduce hospital readmissions at a regional hospital. Appl Clin Inform 2020 Aug;11(4):570-577. [doi: 10.1055/s-0040-1715827] [Medline: 32877943]
10. Chen C, Lam KT, Yip KM, et al. Comparison of an AI chatbot with a nurse hotline in reducing anxiety and depression levels in the general population: pilot randomized controlled trial. JMIR Hum Factors 2025 Mar 6;12:e65785. [doi: 10.2196/65785] [Medline: 40048637]
11. Wang Q, Lee RLT, Hunter S, Chan SWC. The effectiveness of internet-based telerehabilitation among patients after total joint arthroplasty: an integrative review. Int J Nurs Stud 2021 Mar;115:103845. [doi: 10.1016/j.ijnurstu.2020.103845] [Medline: 33360248]
12. Koo TH, Zakaria AD, Ng JK, Leong XB. Systematic review of the application of artificial intelligence in healthcare and nursing care. Malays J Med Sci 2024 Oct;31(5):135-142. [doi: 10.21315/mjms2024.31.5.9] [Medline: 39416729]
13. Yasin YM, Al-Hamad A, Metersky K, Kehyayan V. Incorporation of artificial intelligence into nursing research: a scoping review. Int Nurs Rev 2025 Mar;72(1):e13013. [doi: 10.1111/inr.13013] [Medline: 38967044]
14. El Arab RA, Al Moosa OA, Abuadas FH, Somerville J. The role of AI in nursing education and practice: umbrella review. J Med Internet Res 2025 Apr 4;27:e69881. [doi: 10.2196/69881] [Medline: 40072926]
15. Bressler T, Song J, Kamalumpundi V, Chae S, Song H, Tark A. Leveraging artificial intelligence/machine learning models to identify potential palliative care beneficiaries: a systematic review. J Gerontol Nurs 2025 Jan;51(1):7-14. [doi: 10.3928/00989134-20241210-01] [Medline: 39746126]
16. Gosak L, Martinović K, Lorber M, Stiglic G. Artificial intelligence based prediction models for individuals at risk of multiple diabetic complications: a systematic review of the literature. J Nurs Manag 2022 Nov;30(8):3765-3776. [doi: 10.1111/jonm.13894] [Medline: 36329678]
17. Hu H, Li J, Ge H, et al. Prognostic models for unplanned intensive care unit readmission risk prediction: a systematic review and meta-analysis based on HSROC model. Nurs Crit Care 2025 Mar;30(2):e13306. [doi: 10.1111/nicc.13306] [Medline: 39995380]
18. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Chapter 10: umbrella reviews. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: Joanna Briggs Institute; 2020. [doi: 10.46658/JBIMES-20-11]
19. Robinson KA, Saldanha IJ, McKoy NA. Frameworks for determining research gaps during systematic reviews. : Agency for Healthcare Research and Quality; 2011 URL: https://www.ncbi.nlm.nih.gov/books/NBK62478/pdf/Bookshelf_NBK62478.pdf [accessed 2026-06-23]
20. Gates M, Gates A, Pieper D, et al. Reporting guideline for overviews of reviews of healthcare interventions: development of the PRIOR statement. BMJ 2022 Aug 9;378:e070849. [doi: 10.1136/bmj-2022-070849] [Medline: 35944924]
21. Aromataris E, Fernandez R, Godfrey C, Holly C, Khalil H, Tungpunkom P. Umbrella reviews. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis: JBI (Joanna Briggs Institute); 2024. URL: https://synthesismanual.jbi.global [accessed 2026-06-26] [doi: 10.46658/JBIMES-24-01]

XSL•FO
RenderX

22. Kurniawan MH, Handiyani H, Nuraini T, Hariyati RTS, Sutrisno S. A systematic review of artificial intelligence-powered (AI-powered) chatbot intervention for managing chronic illness. Ann Med 2024 Dec;56(1):2302980. [doi: 10.1080/07853890.2024.2302980] [Medline: 38466897]
23. Li Y, Liang S, Zhu B, et al. Feasibility and effectiveness of artificial intelligence-driven conversational agents in healthcare interventions: a systematic review of randomized controlled trials. Int J Nurs Stud 2023 Jul;143:104494. [doi: 10.1016/j.ijnurstu.2023.104494] [Medline: 37146391]
24. O'Connor S, Vercell A, Wong D, et al. The application and use of artificial intelligence in cancer nursing: a systematic review. Eur J Oncol Nurs 2024 Feb;68:102510. [doi: 10.1016/j.ejon.2024.102510] [Medline: 38310664]
25. Yi N, Baik D, Baek G. The effects of applying artificial intelligence to triage in the emergency department: a systematic review of prospective studies. J Nurs Scholarsh 2025 Jan;57(1):105-118. [doi: 10.1111/jnu.13024] [Medline: 39262027]
26. Kaur K, Kaur P. The application of AI techniques in requirements classification: a systematic mapping. Artif Intell Rev 2024;57(3):57. [doi: 10.1007/s10462-023-10667-1]
27. Munn Z, Aromataris E, Tufanaru C, et al. Chapter 5: systematic reviews of prevalence and incidence. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: Joanna Briggs Institute; 2020. [doi: 10.46658/JBIMES-20-06]
28. Goodfellow I, Bengio Y, Courville A. Deep Learning: MIT Press; 2016. URL: http://www.deeplearningbook.org [accessed 2026-05-07]
29. About chronic diseases. Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/chronic-disease/about/index.html [accessed 2026-04-11]
30. Watson AL, Bond C, Aveyard H, Smith GD, Jackson D. Generative AI at the bedside: an integrative review of applications and implications in clinical nursing practice. J Clin Nurs 2025 Nov 26. [doi: 10.1111/jocn.70151] [Medline: 41293898]
31. Thomas J. Artificial intelligence in nursing research: a narrative review of transforming clinical practice, enhancing patient outcomes, and shaping future care. J Nurs Rep Clin Pract 2025;3(4):368-374. [doi: 10.32598/JNRCP.2410.1189]
32. Shamout F, Zhu T, Clifton DA. Machine learning for clinical outcome prediction. IEEE Rev Biomed Eng 2021;14:116-126. [doi: 10.1109/RBME.2020.3007816] [Medline: 32746368]
33. Hassanein S, El Arab RA, Abdrbo A, et al. Artificial intelligence in nursing: an integrative review of clinical and operational impacts. Front Digit Health 2025;7:1552372. [doi: 10.3389/fdgth.2025.1552372] [Medline: 40124108]
34. Salahshoor S, Renbarger KM. AI-enhanced, digital psychosocial interventions for pediatric developmental disabilities: a scoping review. J Pediatr Nurs 2026;86:199-206. [doi: 10.1016/j.pedn.2025.11.016] [Medline: 41260047]

Abbreviations

AI: artificial intelligence
DL: deep learning
ICN: International Council of Nurses
ICU: intensive care unit
JBI: Joanna Briggs Institute
ML: machine learning
NP: nurse practitioner
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
SR: systematic review
WHO: World Health Organization

Edited by E Borycki, J Sarvestan; submitted 11.Apr.2026; peer-reviewed by E Abdulwahed, F Fernandes; revised version received 02.Jun.2026; accepted 15.Jun.2026; published 15.Jul.2026.

Please cite as:
Joo JY, Liu M, Cho Y, Cho H
Effectiveness of Artificial Intelligence–Based Nursing Interventions for Chronic Illness Care: Umbrella Review
JMIR Nursing 2026;9:e97905
URL: https://nursing.jmir.org/2026/1/e97905
doi:10.2196/97905

© Jee Young Joo, Megan Liu, Youngwoo Cho, Hyungbin Cho. Originally published in JMIR Nursing (https://nursing.jmir.org), 15.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Developing a Best Practice Guideline for Clinical Practice in a Digital Health Environment: Systematic Reviews Based on the Grading of Recommendations, Assessment, Development, and Evaluation Approach

Lauren Bailey[1], RN, BA(Hons), BScN, MPH; Lyndsay Howitt[1], RN, BScN, MPH; Nafsin Nizum[1], RN, BScN, MN; Christine Buchanan[1], RN, BNSc, MN; Maureen Charlebois[2], RN, BScN, MHA, CHE; Jennifer Yoon[3], RN, BScN, MSc; RNAO Expert Panel[4,5]

[1]International Affairs and Best Practice Guidelines Centre, Registered Nurses' Association of Ontario, 4211 Yonge Street #500, Toronto, ON, Canada

[2]Bayshore HealthCare, Toronto, ON, Canada

[3]Humber River Health, Toronto, ON, Canada

[4]See Acknowledgement

[5]

Corresponding Author:
Christine Buchanan, RN, BNSc, MN
International Affairs and Best Practice Guidelines Centre, Registered Nurses' Association of Ontario, 4211 Yonge Street #500, Toronto, ON, Canada

Abstract

Background: Digital health refers to the field of knowledge and practice associated with the development and use of digital technologies to improve clinical practice and health outcomes. Knowledge of digital health technology is becoming essential for all nurses and health providers.

Objective: This study aims to present the results of the systematic reviews that were used to inform the recommendations in a best practice guideline (BPG) following the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) approach. Reviews focused on digital health education for nurses and health providers, peer champion models, and the use of predictive analytics in digital health environments.

Methods: The BPG team, in collaboration with a panel of 17 experts, conducted 5 systematic reviews to address 5 recommendation questions. Systematic searches looked for relevant studies published in English from January 2017 to July 2022 from 10 databases. The GRADE approach was used to synthesize and evaluate the quality of evidence, ensuring the guideline aligned with international reporting standards.

Results: A total of 18 articles across 4 systematic reviews met the inclusion criteria. From these reviews, 4 corresponding recommendations were drafted for nurses and health providers. The strength of the recommendations was determined through discussion and consensus by the expert panel using the GRADE approach. Among all, 1 systematic review resulted in no recommendation due to insufficient evidence.

Conclusions: The BPG on digital health provides 4 evidence-based recommendations for nurses and health providers on how to incorporate digital health technologies into clinical practice. This BPG is intended to be used across all health care settings.

(JMIR Nursing 2026;9:e74942) doi:10.2196/74942

KEYWORDS

digital health; nursing; electronic health; health informatics; clinical guidelines

Introduction

Over the last decade, there has been an increased uptake of digital health technologies across global health care systems [1]. Digital health is a broad term that refers to the field of knowledge and practice associated with the development and use of digital technologies to improve health [2]. Digital health technologies refer to tools, systems, or devices that can generate, create, store, or process data, enabled through microprocesses that are programmed to perform specific functions [3]. Specifically in health care settings (or digital health environments), digital health technologies may encompass eHealth, mHealth, health informatics, artificial intelligence (AI), machine learning, big data, robotics, and advanced computing

XSL•FO
RenderX

sciences [2]. A digital health environment refers to any setting where health providers, informatics professionals, administrators, managers, and persons or families receiving care work in supportive teams to leverage digital tools, technologies, and services to optimize care delivery and empower and activate people to manage their health and wellness [4]. Nurses and health providers use a variety of digital health technologies in practice, including electronic health records, clinical decision support systems (CDSSs) that use predictive analytics, robotics, mobile apps, virtual care platforms, wearable devices, remote monitoring systems, smart home technologies, and others [4]. As nursing practice continues to evolve across all settings and sectors to incorporate these technologies, ongoing education is essential for nurses and health providers to deliver comprehensive clinical care [4,5].

Digital health technologies are advancing at a rapid pace; however, challenges remain in supporting nurses and health providers in using these technologies safely and effectively [6]. Educators and health systems leaders must work to evolve the understanding of novel nurse-patient interactions involving digital health technologies, alongside other core nursing topics [5]. Through further education and training, nurses will have a greater understanding of how both new and existing digital health technologies may impact clinical processes and communication patterns between patients, caregivers, and the interprofessional team [7]. Furthermore, nurses in clinical practice will require initial and ongoing professional development opportunities to aid in the use of digital health technologies [8,9]. Effective training will enable nurses to use these technologies both safely and effectively. Many of the good practice statements, recommendations, and resources within this best practice guideline (BPG) provide guidance on education for nurses and health providers to address this growing need.

The Registered Nurses' Association of Ontario (RNAO) published a new BPG entitled *Clinical Practice in a Digital Health Environment* in March 2024 [4]. The BPG was developed with an expert panel, which included 17 digital health experts representing diverse backgrounds including nursing, education, research, allied health, and people with lived experience. The purpose of the BPG is to provide evidence-based recommendations that foster nurses' ability to maintain, advance, and strengthen professional practice in the context of a digital health environment [4]. The guideline is intended for all nurses (registered nurses, nurse practitioners, and registered practical nurses), nursing students, as well as members of the interprofessional health care team, educators, administrators, executives, policymakers, researchers, and people with lived experience. Within the context of this BPG, people with lived experience refer to patients and family within health systems wherein digital health is used.

The aim of this paper is to describe the BPG development process and the results from 4 systematic reviews that were used to inform the recommendations in the BPG, following the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) approach [10]. Additionally, this paper will reflect on the health equity considerations, research gaps, and limitations noted during guideline development, related to the integration of digital health technologies in clinical practice.

Methods

Development Approach

RNAO's BPG development team used the GRADE approach to develop this guideline, which is in line with international reporting standards [10]. GRADE is a transparent and structured process to evaluate the certainty of a body of evidence from systematic reviews in order to develop sound, evidence-based recommendations in guidelines [10]. The systematic reviews were conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [11] (Checklist 1). The following section will describe how the purpose and scope of the guideline were determined, the 5 systematic reviews that were conducted, and the resulting 4 recommendations that were drafted following completion of the systematic reviews.

Scoping the Best Practice Guideline

To determine the purpose and scope of this BPG, the guideline development team conducted an environmental scan on existing clinical guidelines on this topic and appraised those guidelines. Two guideline development methodologists (GDMs) determined inclusion or exclusion criteria and searched an established list of websites for guidelines and other relevant content (eg, quality standards) published between January 2016 and March 2021 (Multimedia Appendix 1). Expert panel members were also asked to suggest additional guidelines for review. Guidelines were reviewed for content, applicability to health provider scope of practice, accessibility, and quality. Each GDM individually evaluated guideline quality using the Appraisal of Guidelines for Research and Evaluation II instrument [12]. Through this process, it was determined that no guidelines had been developed addressing evidence-based recommendations in this unique and growing area, especially as it relates to nurses and clinical practice.

The team also completed a preliminary literature review to examine available evidence on digital health for nurses, including how digital health technologies are being integrated into the nursing process; how digital health technologies are facilitators and/or barriers for nurses when maintaining and advancing professional practice; and what outcomes are used to measure the impact of using digital health technologies in nursing practice. Two databases were searched for literature (CINAHL and MEDLINE) between January 2016 and May 2021. Screening for eligibility was conducted independently by 2 GDMs with conflicts resolved through consensus. Data extraction was completed for the included studies on a customized Microsoft Excel sheet developed by the GDMs. Elements of data extraction (such as study setting, intervention, and outcomes) were determined by the GDMs. An analysis of themes across the studies was synthesized, and the themes, interventions, and outcomes were later presented to the expert panel.

GDMs also conducted 22 key informant interviews and 2 discussion groups with diverse experts in the field. Key

XSL•FO
RenderX

informants included people with lived experience, direct care health and social service providers, and researchers selected based on their knowledge and expertise related to the BPG topic. Snowball sampling was also used to recruit key informants. See Textbox 1 for a description of the questions asked during the interviews. For the discussion groups, 3 sessions were convened with a total of 18 nursing students, clinical informatics nurses, and frontline nurses to understand the needs of nurses within digital health environments. GDMs used inductive qualitative content analysis to analyze data collected from key informant interviews and discussion groups, and this information was also presented to the expert panel.

Textbox 1. Key informant interview questions.

- How can digital health technologies impact the quality of care a person receives?
- How can digital health technologies promote or hinder the therapeutic nurse-client relationship?
- In what ways can a digital health environment enhance or hinder patient care delivery for underserved populations?
- What skills or competencies do nurses require in order to maintain professional practice in a digital health environment?
- What skills and competencies do nurse leaders require in order to support the interface between nursing clinical practice and digital health technology?
- What challenges do nurses face when working in a digital health environment?
- What challenges or struggles do you face in your current practice related to the use or implementation of digital health technologies?
- What challenges do nurses face when trying to engage in the design, development, and evaluation of digital health environments?
- What policies or practices can help nurses maintain professional practice in a digital health environment?
- What outcomes should we explore in the literature to measure the impact of using digital health technologies in clinical nursing practice?
- What should the scope of this guideline be?
- What should this best practice guideline (BPG) address in order to be most useful in practice for nurses and people receiving care?
- Are there any last thoughts on what is important for us to consider when starting the development of this BPG?

Identifying Priority Recommendation Questions and Outcomes

The BPG development team assembled a panel of 17 experts, including 2 cochairs, from nursing practice, research, education, and policy, as well as other members of the interprofessional team, and people with lived experience representing a range of sectors and practice areas. The BPG was supported by 2 cochairs with relevant clinical and research experience, one of whom was a doctorate-prepared registered nurse, and the other cochair led the pan-Canadian Electronic Health Record Clinical Engagement Strategy for 6 years at Canada Health Infoway. The expert panel also included representatives from different geographical areas, including rural, suburban, and urban. From July to December 2021, 4 panel meetings were held to determine the BPG's purpose, scope, and research questions that informed the systematic reviews. During the first orientation meeting, the expert panel was introduced to RNAO's BPG program, the systematic review process, and the GRADE approach. Additional electronic materials were also sent to the panel to familiarize them with the BPG development process and the GRADE approach. Declarations of conflicts of interest that might be construed as constituting a perceived and/or actual conflict were made by all members of the expert panel prior to their participation in guideline development work, and on an ongoing basis.

During the initial phase of the guideline development process, the expert panel prioritized 4 research questions and corresponding outcomes deemed most important to this topic. An amendment to the PROSPERO registration was made following these initial meetings, once the panel determined through email correspondence that a fifth research question should be added. Textbox 2 displays the final recommendation questions and outcomes that informed focused research questions for the systematic reviews.

XSL•FO
RenderX

Textbox 2. Recommendation questions and outcomes in the clinical practice in a digital health environment best practice guideline.

Recommendation question 1: Should practical (eg, hands-on) professional development education focused on the use of digital health technologies within an organization be recommended or not for all nurses?

Outcomes: nurse competence (with using technology), nurse acceptance of technology, nurse-sensitive outcomes (eg, falls, pressure injuries, and pain), nurse involvement in the technology life cycle, nurse confidence (with using technology), and nurse-person therapeutic relationship.

Recommendation question 2: Should education about relational care and interpersonal communication skills be recommended or not for nurses practicing in virtual care settings and in-person digital health environments?

Outcomes: person or caregiver or family experience or satisfaction, nurse competence (with using technology), nurse confidence (with using technology), nurse-person therapeutic relationship, and person or caregiver or family involvement and engagement in care.

Recommendation question 3: Should the implementation of interdisciplinary peer champion models in health service organizations be recommended or not to facilitate education for health providers on the use of digital health technologies?

Outcomes: health provider competence (with using technology), health provider adoption of technology, health provider confidence (with using technology), health provider sensitive outcomes (eg, pressure injuries and pain), and sustainability of education (ie, knowledge and skills retention).

Recommendation question 4: Should the use of predictive analytics software or systems (eg, command centers and risk assessment software tools) for nurses providing care in all practice settings be recommended or not to inform clinical decision-making and improve clinical outcomes?

Outcomes: proactive or anticipatory care, critical incidents, failure to rescue, consistent application of evidence-based practice, and nurse-sensitive outcomes (eg, falls, pressure injuries, and pain).

Recommendation question 5: Should a distributive model (vs no distributive model or any other type of change management model) be recommended to integrate digital health competencies into the professional practice roles and responsibilities of nurses at all levels within an organization?

Outcomes: nurse competence (with using technology), nurse engagement (with using, developing, acquiring, and participating in education about the technology), nurse confidence (with using technology), person or caregiver or family experience or satisfaction, and nurses being able to define what their role is.

Systematic Retrieval of the Evidence

The systematic reviews for the guideline were registered with PROSPERO in 2022 (CRD42022321580). Upon consultation with the expert panel, 4 amendments were made to the original PROSPERO registered protocol. These included: (1) adding an additional database to search (IEEE Xplore) in April 2022, (2) adding an additional systematic review question (December 2022), (3) conducting indirect evidence searches (January 2023), and (4) publishing the final version of the guideline online (May 2024). All other systematic review methods followed the protocol outlined in the original PROSPERO registration.

Five separate systematic review search strategies were developed and run by an external health sciences librarian from the University Health Network after consulting with 2 GDMs (CB and LH). The systematic searches included peer-reviewed studies of any study design (eg, quantitative, qualitative, mixed methods, and systematic reviews) published in English from January 2017 to July 2022. The following databases were searched: MEDLINE, MEDLINE Epub Ahead of Print and In-Process, Embase, Emcare Nursing, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, APA PsychInfo, CINAHL, and IEEE Xplore. Expert panel members were also asked to review their personal libraries for key studies not found through the above search strategies. For more details and the full search strategy used for each systematic review, please refer to Multimedia Appendix 2.

After conducting the initial searches, it was decided to look for further indirect evidence to support each question. Direct evidence comes from research that directly compares the interventions of interest when applied to the populations of interest and measures outcomes important to patients [13]. Evidence can be indirect if the population differs, the intervention differs, or outcomes differ from those of original interest [13]. The health science librarian conducted additional indirect evidence searches from January 2023 to March 2023 for systematic reviews published in English. The BPG team recognizes that direct evidence allows for more confidence in the results; however, in the absence of direct evidence, GRADE notes that indirect evidence can be used and downgraded accordingly [10,13]. The broader populations and interventions searched were considered sufficiently direct by the expert panel and in line with the original methodology. To ensure the most up-to-date evidence was included in the BPG, an update search was also conducted in English between January 2023 to January 2024 for recommendation questions 1 to 4. However, an update search for question 5 was not completed since a recommendation was not drafted for this area. For the full search strategies, see Multimedia Appendix 2.

Eligibility Criteria

All search results from the librarian were uploaded into DistillerSR software (DistillerSR Inc). All steps of the systematic review process were completed by 2 GDMs (CB and LH for the initial search and CB and LB for the update search). Two GDMs independently completed title and abstract screening using standardized screening guides developed by the GDMs. Screening guides were reviewed by senior members of the RNAO team prior to use. Studies included at this stage had the full text reviewed independently by both GDMs. Final inclusion was deemed appropriate if studies answered the research question, included prioritized outcomes, were published in English, and were accessible for retrieval. See Textbox 3 for inclusion and exclusion criteria, and Multimedia Appendix 2 for further details. Disagreements were settled by consensus. For the initial systematic search, any study design was eligible to be included. For the updated indirect systematic searches,

XSL•FO
RenderX

study designs were limited to systematic reviews and meta-analyses.

Textbox 3. Inclusion and exclusion criteria.

Inclusion criteria:

- A primary focus on the interventions of interest and the prioritized outcomes per research question
- A focus on digital health technologies
- Applicable to nurses or health providers providing care in all practice settings (including registered nurses, registered practical nurses, nursing students, and nurse practitioners)
- Applicable to all health or social service organizations, or academic institutions
- Published after January 2017
- Published in English
- Accessible for retrieval
- Conducted in any geographic region
- Peer-reviewed literature
- Any study design (eg, quantitative, qualitative, mixed methods, and systematic reviews), but when conducting the indirect searches, only systematic reviews and meta-analyses were included.

Exclusion criteria:

- Topic NOT related to the interventions or prioritized outcomes per research question
- Dissertations, commentaries, narratives, discussion papers, case studies, expert reports, consensus documents, and studies with no specific methodology
- Studies not published in English
- Unpublished literature (eg, gray literature)
- Studies published prior to 2017

Data Extraction and Quality Appraisal

Data extraction was completed on the included studies for each research question. The included studies were divided between GDMs and each reviewer independently extracted details from their assigned studies using standardized Excel sheets that were developed by the RNAO team (Multimedia Appendix 3). Each Excel sheet had a designated outcome for which study details were recorded. Details such as the setting, intervention and control description, the outcome and how it was measured, and study results were recorded by 1 GDM. Any harms (such as adverse effects), information on values, preferences, and health equity were also recorded. The second GDM independently reviewed the extracted data for accuracy. Quality appraisal of each article was completed independently by each GDM. The Cochrane Risk of Bias 2.0 tool [14] was used to appraise randomized controlled trials (RCTs), the risk of bias in nonrandomized studies—of interventions (ROBINS-I) tool [15] was used to appraise nonrandomized studies, and the risk of bias in systematic reviews (ROBIS) tool [16] was used to appraise systematic reviews. If a systematic review received a low risk of bias score using the ROBIS tool, and the review's authors completed a risk of bias assessment within the paper, those assessments were also considered when conducting the GRADE consensus. After quality appraisal was completed by both GDMs, GRADE consensus was completed to assess the certainty of evidence for each outcome for each research question. GRADE uses five categories to rate the certainty of evidence as high, moderate, low, and very low by examining (1) risk of bias, (2) inconsistency, (3) imprecision, (4) indirectness, and (5) publication bias [10]. After the 5 categories had been graded, a certainty of evidence was determined for each of the 4 drafted recommendations corresponding to the research questions.

Drafting Recommendations in the BPG

As per the GRADE methodology, the GDMs created an evidence profile (EP) and evidence to decision (EtD) framework for each recommendation [4,10]. The EP outlined details regarding the certainty of evidence across outcomes and the GRADE domains (Multimedia Appendices 4-7). The EtD frameworks provided a narrative summary of the evidence for draft recommendations, described the certainty of evidence, and provided details around values and preferences regarding the intervention, as well as health equity considerations found in the systematic reviews. Expert panel members were provided with the EPs and EtD frameworks to review prior to 3 (virtual) half-day meetings to determine the direction (ie, a recommendation for or against an intervention) and the strength (ie, strong or conditional) of the BPG's recommendations. A conditional recommendation is one for which the desirable effects probably outweigh the undesirable effects, and there is a need to consider more carefully than usual the individual's circumstances, values, and preferences [10]. If there was insufficient direct or indirect evidence to develop a recommendation, the expert panel also had the option not to proceed with a recommendation. The expert panel determined that current evidence was insufficient to assess the certainty of

XSL•FO
RenderX

effects of a distributive model (recommendation question 5) compared to other types of change management models to integrate digital health competencies into the professional practice roles and responsibilities of nurses within an organization; thus, no recommendation was made.

The recommendations and draft BPG also underwent several rounds of internal and external review prior to publication [17]. External reviewers for RNAO BPGs are identified through a public call issued on the RNAO website [17]. For this BPG, the written external review process was completed between September 14, 2023, and October 23, 2023. External reviewers with diverse perspectives, such as nurses and health providers, administrators, researchers, educators, nursing students, and people with lived experience, provided direct feedback.

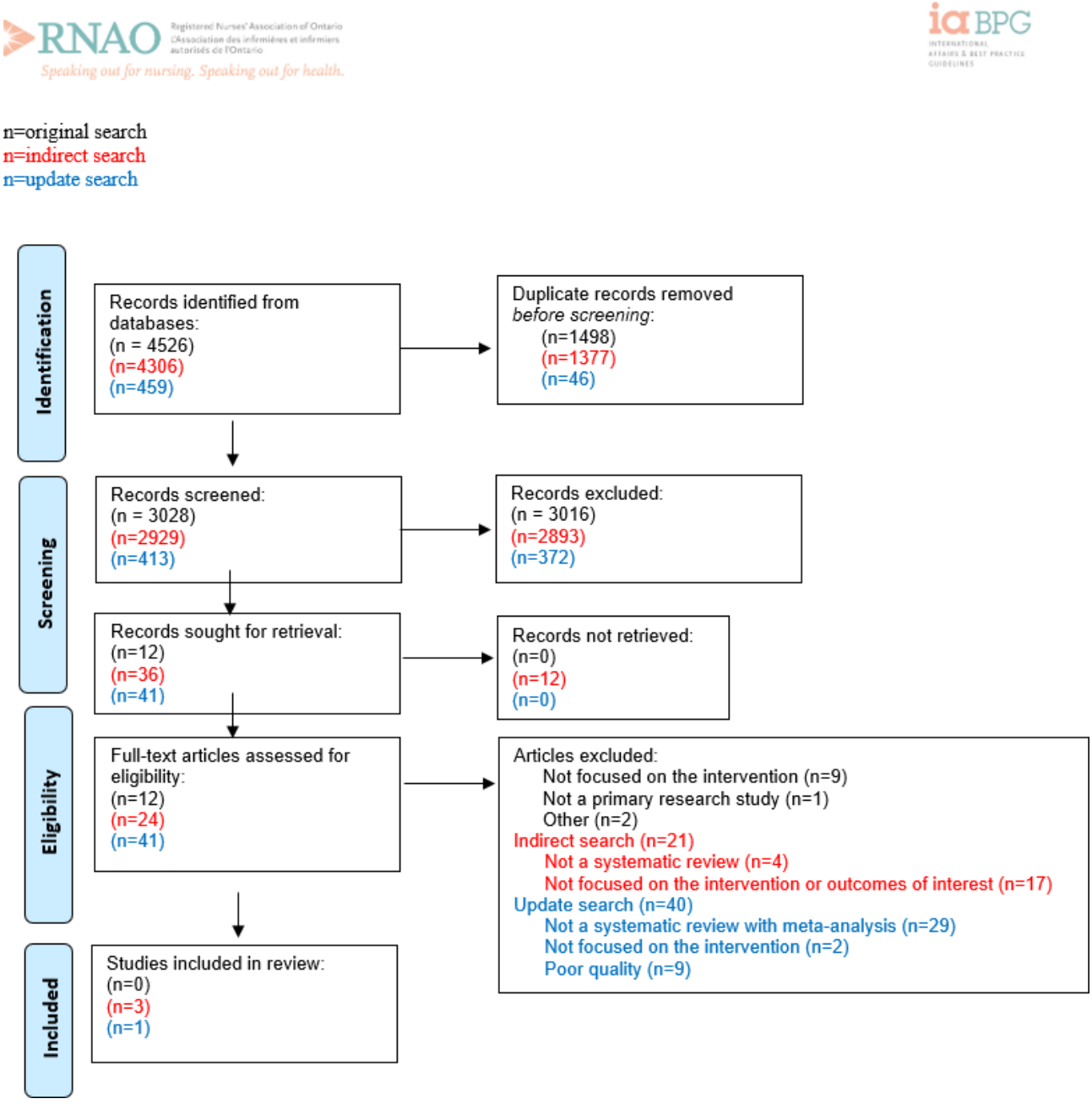
Results

Summary of Results

For PRISMA flow diagrams, see Figures 1-5. Two reviewers screened over 22,500 articles for the 5 original research questions. After screening, the 2 GDMs reviewed 253 full-text articles for relevance to the research questions and outcomes, and 18 articles met the requirements to inform the recommendations. It was determined through consultation with the expert panel that question 5 did not have enough evidence to support the recommendation, so a recommendation was not developed. Thus, 4 recommendations were drafted (one per each of the corresponding systematic reviews), and the strength of the recommendations was determined through discussion and consensus by the expert panel, based on the available evidence (Table 1).

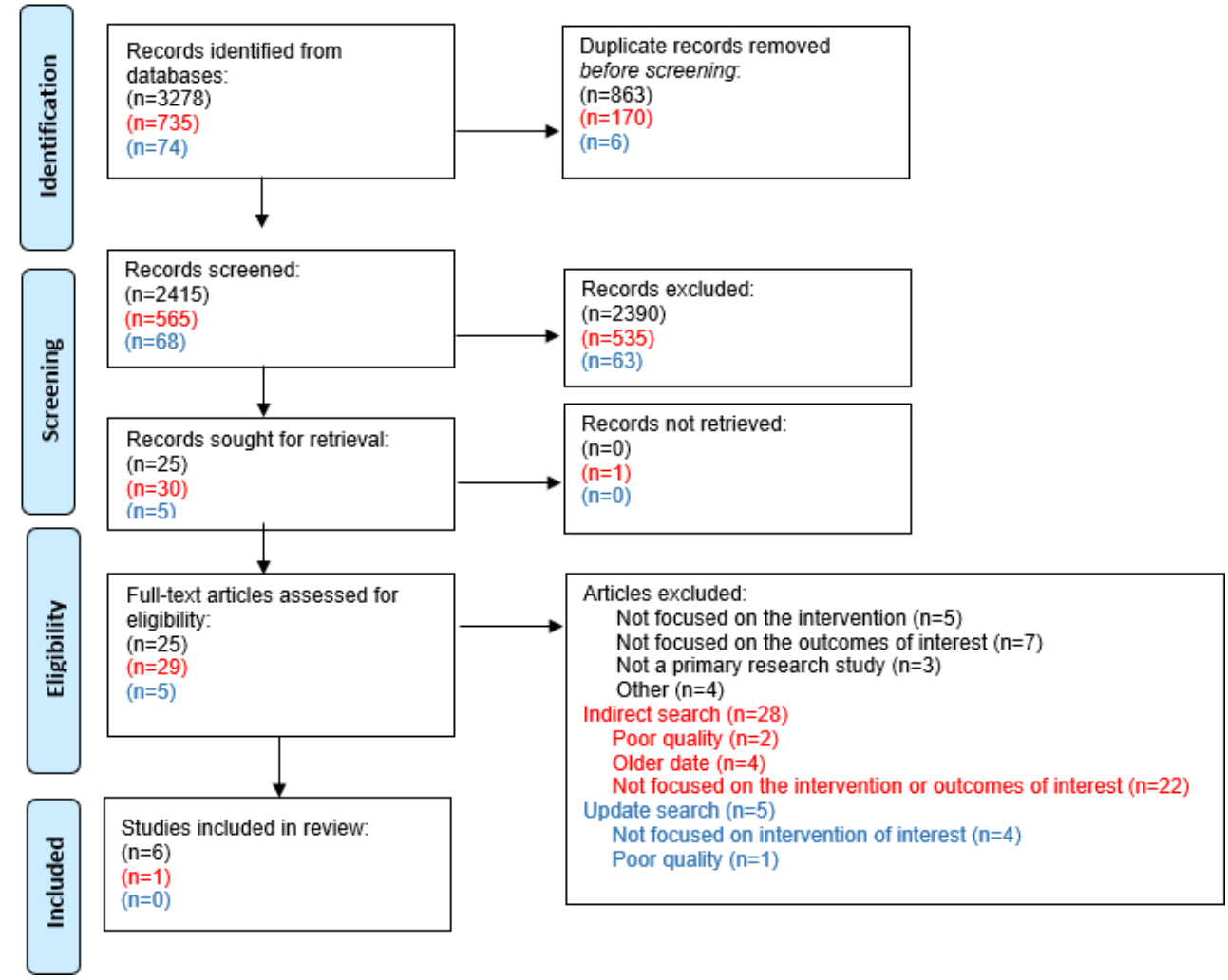
XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for recommendation question 1: "Should practical (eg, hands-on) professional development education be focused on the use of digital health technologies within an organization be recommended or not for all nurses?" Adapted from Page MJ et al [11].



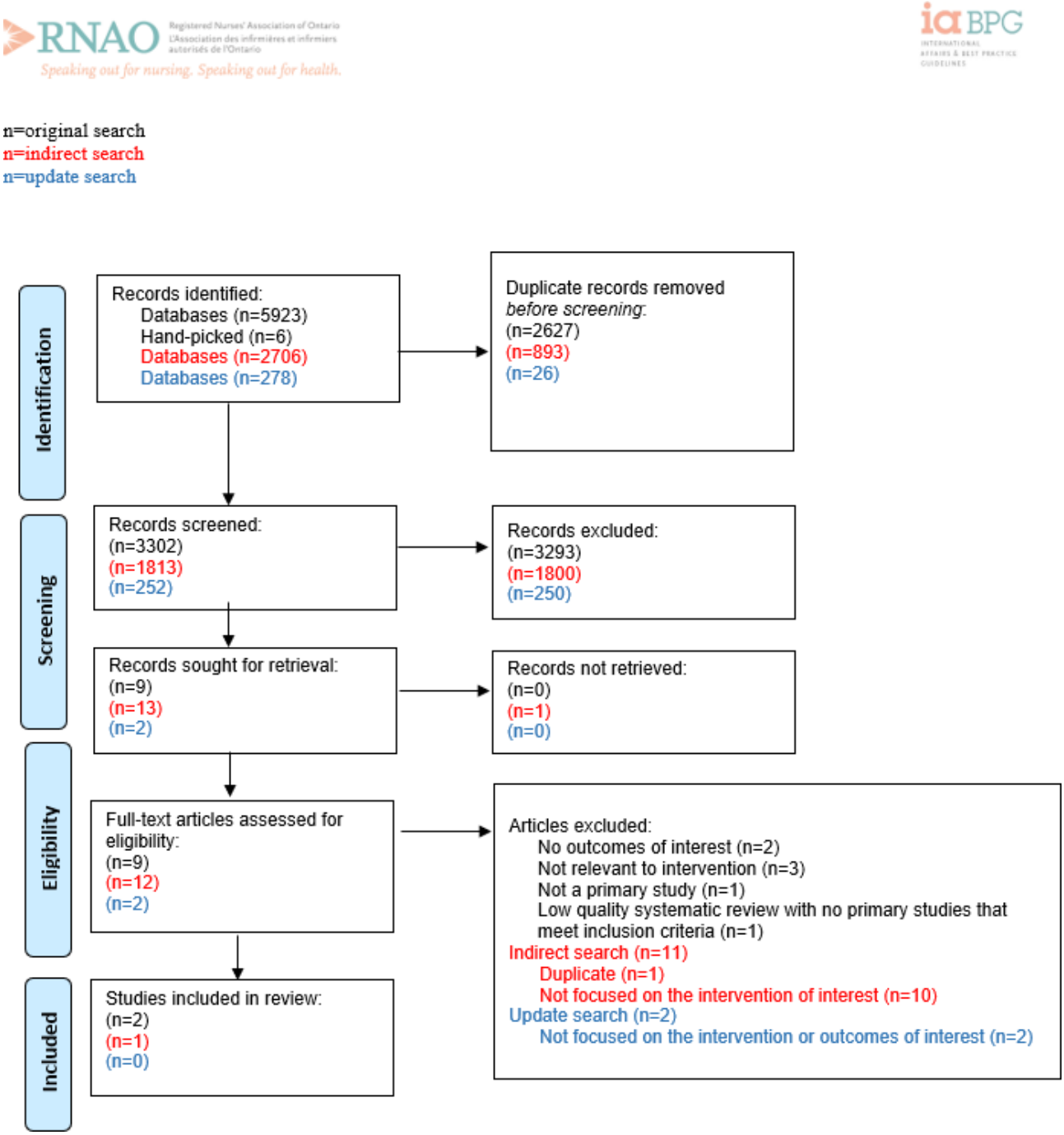
XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for recommendation question 2: "Should education about relational care and interpersonal communication skills be recommended or not for nurses practicing in virtual care settings and in-person digital health environments?" Adapted from Page MJ et al [11].



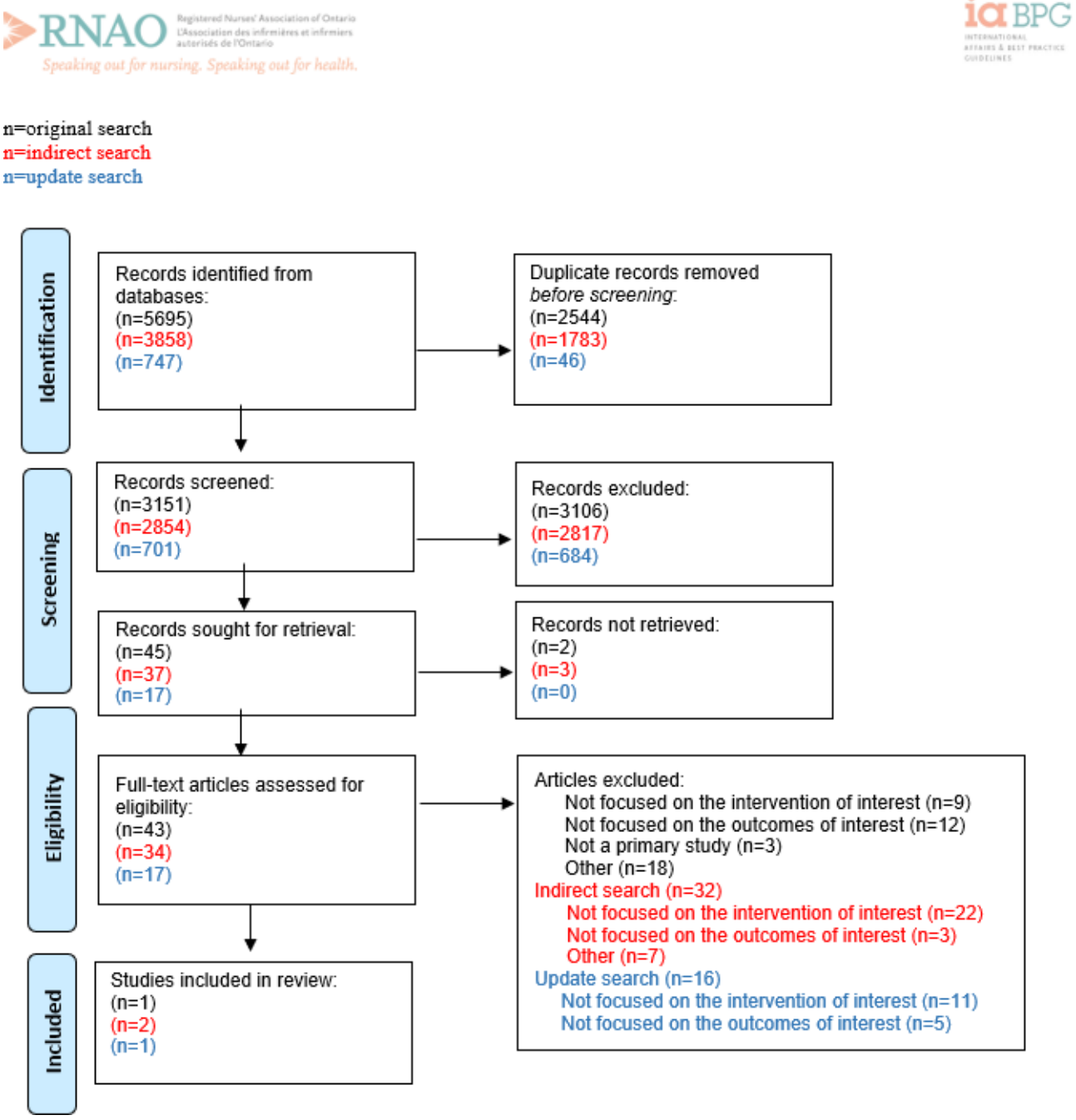
XSL•FO
RenderX

Figure 3. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for recommendation question 3: "Should the implementation of interdisciplinary peer champion models in health service organizations be recommended or not to facilitate education for health providers on the use of digital health technologies?" Adapted from Page MJ et al [11].



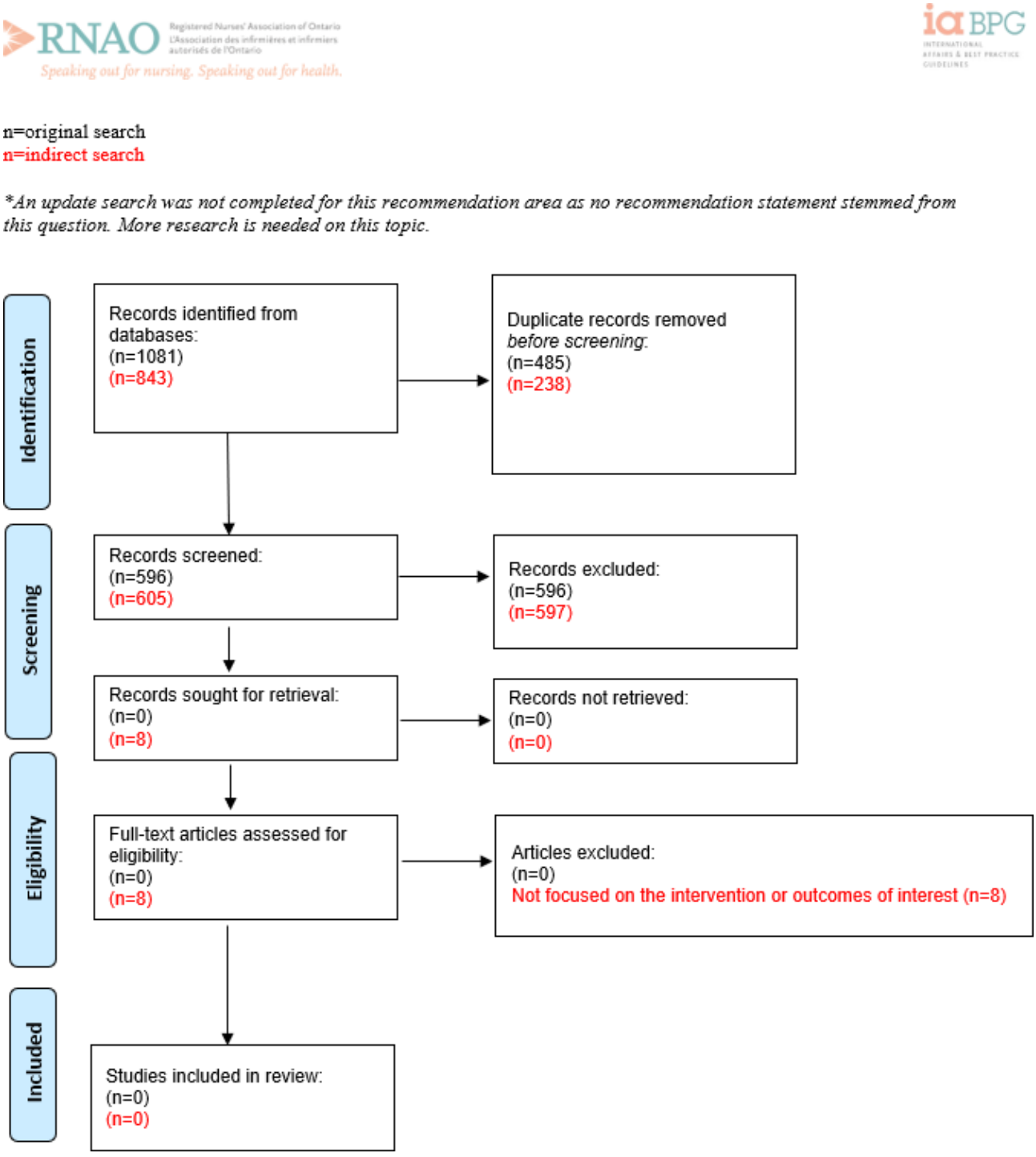
XSL•FO
RenderX

Figure 4. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for recommendation question 4: "Should the use of predictive analytics software or systems (eg, command centers and risk assessment software tools) for nurses providing care in all practice settings be recommended or not to inform clinical decision-making and improve clinical outcomes?" Adapted from Page MJ et al [11].



XSL•FO
RenderX

Figure 5. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for recommendation question 5: "Should a distributive model (vs no distributive model or any other type of change management model) be recommended to integrate digital health competencies into the professional practice roles and responsibilities of nurses at all levels within an organization?" Adapted from Page MJ et al [11].



XSL•FO
RenderX

Table . Summary of recommendations in the best practice guideline.

Recommendation	Strength of recommendation
Recommendation 1.0: the expert panel suggests that health service and academic organizations provide ongoing education to nurses and health providers that includes hands-on training for the use of digital health technologies.	Conditional
Recommendation 2.0: the expert panel suggests that health service and academic organizations provide ongoing education to nurses and health providers that focuses on interpersonal communication skills when using digital health technologies.	Conditional
Recommendation 3.0: the expert panel suggests that health service organizations implement interdisciplinary peer champion models to facilitate education for nurses and health providers on the use of digital health technologies.	Conditional
Recommendation 4.0: the expert panel suggests that health service organizations implement CDSS[a] or early warning systems that use artificial intelligence–driven predictive analytics to support nurses' and health providers' clinical decision-making.	Conditional

[a]CDSS: clinical decision support system.

Recommendation 1.0: The Expert Panel Suggests That Health-Service and Academic Organizations Provide Ongoing Education to Nurses and Health Providers That Includes Hands-on Training for the Use of Digital Health Technologies

Practical or hands-on education refers to deliberate practice, hands-on training, or simulation training (ie, more than just viewing e-learning modules) [4]. The intervention of interest examined whether practical or hands-on education for professional development was more effective than standard education (ie, no-hands-on education component) when training nurses and health providers on the use of digital health technologies [4]. Four meta-analyses informed this recommendation [18-21]. The 4 meta-analyses were assessed for risk of bias using the ROBIS tool, and each one had a low risk of bias [18-21]. Studies included in the meta-analyses were assessed by the authors of the meta-analyses, and they used the Cochrane risk-of-bias 2.0 tool for RCTs, the ROBINS-I tool for nonrandomized studies, and the National Institute for Health and Care Excellence quality appraisal checklist [18-21]. Nine studies within the meta-analyses had a critical risk of bias, 18 studies had high risk of bias, 4 studies had unclear risk of bias, and 1 study had low risk of bias [18-21]. There were concerns noted around allocation concealment, blinding, incomplete outcome data, missing outcome data, selection of the reported results, confounding, allocation concealment, and selection of participants [18-21].

Examples of practical or hands-on professional development education discussed in the studies included nurses practicing using electronic health records while being supervised in a computer lab, and hands-on training for using virtual care platforms [18-21]. For more details on the study designs, the risk of bias assessments, how the interventions were delivered, and outcome measures, refer to the GRADE EP found in Multimedia Appendix 4.

The results of the systematic review suggest that hands-on education for nurses and health providers may improve nurses' competence and confidence, and the nurse-person therapeutic relationship (while the technology is used with the person receiving care). The expert panel determined that the overall evidence was of very low certainty due to the risk of bias in the primary studies, indirectness in the outcomes, inconsistency in the results, and imprecision due to small sample sizes [4]. Based on this certainty of evidence, the panel determined the strength of the recommendation to be conditional.

Recommendation 2.0: The Expert Panel Suggests That Health-Service and Academic Organizations Provide Ongoing Education to Nurses and Health Providers That Focuses on Interpersonal Communication Skills When Using Digital Health Technologies

Interpersonal communication describes the communication between a nurse or health provider and a person receiving care. It includes both verbal and nonverbal communication, as well as leading and listening skills that enable a person to interact positively with others in an effective manner [4,22]. The types of education varied across the studies and included didactic and simulation-based education (eg, simulated patients) to improve medical students' interpersonal communication during consultations; training on incorporating computers or electronic health records into nurse-patient encounters; and education on telehealth communication strategies (eg, phone and video consults) [23-29]. Most studies examined focused on medical students [23-27,29], and 1 study focused on nursing students [28].

Seven studies informed this recommendation, including 1 systematic review, 5 additional nonrandomized studies, and 1 mixed methods study [23-29]. The review was assessed using the ROBIS tool and had a low risk of bias [23]. Studies included in the review were assessed by the review authors in accordance with the Cochrane Handbook for Systematic Reviews of Interventions; none were deemed as having a high risk of bias overall [23]. Nonrandomized studies and the mixed-methods

XSL•FO
RenderX

study were assessed using the ROBINS-I tool, and there was a critical risk of bias related to confounding variables, deviations from the intended interventions, missing data, measurement of outcomes, and selection of the reported results [24-29].

The 7 studies illustrated that there may be benefits when health service and academic organizations provide nurses and other health providers with education about the importance of interpersonal communication when using digital health technologies [4,23-29]. Benefits may include improved person, caregiver, or family experience or satisfaction with care, and increased competence and confidence among nurses; however, the overall certainty of the evidence using the GRADE methodology was very low, due to risk of bias in the seven studies, few participants, and inconsistency in results [4]. Based on these factors, the expert panel determined the strength of the recommendation to be conditional. For more details on the study designs, risk of bias assessments, how the interventions were delivered, and outcome measures, refer to the GRADE EP in Multimedia Appendix 5.

Recommendation 3.0: The Expert Panel Suggests That Health Service Organizations Implement Interdisciplinary Peer Champion Models to Facilitate Education for Nurses and Health Providers on the Use of Digital Health Technologies

Interdisciplinary peer champions refer to super-users or champions that are nurses or other members of the interdisciplinary health care team with expertise and additional training in digital health [4]. These individuals function as a resource for other staff, helping to answer questions and teach staff about new technology during implementation. Peer champions can also help identify gaps in the technology or its implementation in practice. This recommendation examined the effects of organizations implementing peer champion models to facilitate education for staff about digital health technologies.

One systematic review of 6 RCTs and 2 nonrandomized single-arm studies informed this recommendation [30-32]. The review was assessed using the ROBIS tool and had a low risk of bias [30]. Studies included in the review were assessed by the review authors using the Cochrane risk-of-bias tool for RCTs; 5 studies had a high risk of bias and 1 study had an unclear risk of bias [30]. The nonrandomized studies were assessed using the ROBINS-I tool, and there was a critical risk of bias related to confounding variables, missing data, measurement of the outcomes, and selection of the reported results [31,32].

The use of peer champions in health service organizations may increase health providers' adoption of technology and health provider competence [4]. The overall certainty of evidence was low due to a serious risk of bias in the individual studies and a low number of participants [4]. Based on the available evidence, the expert panel determined the recommendation to be conditional. For more details on the study designs, risk of bias assessments, how the interventions were delivered, and outcome measures, refer to the GRADE EP in Multimedia Appendix 6.

Recommendation 4.0: The Expert Panel Suggests That Health Service Organizations Implement Clinical Decision Support Systems or Early Warning Systems That Use AI-Driven Predictive Analytics to Support Nurses' and Health Providers' Clinical Decision-Making

CDSS or early warning systems refer to software found in risk assessment software tools, early warning systems, command centers, and other software systems that use AI machine learning algorithms to interpret data independently [4]. The recommendation question examined whether adding these systems benefits clinical decision-making for nurses and other health providers.

One systematic review of RCTs, 1 nonrandomized single-arm study, and 2 systematic reviews of nonrandomized studies informed this recommendation [33-36]. Included reviews were assessed using the ROBIS tool and had a low risk of bias [33,35,36]. Studies included in 1 review were assessed by the review authors using the Critical Appraisal Skills Programme checklist for RCTs; 2 studies had a low risk of bias and 1 study had a high risk of bias [33]. Concerns were noted around the lack of details describing the methods, and the lack of blinding [33]. Studies included in another review were assessed by the review authors using the Prediction model Risk Of Bias Assessment Tool; all 10 studies had high or unclear risk of bias [36]. The nonrandomized study was assessed using the ROBINS-I tool and had a critical risk of bias due to lack of control for confounding variables, deviations from the intended intervention, and selection of the reported results [34]. Studies in the final review were assessed by the review authors using the ROBINS-I tool; all 5 included studies had a critical risk of bias [35]. Concerns were noted around confounding, selection of participants, missing data, measurement of outcomes, and selection in reported results [35].

There may be benefits when implementing CDSS or early warning systems that use AI-driven predictive analytics to inform nurses' clinical decision-making, such as improved proactive or anticipatory care, decreased failure to rescue, consistent application of evidence-based practice, and improved nurse-sensitive outcomes [4]. The overall certainty of evidence was low due to risk of bias and few participants [4]. As evidence is still emerging on this topic and the results were mixed, the expert panel determined the strength of the recommendation to be conditional. For more detail on the study designs, risk of bias assessments, how the interventions were delivered, and outcome measures, refer to the GRADE EP in Multimedia Appendix 7.

Discussion

Digital Health Considerations

In 2019, the World Health Organization released a global strategy on digital health acknowledging the vital role digital health plays in planning and providing health services [2]. As digital health technologies become increasingly integrated into health care, nurses need leadership and guidance to safely and effectively use technology in practice. RNAO's BPG provides evidence-based recommendations to foster nurses' ability to maintain, advance, and strengthen professional practice in the context of a digital health environment [4]. The guideline's

XSL•FO
RenderX

recommendations focus on (1) hands-on education related to the use of digital health technologies, (2) education about interpersonal communication skills when using digital health technologies, (3) using interdisciplinary peer-champion models to provide education about digital health technologies, and (4) implementing CDSS that uses AI to support but not replace clinical decision-making. While not discussed in this article, additional good practice statements are also provided in the guideline [4].

While digital health has the potential to enhance the quality of care and address key health system challenges, the importance of considering the digital determinants of health, including digital literacy and the digital divide, to ensure equitable delivery of care must be considered. Digital literacy refers to a person's ability to effectively interact with digital technology, using skills required to find, understand, appraise, and apply health information specifically from electronic sources [37]. The digital divide refers to the gap between those who have access to digital technologies, including the internet, accessible health websites and portals, versus those who do not [38]. The World Health Organization's global strategy on digital health notes that digital technologies are to be adaptable to different countries and contexts to help address key health system challenges, while incorporating equity, diversity, and inclusion principles [2]. Unfortunately, the use of certain digital health technologies such as CDSS that use AI may be difficult to implement in less affluent health care systems due to the digital divide [39]. The effectiveness of implementing CDSS that use AI to detect changes in a patient's condition is also dependent on having staff who respond appropriately to these digital tools as well as nursing leadership to continuously oversee the refinement of CDSS and algorithms as needed. As outlined by Richardson et al [40] in their framework for digital health equity, there are several domains of equity including biological, behavioral, physical/built environment, sociocultural environment, and the health care system. The framework can help support the work of digital health technology developers to think about and incorporate principles of digital health equity from the very beginning of the technology development process [4,40]. The framework is also important for end-users, researchers, and health systems leaders, as digital health transformation requires health leaders at all levels to understand how the digital determinants impact health equity [4,40].

In addition to considering the digital determinants of health, when discussing the use of digital health technology with a person receiving care and/or their family, nurses must consider: their preferences and goals; capability and motivation for using technology; how the technology fits into their current care routines; and any costs associated with using the technology [4,41]. Digital health technologies have the potential to enhance a person's experience of the care they receive [5]; however, nurses must consider a person's values and preferences for using technology and ensure that using the technology does not negatively impact or compromise the nurse-patient therapeutic relationship [2,5].

Implementation and Evaluation Considerations

Evidence-based guidelines are effective when there are tools and strategies in place to facilitate their implementation into practice [12]. RNAO uses an integrated approach to ensure that guidelines are both trustworthy and applicable in real-world settings [17]. This BPG includes several tools to support its implementation, including implementation tips, supporting resources, appendices related to the recommendations, and good practice statements. The BPG also directs readers to RNAO's Leading Change Toolkit, which can be used to guide change initiatives, including the implementation of BPGs [42]. RNAO has a network of best practice champions who are the change agents that aid in the implementation of the guidelines, and Best Practice Spotlight Organizations® (BPSO®) internationally from over 13 different countries that partner with RNAO to systematically implement and evaluate RNAO's BPGs [17].

Finally, a monitoring and evaluation table outlines structure, process, and outcome indicators that health service organizations can use to monitor the impact of BPG implementation. Ongoing evaluation is crucial to support the uptake and impact of BPGs on person, organizational, and health systems outcomes [17]. RNAO houses 2 data systems to support BPSOs to monitor and evaluate BPGs: MyBPSO and Nursing Quality Indicators for Reporting and Evaluation® [17]. These 2 data systems are used by BPSOs to report evaluation and monitoring data. As of November 2025, implementation of this BPG has begun in BPSOs, and 1 large Canadian community hospital BPSO has demonstrated 99% (555/562) of nursing staff were compliant with orientation to technologies. Evaluation has also indicated that 84% (474/562) of nurses at this hospital reported comfort with hospital-based technologies to deliver care, and 30% (20,188/72,342) of patients enrolled in a digital patient portal over an 18-month period. Evaluation and monitoring of outcomes is ongoing, and it is anticipated that in the coming years, more BPSOs will implement this valuable BPG.

Future Research Considerations

The expert panel noted that although rigorous RCTs are needed, more exploration including qualitative research is also needed in the area of digital health as it pertains to nursing and clinical practice. For example, studies that examine the efficacy, accuracy, and generalizability of AI-driven predictive analytics, and qualitative studies exploring how nurses and health providers adapt their communication skills in digital health environments. National and international research institutes focused specifically on advancing digital health technologies and integrating digital health practices into clinical care for nurses and health providers would also be beneficial.

Limitations

A few limitations were noted by the expert panel and GDMs during the development of the BPG. First, research in digital health that is specific to nurses and clinical practice is an emerging area. As research is yet to be well established, most evidence for the prioritized research questions was of low or very low certainty; thus, all the recommendations contained in the BPG were deemed conditional. There were few well-designed RCTs, and many of the nonrandomized studies

XSL•FO
RenderX

had a high risk of bias, small sample sizes, and inconsistent results. In addition, due to the paucity of research evidence focused on nurses and digital health, the expert panel considered indirect evidence. According to the GRADE methodology, directness is assessed based on the relevance to the target population, intervention, and outcomes of interest [10]. Although GRADE methods allow for the use of indirect evidence, the reliance on indirect evidence due to insufficient direct evidence is a limitation in this BPG, recognizing that indirect evidence may introduce potential biases or uncertainties. The absence of research and use of indirect evidence is noted in the BPG as research gaps, stressing areas for further exploration.

Despite these limitations, expert panel members and additional external reviewers noted the need for guidance on this topic and the importance of publishing this guideline. Conditional recommendations are not to be seen as less important or less trustworthy; they simply imply that there is a need to consider more carefully than usual the individual person or family's circumstances, preferences, and values [10]. When implementing conditional recommendations, health providers need to allocate more time to shared decision-making and comprehensively explain the potential benefits and harms to people and their families [10]. It is becoming increasingly common for clinical guidelines to only include conditional recommendations, as guideline panels and developers recognize the importance of thinking holistically [43,44]. As evidenced by the COVID-19 pandemic, guideline developers also must balance the need for guidance with rapidly evolving research topics [43]. In this BPG specifically, conditional recommendations allow for guidance on an emerging topic (clinical practice in a digital health environment) while recognizing the need for nurses and health providers to consider the implications within their own health care context. Additionally, it has been argued that the implementation of all recommendations, including strong recommendations, depends on social and relational processes governing decision-making for individuals [43]. With this argument in mind, end users of all guidelines should think about contextual implications and the values and preferences of patients when implementing both strong and conditional recommendations.

A final limitation is that the authors only included studies published in English from 2017 onwards. They did not search for gray literature or search reference lists of included studies for further evidence due to timelines and feasibility. Therefore, it is possible that some additional studies were missed.

Conclusions

Digital health within the context of the clinical environment is an emerging topic. This BPG provides 4 evidence-based recommendations, along with good practice statements, implementation and evaluation, and monitoring resources. At the time of BPG development, no guidelines had been developed addressing evidence-based recommendations in this unique and growing area, especially as it relates to nurses and health providers. It is anticipated that this BPG can support nurses, other health providers, and health and academic organizations to make informed decisions about education and care related to digital health that can ultimately improve provider, patient, and system outcomes.

Acknowledgments

The authors would like to acknowledge the contributions of the 17 expert panel members for this BPG: Rob Beleno, Michelle Beyn (The Hospital for Sick Children), Rebecca Bliss (Hamilton Family Health Team), Elizabeth Borycki (University of Victoria), Margarita Cleghorne (Primacare Living Solutions Inc), Sandra Davidson (University of Calgary), Lorie Donelle (University of South Carolina, Western University), Alexander Hamilton (Kingston General Hospital), Tsz (Sally) Yee Ip (North York General Hospital, Holland Bloorview Kids Rehabilitation Hospital), Margery Konan (Michael Garron Hospital), Daile Moffat (Path2Care Inc), Sanaz Riahi (Ontario Shores Center for Mental Health Sciences), Tiziana Rivera (William Osler Health System), Tammy Rooke (CarePartners), Vaishali Sengar (Unity Health), Steve Sottile (Humber River Health), and Meredith Whitehead (Chatham-Kent Health Alliance). The authors would also like to acknowledge the external reviewers who provided feedback on the guideline before publication and additional Registered Nurses' Association of Ontario staff who assisted in creating this best practice guideline (BPG). The authors also acknowledge the leadership of the cofounder of this BPG, Humber River Health, and the significant contributions of expert panel cochair Vanessa Burkoski (Humber River Health), who passed away during the development of this BPG. The expert panel chose to dedicate the guideline to Vanessa Burkoski. Generative artificial intelligence was not used in any portion of the manuscript generation, or any portion of this work.

Funding

This work was funded primarily by the Government of Ontario. All work produced by the Registered Nurses' Association of Ontario is editorially independent from its funding source.

Data Availability

All data generated or analyzed during this review are included in this published article and its supplementary multimedia appendices. This information is also publicly available on the BPG website [4].

Authors' Contributions

Conceptualization: LB, LH, NN, CB, MC, JY

XSL•FO
RenderX

Formal analysis: LB, CB, LH
Funding acquisition: NN, JY
Project administration: LH, NN, CB
Supervision: MC, JY
Writing – original draft: LB, CB, LH, NN
Writing – review & editing: LB, LH, NN, CB, MC, JY

Conflicts of Interest

None declared.

Multimedia Appendix 1
Guideline Search Strategy.
[PDF File, 201 KB - nursing_v9i1e74942_app1.pdf]

Multimedia Appendix 2
Systematic Review Search Strategies.
[PDF File, 4325 KB - nursing_v9i1e74942_app2.pdf]

Multimedia Appendix 3
Sample Data Extraction Tables.
[XLSX File, 30 KB - nursing_v9i1e74942_app3.xlsx]

Multimedia Appendix 4
Recommendation 1 Evidence Profile.
[PDF File, 246 KB - nursing_v9i1e74942_app4.pdf]

Multimedia Appendix 5
Recommendation 2 Evidence Profile.
[PDF File, 333 KB - nursing_v9i1e74942_app5.pdf]

Multimedia Appendix 6
Recommendation 3 Evidence Profile.
[PDF File, 259 KB - nursing_v9i1e74942_app6.pdf]

Multimedia Appendix 7
Recommendation 4 Evidence Profile.
[PDF File, 310 KB - nursing_v9i1e74942_app7.pdf]

Checklist 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File, 193 KB - nursing_v9i1e74942_app8.pdf]

References

1. Abernethy A, Adams L, Barrett M, et al. The promise of digital health: then, now, and the future. NAM Perspect 2022;2022:10. [doi: 10.31478/202206e] [Medline: 36177208]
2. Global strategy on digital health 2020-2025. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2025-02-03]
3. Johnston K, Kervin L, Wyeth P. Defining digital technology. Centre of Excellence for the Digital Child. 2022. URL: https://www.digitalchild.org.au/blog/defining-digital-technology/#:~:text=We%20refer%20to%20digital%20technology,programmed%20to%20perform%20various%20functions [accessed 2025-02-03]
4. Clinical practice in a digital health environment. : Registered Nurses' Association of Ontario (RNAO); 2024 URL: https://rnao.ca/media/7163/download [accessed 2025-02-03]
5. Registered Nurses' Association of Ontario nursing and compassionate care in the age of artificial intelligence: engaging the emerging future. : Registered Nurses' Association of Ontario; 2020 URL: https://rnao.ca/sites/rnao-ca/files/

XSL•FO
RenderX

RNAO-AMS_Report-Nursing_and_Compassionate_Care_in_the_Age_of_AI_Final_For_Media_Release_10.21.2020.pdf [accessed 2025-02-03]
6. Booth RG, Strudwick G, McBride S, O'Connor S, Solano López AL. How the nursing profession should adapt for a digital future. BMJ 2021;373:n1190. [doi: 10.1136/bmj.n1190]
7. Fernandes MDF, Esteves RB, Teixeira CAB, Gherardi-Donato EDS. The present and the future of nursing in the brave new world. Rev Esc Enferm USP 2018 Jul 23;52:e03356. [doi: 10.1590/S1980-220X2017031603356] [Medline: 30043931]
8. Carrière R, MacDonald A, Chan Y. Past, present and future: the outlook from mid-career nurse informaticians. Nurs Leadersh (Tor Ont) 2016;28(4):8-17. [doi: 10.12927/cjnl.2016.24556] [Medline: 27122085]
9. Perry L. Machine learning: great opportunities, but will it replace nurses? Int J Nurs Pract 2019 Feb;25(1):e12725. [doi: 10.1111/ijn.12725] [Medline: 30734415]
10. Schunemann HJ, Brozek J, Guyatt G, Oxman A, editors. Handbook for Grading the Quality of Evidence and the Strength of Recommendations Using the GRADE Approach: GRADE Working Group; 2013. URL: https://gdt.gradepro.org/app/handbook/handbook.html [accessed 2025-02-03]
11. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
12. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. Can Med Assoc J 2010 Dec 14;182(18):E839-E842. [doi: 10.1503/cmaj.090449]
13. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence-indirectness. J Clin Epidemiol 2011 Dec;64(12):1303-1310. [doi: 10.1016/j.jclinepi.2011.04.014] [Medline: 21802903]
14. A revised cochrane risk-of-bias tool for randomized trials. : The Cochrane Collaboration; 2020 URL: https://www.cochrane.de/sites/cochrane.de/files/uploads/RoB_2.0_guidance_2019.pdf [accessed 2025-02-03]
15. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919. [doi: 10.1136/bmj.i4919] [Medline: 27733354]
16. Whiting P, Savović J, Higgins JPT, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2016 Jan;69:225-234. [doi: 10.1016/j.jclinepi.2015.06.005] [Medline: 26092286]
17. Howitt L, Burt A, Nizum N, Buchanan C, Rey M, Grinspun D. Best practice guideline development methods: an integrated approach based on the Registered Nurses' Association of Ontario's methods. Clinical and Public Health Guidelines 2026 Jan;3(1):1-8 [FREE Full text] [doi: 10.1002/gin2.70052]
18. Li YY, Au ML, Tong LK, Ng WI, Wang SC. High-fidelity simulation in undergraduate nursing education: a meta-analysis. Nurse Educ Today 2022 Apr;111:105291. [doi: 10.1016/j.nedt.2022.105291] [Medline: 35158134]
19. Hegland PA, Aarlie H, Strømme H, Jamtvedt G. Simulation-based training for nurses: systematic review and meta-analysis. Nurse Educ Today 2017 Jul;54:6-20. [doi: 10.1016/j.nedt.2017.04.004] [Medline: 28456053]
20. Oliveira Silva G, Oliveira FSE, Coelho ASG, et al. Effect of simulation on stress, anxiety, and self-confidence in nursing students: systematic review with meta-analysis and meta-regression. Int J Nurs Stud 2022 Sep;133:104282. [doi: 10.1016/j.ijnurstu.2022.104282] [Medline: 35679634]
21. Liu K, Zhang W, Li W, Wang T, Zheng Y. Effectiveness of virtual reality in nursing education: a systematic review and meta-analysis. BMC Med Educ 2023 Sep 28;23(1):710. [doi: 10.1186/s12909-023-04662-x] [Medline: 37770884]
22. Person- and family-centred care. : Registered Nurses' Association of Ontario; 2015 URL: https://rnao.ca/sites/rnao-ca/files/FINAL_Web_Version_0.pdf [accessed 2025-02-03]
23. Gilligan C, Powell M, Lynagh MC, et al. Interventions for improving medical students' interpersonal communication in medical consultations. Cochrane Database Syst Rev 2021 Feb 8;2(2):CD012418. [doi: 10.1002/14651858.CD012418.pub2] [Medline: 33559127]
24. Lee WW, Alkureishi ML, Wroblewski KE, Farnan JM, Arora VM. Incorporating the human touch: piloting a curriculum for patient-centered electronic health record use. Med Educ Online 2017;22(1):1396171. [doi: 10.1080/10872981.2017.1396171] [Medline: 29103366]
25. Gunner CK, Eisner E, Watson AJ, Duncan JL. Teaching webside manner: development and initial evaluation of a video consultation skills training module for undergraduate medical students. Med Educ Online 2021 Dec;26(1):1954492. [doi: 10.1080/10872981.2021.1954492] [Medline: 34313579]
26. Newcomb AB, Duval M, Bachman SL, Mohess D, Dort J, Kapadia MR. Building rapport and earning the surgical patient's trust in the era of social distancing: teaching patient-centered communication during video conference encounters to medical students. J Surg Educ 2021;78(1):336-341. [doi: 10.1016/j.jsurg.2020.06.018] [Medline: 32709566]
27. Mahabamunuge J, Farmer L, Pessolano J, Lakhi N. Implementation and assessment of a novel telehealth education curriculum for undergraduate medical students. J Adv Med Educ Prof 2021 Jul;9(3):127-135. [doi: 10.30476/jamp.2021.89447.1375] [Medline: 34277843]
28. Riley E, McCance C, Ward N, deGravelles P. Evaluation of a simulation-based learning experience using a prenatal telehealth scenario with prelicensure nursing students. Teach Learn Nurs 2022 Apr;17(2):220-224. [doi: 10.1016/j.teln.2022.01.006]

XSL•FO
RenderX

29. Lanier C, Dominicé Dao M, Hudelson P, Cerutti B, Junod Perron N. Learning to use electronic health records: can we stay patient-centered? A pre-post intervention study with family medicine residents. BMC Fam Pract 2017 May 26;18(1):69. [doi: 10.1186/s12875-017-0640-2] [Medline: 28549460]
30. Hall AM, Flodgren GM, Richmond HL, et al. Champions for improved adherence to guidelines in long-term care homes: a systematic review. Implement Sci Commun 2021 Aug 3;2(1):85. [doi: 10.1186/s43058-021-00185-y] [Medline: 34344486]
31. Walsh KE, Secor JL, Matsumura JS, et al. Secure provider-to-provider communication with electronic health record messaging: an educational outreach study. J Healthc Qual 2018;40(5):283-291. [doi: 10.1097/JHQ.0000000000000115] [Medline: 29280777]
32. Kadish SS, Mayer EL, Jackman DM, et al. Implementation to optimization: a tailored, data-driven approach to improve provider efficiency and confidence in use of the electronic medical record. J Oncol Pract 2018 Jul;14(7):e421-e428. [doi: 10.1200/JOP.18.00093] [Medline: 29939808]
33. Cresswell K, Callaghan M, Khan S, Sheikh Z, Mozaffar H, Sheikh A. Investigating the use of data-driven artificial intelligence in computerised decision support systems for health and social care: a systematic review. Health Informatics J 2020 Sep;26(3):2138-2147. [doi: 10.1177/1460458219900452] [Medline: 31964204]
34. Manaktala S, Claypool SR. Evaluating the impact of a computerized surveillance algorithm and decision support system on sepsis mortality. J Am Med Inform Assoc 2017 Jan;24(1):88-95. [doi: 10.1093/jamia/ocw056] [Medline: 27225197]
35. Klarenbeek SE, Weekenstroo HHA, Sedelaar JPM, Fütterer JJ, Prokop M, Tummers M. The effect of higher level computerized clinical decision support systems on oncology care: a systematic review. Cancers (Basel) 2020 Apr 22;12(4):1032. [doi: 10.3390/cancers12041032] [Medline: 32331449]
36. Frondelius T, Atkova I, Miettunen J, et al. Early prediction of ventilator-associated pneumonia with machine learning models: a systematic review and meta-analysis of prediction model performance☆. Eur J Intern Med 2024 Mar;121:76-87. [doi: 10.1016/j.ejim.2023.11.009] [Medline: 37981529]
37. Chidambaram S, Jain B, Jain U, et al. An introduction to digital determinants of health. PLOS Digit Health 2024 Jan;3(1):e0000346. [doi: 10.1371/journal.pdig.0000346] [Medline: 38175828]
38. Soomro KA, Kale U, Curtis R, Akcaoglu M, Bernstein M. Digital divide among higher education faculty. Int J Educ Technol High Educ 2020;17(1):21. [doi: 10.1186/s41239-020-00191-5] [Medline: 41001069]
39. Digital divide. Stanford University. URL: https://cs.stanford.edu/people/eroberts/cs181/projects/digital-divide/start.html [accessed 2025-02-03]
40. Richardson S, Lawrence K, Schoenthaler AM, Mann D. A framework for digital health equity. NPJ Digit Med 2022 Aug 18;5(1):119. [doi: 10.1038/s41746-022-00663-0] [Medline: 35982146]
41. Behaviour change: digital and mobile health interventions. National Institute for Health and Care Excellence.: NICE; 2020. URL: https://www.nice.org.uk/guidance/NG183 [accessed 2025-02-03]
42. Leading change toolkit. Registered Nurses' Association of Ontario. 2024. URL: https://www.RNAO.ca/leading-change-toolkit [accessed 2025-02-03]
43. Wyer PC, Gabbay J, Suh EH. Why all clinical guideline recommendations are 'conditional'. Clin Public Health Guidel 2024 May;1(2):e12013 [FREE Full text] [doi: 10.1002/gin2.12013]
44. WHO guideline on the use of glucagon-like peptide-1 (GLP-1) therapies for the treatment of obesity in adults. World Health Organization (WHO). 2025. URL: https://app.magicapp.org/#/guideline/LrRxrL [accessed 2025-12-12]

Abbreviations

AI: artificial intelligence
BPG: best practice guideline
BPSO: Best Practice Spotlight Organization
CDSS: clinical decision support system
EP: evidence profile
EtD: evidence to decision
GDM: guideline development methodologist
GRADE: Grading of Recommendations, Assessment, Development, and Evaluation
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews
RCT: randomized controlled trial
RNAO: Registered Nurses' Association of Ontario
ROBINS-I: risk of bias in nonrandomized studies—of interventions
ROBIS: risk of bias in systematic reviews

XSL•FO
RenderX

Edited by J Sarvestan; submitted 25.Mar.2025; peer-reviewed by F Lehocki, S Hatch, T Vagg; revised version received 15.Dec.2025; accepted 16.Dec.2025; published 23.Jan.2026.

Please cite as:
Bailey L, Howitt L, Nizum N, Buchanan C, Charlebois M, Yoon J, RNAO Expert Panel
Developing a Best Practice Guideline for Clinical Practice in a Digital Health Environment: Systematic Reviews Based on the Grading of Recommendations, Assessment, Development, and Evaluation Approach
JMIR Nursing 2026;9:e74942
URL: https://nursing.jmir.org/2026/1/e74942
doi:10.2196/74942

© Lauren Bailey, Lyndsay Howitt, Nafsin Nizum, Christine Buchanan, Maureen Charlebois, Jennifer Yoon, RNAO Expert Panel. Originally published in JMIR Nursing (https://nursing.jmir.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Natural Language Processing in Nursing Practice for Clinical Reasoning and Workflow: Scoping Review

Keisuke Iwase[1,2], MSN, RN; Taiki Teshima[2,3], MSN, RN; Miyae Yamakawa[2], RN, PhD; Yasushi Takeya[2], MD, PhD

[1]Faculty of Nursing, School of Health Sciences, Fujita Health University, Toyoake, Aichi, Japan

[2]Department of Evidence-Based Clinical Nursing, Division of Health Sciences, Graduate School of Medicine, Osaka University, 1-7 Yamadaoka, Suita, Osaka, Japan

[3]Faculty of Nursing, Kansai Medical University, Hirakata, Osaka, Japan

Corresponding Author:
Miyae Yamakawa, RN, PhD
Department of Evidence-Based Clinical Nursing, Division of Health Sciences, Graduate School of Medicine, Osaka University, 1-7 Yamadaoka, Suita, Osaka, Japan

Abstract

Background: Recent advancements in digital health technologies have led to the accumulation of large volumes of unstructured text data, particularly in electronic health records. Given that nursing documentation frequently includes detailed descriptions of daily living activities and psychosocial contexts, existing medical natural language processing (NLP) tools may not be directly applicable without adaptation. Moreover, the integration of standardized nursing terminologies into NLP systems remains limited. Although NLP has the potential to support risk detection, information management, and care planning, its practical applications in nursing have not yet been systematically mapped.

Objective: This scoping review aimed to map the practical applications of NLP in nursing practice, with a specific focus on real-world clinical use, particularly for assessment, documentation, and interventions, and to identify the study characteristics, application domains, outcomes, and challenges.

Methods: Following the Joanna Briggs Institute methodology and PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines, a systematic search was conducted across PubMed, CINAHL, the Cochrane Library, Scopus, Embase, and Ichushi-Web in October 2023 and June 2025. The population-concept-context framework (population: nurses; concept: NLP applications; context: clinical settings) was used to guide study selection. Two independent reviewers screened 2889 unique records and reviewed the full texts of 72 studies. Data were descriptively synthesized and thematically classified.

Results: Six studies met the inclusion criteria, indicating that empirically evaluated applications of NLP in real-world nursing practice remain limited. NLP applications were categorized into three domains: (1) text summarization and generation, (2) information extraction and classification, and (3) real-time detection and screening. NLP was applied across multiple stages of nursing practice, including handovers, assessment, care planning, continuity of care, and patient safety management. Notably, NLP was used as a form of augmented intelligence to complement, rather than replace, nurses' clinical judgment and to support clinical decision-making. These findings also highlight a substantial gap between rapid technological advances and the limited implementation of NLP in real-world nursing practice.

Conclusions: This review highlights the substantial gap between rapid technological advances and the limited implementation of NLP in real-world nursing practice. NLP may help reduce manual review workload, support early risk identification, and facilitate individualized care planning in selected nursing contexts. However, direct empirical evidence regarding its effects on nursing workflows, nurses' clinical decision-making, and patient-level outcomes remains limited. Future research should prioritize real-world implementation studies and evaluate nurses' acceptance, adoption barriers, workflow integration, and the influence of NLP on clinical decision-making.

(JMIR Nursing 2026;9:e94778) doi:10.2196/94778

KEYWORDS

clinical decision support; nursing documentation; natural language processing; scoping review; AI

XSL•FO
RenderX

Introduction

Recent advances in information technology have resulted in the daily generation of vast amounts of health care–related data through digital systems, such as electronic health records (EHRs). EHRs comprise longitudinal data that are routinely collected in clinical settings, including patient demographic information, vital signs, clinical documentation, insurance claims data, patient-reported outcomes, and data obtained from home-based monitoring devices [1]. A previous review examined applications of the Internet of Things in health care and demonstrated that the integration of Internet of Things devices with cloud computing and AI technologies enables real-time remote monitoring and data-driven clinical insights, thereby contributing to transformative changes in health care delivery [2]. These technological developments have made the acquisition of clinical data easier. Consequently, the effective use of health care data in medical and nursing practice has become increasingly important, highlighting the need for advanced analytical approaches such as natural language processing (NLP).

NLP enables computers to analyze text and speech in natural language and transform unstructured information into structured data that can be computationally processed [3]. In health care, NLP has attracted increasing attention as a means of extracting meaningful information from text data in EHRs, and research in this field has expanded rapidly in recent years [4,5]. However, a substantial proportion of health care data, including EHRs, consists of unstructured text, which is estimated to account for approximately 80% of all medical data [6].

Although these textual data contain rich information about patients' conditions, clinical trajectories, and interventions, they are more difficult to analyze and use than structured data [7]. Specifically, nursing documentation generally describes patients' activities of daily living and psychosocial contexts in greater detail than medical diagnoses, which limits the direct applicability of existing medical NLP tools to nursing records. Existing NLP tools are primarily optimized for diagnostic terminologies and may not adequately support the analysis of nuanced expressions, domain-specific vocabulary, and clinical reasoning inherent in nursing documentation. Furthermore, variations in documentation styles across institutions and individual nurses pose a fundamental challenge to the standardization of NLP systems. The integration of standardized nursing languages—such as NANDA-International (NANDA-I), the Nursing Outcomes Classification, and the Nursing Interventions Classification—into EHRs enables the systematic documentation of nursing practice and facilitates the visualization and evaluation of nursing expertise and professional contributions [8].

Previous studies [9] have reported various applications of NLP in health care, including patient risk prediction and the analysis of clinical conversations. In nursing, NLP has also been applied to tasks such as symptom detection from home care records [10], hospitalization risk prediction based on nursing documentation [11], extraction of symptom-related information from EHRs [12], and fall risk prediction [13]. However, previous reviews have reported that most studies have focused primarily on technical aspects, such as disease classification and algorithm development, with limited attention to practical implementation for clinical interpretation and nursing assessment [12,14].

In recent years, the application of generative AI and large language models (LLMs) in health care has expanded rapidly, attracting considerable attention. However, a gap between technological advancement and real-world implementation has also been highlighted. According to a recent review [15], numerous LLM-related studies in health care have been published between 2023 and mid-2024. Nevertheless, most of these studies have focused on tasks such as question answering, with only limited research addressing clinical workflow support. Furthermore, concerns have been raised regarding the methodological rigor of many studies focusing on these emerging technologies [16]. Therefore, despite rapid advancements in AI, structural barriers to integration into clinical practice remain unresolved. Accordingly, rather than focusing on emerging technologies such as LLMs or ambient AI, this study adopts an implementation-oriented scoping review that prioritizes NLP applications implemented and evaluated within real-world nursing workflows.

Therefore, this scoping review aimed to comprehensively map how NLP has been applied in nursing practice, particularly to support nursing assessment, documentation, and interventions, with a focus on real-world clinical implementation. The present study also aimed to characterize previous research, identify application domains, and clarify reported outcomes and challenges. Specifically, we sought to (1) categorize the distinct types and roles of NLP approaches, (2) identify their impact across various stages of the nursing process, and (3) summarize the reported outcomes and barriers to implementation. By doing so, this study provides foundational insights for integrating "augmented intelligence" into professional nursing care.

Methods

Study Design

This scoping review followed the Joanna Briggs Institute methodology [17] and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines [18] (Checklist 1). This design was selected to map the existing evidence base and identify gaps in the evolving domain of NLP in nursing practice.

Search Strategy

The literature search was conducted using PubMed, CINAHL, the Cochrane Library, Scopus, Embase, and Ichushi-Web (Japan). The initial search was conducted on October 5, 2023. In the updated search, performed on June 13, 2025, the phrase "data mining" was added to enhance the search's sensitivity and comprehensiveness. The comprehensive search strategies for each database are tabulated in Multimedia Appendix 1. Equivalent search strategies were adapted for each database, as appropriate. The search strategy was designed to prioritize foundational NLP and data mining approaches with established evidence of clinical integration, reflecting the implementation-oriented scope of this review. Accordingly, the

XSL•FO
RenderX

review focuses on empirically evaluated applications within nursing workflows, rather than capturing the broad spectrum of emerging technologies such as generative AI or LLMs.

Eligibility Criteria

Using the Joanna Briggs Institute population–concept-context framework, the inclusion criteria were as follows: nurses (population), NLP applications in nursing practice (concept), and clinical settings (context). Eligible studies were peer-reviewed original articles in English or Japanese that described NLP applications in nursing practice (eg, assessment, documentation, or interventions). The eligibility criteria were specifically defined to identify studies describing the use of NLP within real-world nursing practice and clinical workflows. Studies were excluded if they (1) focused solely on technical model performance or diagnostic accuracy without clinical application or nursing involvement, (2) focused on physician-centered or other nonnursing applications without explicit involvement of nurses, or (3) were gray literature (eg, conference abstracts or proceedings) lacking sufficient methodological detail. Among studies addressing similar topics (eg, automated identification from clinical or radiology reports), only those that described applications within nursing-related clinical workflows (eg, assessment, documentation, or interventions) were included.

Study Selection

Two independent reviewers screened studies in 2 stages: screening of titles and abstracts, followed by full-text review. Disagreements were resolved through consultation with a third author. 19-2119-24

Data Extraction and Analysis

Data were extracted using a standardized form (see Multimedia Appendix 2) that included author, year, country, design, population, nursing practice domain, NLP techniques, performance metrics, human intervention points, and outcomes. The extracted data were descriptively synthesized. NLP techniques were categorized according to their methodological approach (eg, rule-based and machine learning), and applications were thematically classified according to clinical context (assessment, documentation, and interventions). Common themes among outcomes and challenges were aggregated.

Ethical Considerations

This study is a scoping review based on previously published literature and did not involve human participants. Therefore, ethics approval was not required in accordance with the authors' institutional policies.

Results

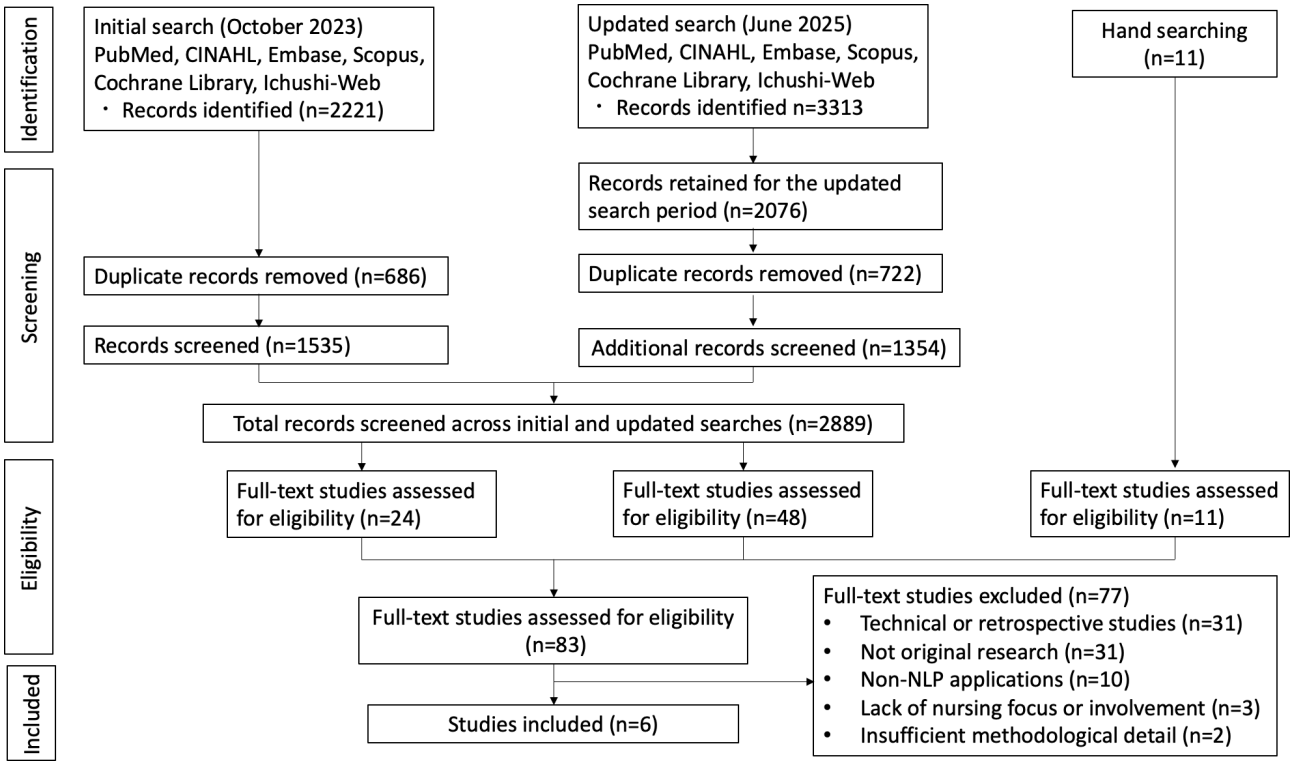
Study Selection Results

From the 1535 unique records in the initial search, 24 studies underwent full-text review, and 3 studies were included [19-21]. The updated search yielded an additional 1354 unique records and 48 full-text reviews, adding 3 more studies. Overall, 2889 unique records were screened, 72 full-text studies were reviewed, and 6 studies were included [19-24]. Hand searching the 11 references from relevant reviews identified no further studies.

As illustrated in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram, most of the records excluded during the title and abstract screening and the full-text review stages focused on the technical development or performance evaluation of NLP models and did not describe nursing practice or clinical workflow applications. The screening and selection process was conducted systematically and transparently based on predefined inclusion and exclusion criteria (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of study selection.



Characteristics of Included Studies

Table 1 summarizes the characteristics of the included studies. The examined studies were published between 2012 and 2025 [19-24]. Four studies were conducted in the United States [19,21,23,24], 1 in the United Kingdom [22], and 1 in Taiwan [20]. The study designs included a pilot implementation study [20], retrospective and prospective implementation evaluation studies [19,21-23], and an observational implementation study [24].

Table . Characteristics of included studies.

Author (year)	Setting	Population	Study design	Data sources
Hunter et al [22], 2012	NICU[a]	Neonates	Prospective	EHR[b] (physiological data, notes)
Evans et al [19], 2016	Emergency department	Heart failure patients	Pre-post	EHR (physiological data, reports)
Tang et al [20], 2019	LTCF[c]	Older adults	Pilot	Nursing records
Boitano et al [21], 2023	Academic hospital	AAA[d]	Retrospective and prospective	Radiology reports (EHR)
Waltzman et al [23], 2024	Community	HAV[e] cases	Cross-sectional	Nursing notes
Nguyen et al [24], 2025	Integrated system	Patient safety cases	Observational	EHR notes

[a]NICU: neonatal intensive care unit.

[b]EHR: electronic health record.

[c]LTCF: long-term care facility.

[d]AAA: abdominal aortic aneurysm.

[e]HAV: health care–associated violence.

The primary data sources used across these studies consisted of unstructured text documented in EHRs, including clinical notes, nursing notes, incident reports, and handover records. The NLP techniques involved rule-based approaches [22-24], machine learning–based classification and prediction methods [19,20,23], and natural language generation [22]. Notably, all included studies implemented NLP as part of existing clinical systems involving nurses, rather than as a stand-alone technical intervention. For a more comprehensive description of study characteristics, including detailed setting information and specific data extraction methods, see Multimedia Appendix 3.

XSL•FO
RenderX

Categorization and Roles of Applied NLP Approaches

Overview

The applied NLP approaches were further categorized into 3 distinct types based on their clinical roles (Table 2). For a more comprehensive description of these approaches, see Multimedia Appendix 4.

Table . Application of natural language processing (NLP) in nursing practice.

Author (year)	NLP task (type)	Role of NLP	Direct impact on nursing	Key challenges
Hunter et al [22], 2012	Summarization	Automated summarization of clinical information	Handover and information sharing	Domain-specific (respiratory and cardiovascular)
Tang et al [20], 2019	Summarization	Care plan generation based on clinical data	Support for chronic disease management	Data privacy and security
Evans et al [19], 2016	Extraction	Identification of high-risk patients from EHR[a] data	Clinical decision support	Multicenter validation needed
Boitano et al [21], 2023	Extraction	Identification of patients with AAA[b] from radiology reports	Population health and nurse navigation	Dependence on explicit documentation
Waltzman et al [23], 2024	Extraction	Extraction of violent incidents from nursing notes	HAV[c] prevention and safety management	Manual review burden
Nguyen et al [24], 2025	Detection	Real-time triage of patient safety concerns	Telephone nursing triage	Limited infrastructure

[a]EHR: electronic health record.

[b]AAA: abdominal aortic aneurysm.

[c]HAV: healthcare-associated violence.

Type 1: Text Summarization and Generation

This approach involves generating standardized information from free-text nursing documentation to serve specific clinical purposes. One system [22] used natural language generation techniques to automatically produce standardized nursing shift summaries from EHR notes. This system reduced information loss during handovers and improved the quality of clinical information transfer. Similarly, another system [20] was developed to integrate historical nursing records and admission assessment documents to generate individualized nursing care plans. Notably, the system automatically retrieved up-to-date evidence, such as best practice guidelines, from web-based sources and incorporated this information into the generated plans, supporting the integration of evidence-based practice into routine nursing care.

Type 2: Information Extraction and Classification

This approach focuses on systematically extracting and identifying information related to specific diseases, risks, or events from unstructured nursing documentation or clinical reports. One study [19] applied a proprietary NLP algorithm to analyze comprehensive medical records, significantly improving the sensitivity for identifying patients with heart failure. Implementation of the system was associated with reductions in clinicians' diagnostic time and the 30-day mortality rate. Similarly, an existing NLP system [21] was used to automatically identify patients with abdominal aortic aneurysms (AAA) from radiology reports, enabling reliable referral of patients requiring follow-up and resulting in timely surgical interventions. Furthermore, an NLP-based method [23] extracted health care–associated violence (HAV) events from nursing notes. This facilitated high-accuracy identification of violent incidents that were frequently underreported in conventional incident reporting systems, thereby enhancing the visibility of workplace safety issues.

Type 3: Real-Time Detection and Screening

This approach involves the real-time analysis of free-text data to prioritize triage decisions or identify patient safety concerns. A real-time analysis system [24] examined appointment notes and triage documentation during outpatient scheduling to identify safety concerns. When detecting high-risk signals, nurses conducted immediate telephone follow-ups. This mechanism served as a clinical "safety catch" by supporting the early identification and follow-up of serious safety concerns. Furthermore, the system reduced the number of cases requiring manual review by approximately 80% (from a monthly average of 4215 to 705 cases), demonstrating improved operational efficiency.

Applications and Outcomes Across the Stages of the Nursing Process

The applications of NLP identified in the included studies were categorized according to the nursing process stages (Tables 1 and 2).

Information Sharing and Handover

For information sharing during shift changes, NLP-based automated summarization supported nurses by presenting patient information in a structured and comprehensible format. One study [22] reported that automatically generated nursing handover summaries were highly rated by clinicians, with 90% considered easy to understand, 70% considered accurate, and 59% considered helpful. These findings suggest that

XSL•FO
RenderX

NLP-mediated summarization may support effective information sharing during handovers.

Assessment and Discharge Planning

The use of NLP enabled early identification of high-risk patients and contributed to the acceleration of discharge support processes. An integrated NLP system [19] was used to analyze unstructured clinical notes alongside existing structured data (such as *International Classification of Diseases* codes), improving the sensitivity of patient identification for heart failure. This approach reduced clinicians' chart review time (10 vs 40 min), facilitating earlier assessment and intervention.

Care Planning

NLP supported the more rapid development of individualized care plans. One study [20] reported a reduction in the time required for care planning from 7 to 4 days (a 42.8% improvement) and an increase in the acceptability of these plans from 39% to 78%. The system also supported more consistent care planning, particularly among less experienced staff.

Follow-Up and Continuity of Care

Continuous system-based monitoring optimized nurse navigation and enabled proactive outreach. A large-scale analysis [21] identified 6495 patients with AAA from 6.3 million radiology reports and facilitated 29 follow-up outpatient visits for patients whose surveillance had lapsed, as well as 4 surgical interventions for 3 patients. These outcomes helped improve adherence to guideline-recommended surveillance.

Safety Management and Surveillance

NLP enhanced the visibility of safety threats, which were often overlooked by conventional reporting systems. Using an NLP-assisted approach [23], HAV events were extracted from nursing notes with a sensitivity of 96.8%. In addition, a real-time safety screening [24] facilitated the identification of serious concerns and reduced the number of cases requiring manual review by approximately 80% (from a monthly average of 4215 to 705 cases), demonstrating improved operational efficiency.

Outcomes and Challenges of NLP Implementation

Across the included studies, the primary reported benefits of NLP implementation included reduced nursing workload, enhanced identification of specific risks or target populations, and support for nurses' clinical decision-making (Table 2). Specifically, the studies reported standardized handovers and reduced information loss [22], earlier identification of high-risk patients and acceleration of discharge planning [19], improved efficiency of care plan development and integration of up-to-date clinical knowledge [20], proactive follow-up of patients with AAA who have not received recommended surveillance [21], increased visibility of HAV events [23], and facilitation of concern identification with a substantial reduction in nurses' time spent on manual record review [24]. Notably, these functions did not replace nurses' decision-making; rather, NLP operated as a form of augmented intelligence that complemented professional judgment and supported the translation of insights into appropriate nursing interventions.

All included studies reported notable challenges and limitations. Some studies were conducted within a single institution [19], emphasizing the need for external validation across diverse clinical settings. Studies using rule-based or keyword-dependent approaches [21,22,24] reported that inaccuracies in documentation and variability in language use resulted in false positives or missed cases. To mitigate these technical limitations and the risk of erroneous outputs, the included studies implemented human-in-the-loop mechanisms to safeguard clinical judgment. These mechanisms functioned as important safeguards against algorithmic rigidity and potential errors in automated classification, ensuring that qualified clinicians, such as specialist nurses or nurse navigators, retained responsibility for clinical decisions by reviewing and validating NLP outputs (Table 2). Additionally, several studies emphasized that translating NLP-derived information into actual clinical interventions necessitated the availability of dedicated nursing personnel, such as nurse navigators or specialized teams [21,23]. Overall, these findings suggest that NLP complements rather than replaces nursing judgment. By structuring narrative data, NLP may facilitate informed decision-making. Therefore, future NLP systems should be designed as robust human-in-the-loop tools that empower nurses as final arbiters of clinical safety, ensuring seamless integration into nursing workflows.

Discussion

Summary of Findings

In this scoping review, only 6 studies [19-24] met the inclusion criteria, highlighting a substantial implementation gap between the rapid advancement of NLP technologies and their integration into real-world nursing practice. The limited number of eligible studies is an important finding in itself. Rather than focusing solely on whether NLP can technically process nursing documentation, this review highlights the broader question of why implementation-oriented NLP research in nursing remains so limited. The findings of this review suggest that this implementation gap reflects interrelated technical, organizational, and professional and cultural barriers to implementation.

Factors Contributing to the Implementation Gap

Technical Barriers

Nursing documentation often consists of heterogeneous free-text narratives describing patients' health status, symptoms, functional status, and other contextual information, which makes standardization challenging [8,14]. Standardized nursing terminologies are therefore fundamental to ensuring the comparability, interoperability, and secondary use of nursing data [8]. However, a recent review found that the use of standardized nursing terminologies in nursing NLP research remains limited [14]. Furthermore, a recent study reported that nurses perceived institutional and technical barriers to implementing NANDA-I, Nursing Interventions Classification, and Nursing Outcomes Classification in clinical practice [25]. These limitations in the adoption of standardized nursing terminologies may hinder the development and external validation of NLP systems for nursing documentation, thereby

XSL•FO
RenderX

limiting their implementation and generalizability. In addition, the limited representation of standardized nursing terminologies in NLP training datasets may constrain the ability of NLP models to recognize nursing-specific concepts, relationships, and clinical reasoning patterns.

Organizational Barriers

EHR systems continue to face challenges related to data access, interoperability, privacy protection, and data governance [1,26]. Access to nursing record datasets may also be restricted by institutional ethics review requirements, data-use approval processes, and concerns regarding the secondary use of sensitive clinical documentation. Similar implementation barriers, including inadequate technical support, poor system integration, and usability issues, have also been identified in a scoping review of EHR implementation [27]. Furthermore, insufficient involvement of end users during the planning, development, and implementation of EHR systems has been recognized as a common implementation barrier [27]. These organizational challenges may hinder the integration of NLP into routine nursing workflows and limit its implementation in clinical practice.

Professional and Cultural Barriers

Although physician-centered structured data have been the primary focus of EHR and NLP research, nursing documentation may have been underused. This imbalance may partly explain why implementation studies of NLP in nursing remain limited. Consistent with this concern, a previous review [28] found that nurses' involvement was limited, particularly during the early phases of AI development, and that ethical considerations related to the implementation and use of AI in nursing received limited attention. In addition, publication bias may have contributed to the underreporting of unsuccessful implementation efforts or negative findings, thereby skewing the available evidence toward successful cases. Taken together, these professional and cultural barriers may help explain why implementation-oriented NLP research in nursing remains limited. Addressing these barriers may require involving nurses as coinvestigators throughout the development and evaluation of NLP systems to ensure that these technologies are aligned with nursing practice.

Implications for Implementation

Unlike previous reviews, which have predominantly focused on the technical aspects of NLP or broader implementation and ethical issues related to AI in health care [12,14,29], this review examined how NLP has been integrated into nursing workflows and the extent to which nurses are involved in these systems in real-world clinical settings. Although recent interest has shifted toward LLM-based applications, empirical evidence regarding the implementation of these emerging technologies remains limited. Consistent with previous comprehensive reviews [28], there is limited empirical evidence regarding the effects of NLP on clinical outcomes, nurses' decision-making, and integration into routine nursing practice. This pattern is also evident in more recent research. For example, a recent review [15] reported that only 12.5% of 415 studies using advanced LLMs focused on clinical workflow support, whereas most applications addressed clinical decision support or patient information support. Similarly, although ambient AI technologies have recently received substantial attention, peer-reviewed empirical evidence regarding their implementation in nursing practice remains limited. Notably, although some studies have reported usability or perceived helpfulness [22], nurse acceptance and implementation barriers have rarely been examined explicitly, highlighting another important gap in the current evidence base. These findings suggest that the successful implementation of NLP in nursing depends not only on technological performance but also on organizational support, professional engagement, and well-defined implementation processes.

NLP as Augmented Intelligence

Previous reviews have noted that health care NLP has primarily emphasized algorithm development, data extraction, and classification tasks, with relatively limited attention to supporting nurses' judgment, decision-making, and care processes [12,14]. Specifically, one review [12] reported that many studies primarily used NLP for classification, whereas an integrative review [14] highlighted the need for improved evaluation, standardization, and broader application of NLP in nursing. Building on these observations, this review shifted the focus from technical accuracy alone to empirical evidence of how NLP is used in nursing practice and how it may support clinical decision-making. The included studies suggest that NLP may reduce manual review workload, improve patient identification, and facilitate clinical decision-making in selected nursing contexts [19,21,24]. Collectively, these findings indicate that NLP may support clinical workflows by functioning as an analytical tool and a decision support resource. However, direct empirical evidence regarding patient outcomes remains limited.

These findings suggest that NLP should not be conceptualized as a technology that fully automates nurses' clinical judgment. Rather, it may be better understood as a form of augmented intelligence that complements and supports cognitive processes in nursing practice. In particular, NLP may support clinical reasoning by processing large volumes of unstructured data during the early stages of assessment and information processing [30]. This perspective aligns with the principle that clinical decision support systems should be integrated into clinical workflows while preserving professional judgment [31]. The reported usefulness of automatically generated summaries [22] is consistent with this perspective and with the broader concept of the convergence of human and artificial intelligence [32].

Previous studies have identified several challenges associated with clinical decision support systems, including omission and commission errors [33] and concerns regarding data security and patient privacy [20]. As NLP becomes increasingly incorporated into clinical decision support and nursing workflows, related risks, including automation bias, warrant careful consideration. However, empirical evidence on how these challenges affect the implementation of NLP in nursing practice remains limited.

Accordingly, future implementation efforts should focus on the safe and responsible integration of NLP into nursing practice. Because NLP applications process sensitive personal health information, robust privacy protection and data security are essential [20,34,35]. Previous studies have also emphasized the

XSL•FO
RenderX

importance of explainability and responsible AI implementation [36]. Building on these considerations, future research should explore strategies to improve AI literacy among nurses and thereby facilitate the responsible implementation of NLP in clinical practice.

Taken together, these findings suggest that NLP should be considered not merely as a technical tool but as part of clinical systems that involve nurse engagement. This perspective is conceptually aligned with the Non-Adoption, Abandonment, Scale-Up, Spread, and Sustainability framework [26], particularly regarding NLP adoption, integration into routine clinical workflows, and long-term sustainability in nursing practice. The technical, organizational, and professional barriers identified in this review may represent key challenges to the adoption, routine integration, and long-term sustainability of NLP in nursing practice. This review, therefore, provides an implementation-oriented perspective that may guide future research and support the integration of NLP into nursing practice.

Limitations

This scoping review has several limitations. First, the literature search was restricted to studies published in English and Japanese, potentially introducing language bias. Second, the number of databases searched was limited, and gray literature, including doctoral dissertations, conference proceedings, and unpublished reports, was excluded. Consequently, some relevant studies may have been missed. Furthermore, as this review specifically focused on NLP applications in nursing practice, it excluded studies that did not explicitly involve nurses in clinical workflows (eg, physician-centered applications). Future reviews should incorporate broader database coverage and gray literature to provide a more exhaustive overview of the field. Third, the number of included studies was small, reflecting the limited implementation of NLP in nursing practice and the early stage of implementation-oriented research in this field. Accordingly, this review provides a map of current research trends rather than definitive conclusions, and further empirical studies are warranted as the field evolves. Fourth, the search strategy was predominantly designed to capture foundational NLP and data mining techniques. Therefore, it may not have fully captured the most recent studies focusing on generative AI and ambient AI technologies, which have proliferated rapidly in recent years. Given the recent surge in studies using these technologies for nursing documentation, summarization, and care planning, this review's findings may not fully reflect this emerging body of work. Future studies should compare traditional NLP approaches with newer generative AI technologies to explore their distinct impacts on nursing practice. Fifth, a geographic bias exists, as the included studies were conducted in a limited number of regions, specifically the United Kingdom, the United States, and Taiwan. Consequently, the findings may have limited generalizability to diverse nursing environments with different health care systems and cultural contexts. Finally, in accordance with the methodological nature of a scoping review, this study did not include a formal appraisal of the methodological quality or the risk of bias of the included studies. Therefore, caution should be exercised when interpreting the findings. Future research should incorporate systematic quality assessment to strengthen the evidence base and validate the conclusions drawn.

Conclusions

This scoping review suggests that NLP may support nursing practice by extracting and generating clinically relevant information from unstructured nursing-related text. However, empirical evidence regarding its implementation in real-world nursing practice remains limited. Rather than replacing professional judgment, NLP may function as a form of augmented intelligence that supports nurses' clinical decision-making. In selected nursing contexts, NLP may help reduce manual review workload, facilitate early risk identification, and support information processing. Nevertheless, direct empirical evidence regarding its effects on nursing workflows, nurses' clinical decision-making, and patient-level outcomes is limited. Future research should examine nurses' acceptance, adoption barriers, and implementation processes in real-world clinical settings.

Although previous studies have often focused on system development and technical performance, future research should prioritize implementation-oriented evaluations and assess how NLP influences nurses' clinical decision-making in practice. To support the responsible integration of AI technologies into nursing practice, nurses may require sufficient AI literacy to critically appraise and interpret AI-generated outputs. Developing this knowledge and these skills may help nurses use AI technologies while maintaining professional judgment in an increasingly digital health care environment.

The principal contribution of this review is to identify the substantial gap between rapid technological advances and the limited implementation of NLP in real-world nursing practice, rather than to establish the effectiveness of NLP.

Acknowledgments

We express our sincere gratitude to all contributors to this study. We also thank Editage for their English language editing services.

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) and Gemini (Google) for English translation and to improve the manuscript's readability. The authors take full responsibility for the content.

Funding

This study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

XSL•FO
RenderX

Data Availability

Data sharing is not applicable to this study as no data sets were generated or analyzed during this study.

Authors' Contributions

Conceptualization: KI, MY
Data curation: KI, TT
Formal Analysis: KI, TT
Methodology: KI, MY, YT
Supervision: MY, YT
Writing – original draft: KI
Writing – review & editing: MY, YT
All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies for all databases.
[DOCX File, 24 KB - nursing_v9i1e94778_app1.docx]

Multimedia Appendix 2
Codebook.
[DOCX File, 21 KB - nursing_v9i1e94778_app2.docx]

Multimedia Appendix 3
Characteristics of included studies.
[DOCX File, 17 KB - nursing_v9i1e94778_app3.docx]

Multimedia Appendix 4
Application of natural language processing in nursing practice.
[DOCX File, 19 KB - nursing_v9i1e94778_app4.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 25 KB - nursing_v9i1e94778_app5.docx]

References

1. Cowie MR, Blomster JI, Curtis LH, et al. Electronic health records to facilitate clinical research. Clin Res Cardiol 2017 Jan;106(1):1-9. [doi: 10.1007/s00392-016-1025-6] [Medline: 27557678]
2. Li C, Wang J, Wang S, Zhang Y. A review of IoT applications in healthcare. Neurocomputing 2024 Jan;565:127017. [doi: 10.1016/j.neucom.2023.127017]
3. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. J Am Med Inform Assoc 2011;18(5):544-551. [doi: 10.1136/amiajnl-2011-000464] [Medline: 21846786]
4. Wu S, Roberts K, Datta S, et al. Deep learning in clinical natural language processing: a methodical review. J Am Med Inform Assoc 2020 Mar 1;27(3):457-470. [doi: 10.1093/jamia/ocz200] [Medline: 31794016]
5. Wu H, Wang M, Wu J, et al. A survey on clinical natural language processing in the United Kingdom from 2007 to 2022. NPJ Digit Med 2022 Dec 21;5(1):186. [doi: 10.1038/s41746-022-00730-6] [Medline: 36544046]
6. Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA 2013 Apr 3;309(13):1351-1352. [doi: 10.1001/jama.2013.393] [Medline: 23549579]
7. Durieux BN, Tarbi EC, Lindvall C. Opportunities for computational tools in palliative care: supporting patient needs and lowering burden. Palliat Med 2022 Sep;36(8):1168-1170. [doi: 10.1177/02692163221122261] [Medline: 36062726]
8. Jones D, Lunney M, Keenan G, Moorhead S. Standardized nursing languages: essential for the nursing workforce. Annu Rev Nurs Res 2010;28:253-294. [doi: 10.1891/0739-6686.28.253] [Medline: 21639030]
9. Hossain E, Rana R, Higgins N, et al. Natural language processing in electronic health records in relation to healthcare decision-making: a systematic review. Comput Biol Med 2023 Mar;155:106649. [doi: 10.1016/j.compbiomed.2023.106649] [Medline: 36805219]

XSL•FO
RenderX

10. Woo K, Adams V, Wilson P, et al. Identifying urinary tract infection–related information in home care nursing notes. J Am Med Dir Assoc 2021 May;22(5):1015-1021.e2. [doi: 10.1016/j.jamda.2020.12.010] [Medline: 33434568]
11. Topaz M, Koleck TA, Onorato N, Smaldone A, Bakken S. Nursing documentation of symptoms is associated with higher risk of emergency department visits and hospitalizations in homecare patients. Nurs Outlook 2021;69(3):435-446. [doi: 10.1016/j.outlook.2020.12.007] [Medline: 33386145]
12. Wieland-Jorna Y, van Kooten D, Verheij RA, de Man Y, Francke AL, Oosterveld-Vlug MG. Natural language processing systems for extracting information from electronic health records about activities of daily living. A systematic review. JAMIA Open 2024;7(2):ooae044. [doi: 10.1093/jamiaopen/ooae044] [Medline: 38798774]
13. Trinh VQN, Zhang S, Kovoor J, et al. The use of natural language processing in detecting and predicting falls within the healthcare setting: a systematic review. Int J Qual Health Care 2023 Oct 17;35(4):mzad077. [doi: 10.1093/intqhc/mzad077] [Medline: 37758209]
14. Mitha S, Schwartz J, Hobensack M, et al. Natural language processing of nursing notes: an integrative review. Comput Inform Nurs 2023 Jun 1;41(6):377-384. [doi: 10.1097/CIN.0000000000000967] [Medline: 36730744]
15. Zhang Z, Momeni Nezhad MJ, Bagher Hosseini SM, et al. Advancing healthcare with large language models: a scoping review of applications and future directions. Int J Med Inform 2026 Mar 15;208:106231. [doi: 10.1016/j.ijmedinf.2025.106231] [Medline: 41443123]
16. Lai X, Lai Y, Chen J, Huang S, Gao Q, Huang C. Evaluation strategies for large language model-based models in exercise and health coaching: scoping review. J Med Internet Res 2025 Oct 14;27:e79217. [doi: 10.2196/79217] [Medline: 41086432]
17. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
18. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
19. Evans RS, Benuzillo J, Horne BD, et al. Automated identification and predictive tools to help identify high-risk heart failure patients: pilot evaluation. J Am Med Inform Assoc 2016 Sep;23(5):872-878. [doi: 10.1093/jamia/ocv197] [Medline: 26911827]
20. Tang V, Siu PKY, Choy KL, et al. An adaptive clinical decision support system for serving the elderly with chronic diseases in healthcare industry. Expert Systems 2019 Apr;36(2):e12369. [doi: 10.1111/exsy.12369]
21. Boitano LT, DeVivo G, Robichaud DI, et al. Successful implementation of a nurse-navigator-run program using natural language processing identifying patients with an abdominal aortic aneurysm. J Vasc Surg 2023 Mar;77(3):922-929. [doi: 10.1016/j.jvs.2022.10.034] [Medline: 36328142]
22. Hunter J, Freer Y, Gatt A, Reiter E, Sripada S, Sykes C. Automatic generation of natural language nursing shift summaries in neonatal intensive care: BT-Nurse. Artif Intell Med 2012 Nov;56(3):157-172. [doi: 10.1016/j.artmed.2012.09.002] [Medline: 23068882]
23. Waltzman M, Ozonoff A, Fournier KA, et al. Surveillance of health care-associated violence using natural language processing. Pediatrics 2024;154(2):e2023063059. [doi: 10.1542/peds.2023-063059] [Medline: 38973359]
24. Nguyen K, Springob P, Evans A, et al. A safety catch system for patient self-service appointment booking. NEJM Catal 2025 Feb;6(3). [doi: 10.1056/CAT.24.0162]
25. Gilart E, Bocchino A, Gilart-Cantizano P, et al. The integration of AI into the nursing process: a comparative analysis of NANDA, NOC, and NIC-based care plans. Nurs Rep 2025 May 27;15(6):186. [doi: 10.3390/nursrep15060186] [Medline: 40559477]
26. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
27. Tsai CH, Eghdam A, Davoody N, Wright G, Flowerday S, Koch S. Effects of electronic health record implementation and barriers to adoption and use: a scoping review and qualitative analysis of the content. Life (Basel) 2020 Dec 4;10(12):327. [doi: 10.3390/life10120327] [Medline: 33291615]
28. von Gerich H, Moen H, Block LJ, et al. Artificial intelligence -based technologies in nursing: a scoping literature review of the evidence. Int J Nurs Stud 2022 Mar;127:104153. [doi: 10.1016/j.ijnurstu.2021.104153] [Medline: 35092870]
29. Abdelmohsen SA, Al-Jabri MM. Artificial intelligence applications in healthcare: a systematic review of their impact on nursing practice and patient outcomes. J Nurs Scholarsh 2025 Nov;57(6):957-966. [doi: 10.1111/jnu.70040] [Medline: 40836587]
30. Banning M. Clinical reasoning and its application to nursing: concepts and research studies. Nurse Educ Pract 2008 May;8(3):177-183. [doi: 10.1016/j.nepr.2007.06.004] [Medline: 17869587]
31. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. JAMA 2018 Dec 4;320(21):2199-2200. [doi: 10.1001/jama.2018.17163] [Medline: 30398550]
32. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019 Jan;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
33. Skitka LJ, Mosier KL, Burdick M. Does automation bias decision-making? Int J Hum Comput Stud 1999 Nov;51(5):991-1006. [doi: 10.1006/ijhc.1999.0252]

XSL•FO
RenderX

34. Albouq SS, Sen AAA, Almashf N, Yamin M, Alshanqiti A, Bahbouh NM. A survey of interoperability challenges and solutions for dealing with them in IoT environment. IEEE Access 2022;10:36416-36428. [doi: 10.1109/ACCESS.2022.3162219]
35. Sharma A, Bhushan K. A comprehensive survey on IoT security: challenges, security issues, and countermeasures. Comput Sci Rev 2026 Feb;59:100839. [doi: 10.1016/j.cosrev.2025.100839]
36. London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. Hastings Cent Rep 2019 Jan;49(1):15-21. [doi: 10.1002/hast.973] [Medline: 30790315]

Abbreviations

AAA: abdominal aortic aneurysms
EHR: electronic health record
HAV: health care–associated violence
LLM: large language model
NLP: natural language processing
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by K Cato; submitted 06.Mar.2026; peer-reviewed by A K L, M Torii; revised version received 14.Jul.2026; accepted 20.Aug.2026; published 16.Oct.2026.

Please cite as:
Iwase K, Teshima T, Yamakawa M, Takeya Y
Natural Language Processing in Nursing Practice for Clinical Reasoning and Workflow: Scoping Review
JMIR Nursing 2026;9:e94778
URL: https://nursing.jmir.org/2026/1/e94778
doi:10.2196/94778

© Keisuke Iwase, Taiki Teshima, Miyae Yamakawa, Yasushi Takeya. Originally published in JMIR Nursing (https://nursing.jmir.org), 16.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Integration of Virtual and Augmented Reality Into Obstetric Nursing and Midwifery Education: Systematic Review

Mariana Luisa Firmiano[1], MSc; Ana Paula Assunção Moreira[2], PhD; Flaviana Vely Mendonça Vieira[3], PhD; Carla Sa-Couto[4], PhD; Alexandrina Cardoso[5], PhD

[1]School of Medicine and Biomedical Sciences, University of Porto, Rua de Jorge Viterbo Ferreira, 228, Porto, Portugal

[2]Nursing School, Federal University of São Paulo, São Paulo, São Paulo, Brazil

[3]Faculty of Nursing, Federal University of Goiás, Goiânia, Goiás, Brazil

[4]RISE-Health, Faculty of Medicine, University of Porto, Porto, Portugal

[5]RISE-Health, Nursing School of University of Porto, University of Porto, Porto, Portugal

Corresponding Author:
Mariana Luisa Firmiano, MSc
School of Medicine and Biomedical Sciences, University of Porto, Rua de Jorge Viterbo Ferreira, 228, Porto, Portugal

Abstract

Background: Technologies such as virtual reality (VR) and augmented reality (AR) have been increasingly incorporated into nursing education to support the development of cognitive, psychomotor, and behavioral competencies. In midwifery, immersive environments offer opportunities to simulate high-risk or low-frequency clinical scenarios, strengthening students' confidence and preparedness for professional practice. However, there is still significant variability in how VR and AR are pedagogically implemented, with limited understanding of their theoretical grounding, instructional design, and educational outcomes.

Objective: This systematic review aimed to synthesize and critically evaluate the characteristics of educational programs using VR and AR to develop competencies among nursing and midwifery students and professionals.

Methods: This review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 and Synthesis Without Meta-Analysis (SWiM) guidelines and was registered in PROSPERO (International Prospective Register of Systematic Reviews). A comprehensive search was conducted in August 2024 across PubMed, Web of Science, Scopus, and Cumulative Index to Nursing and Allied Health Literature (CINAHL) databases. Eligible studies were experimental, quasi-experimental, or observational, and involved undergraduate or postgraduate students in nursing, obstetric nursing, or midwifery, as well as certified midwives or nurse-midwives. Two independent reviewers conducted screening, data extraction, and quality appraisal using Joanna Briggs Institute (JBI) tools and the Medical Education Research Study Quality Instrument (MERSQI) scale. Data were synthesized narratively according to predefined thematic categories.

Results: Four studies published between 2021 and 2024 met the inclusion criteria, encompassing 634 participants (360 midwifery students and 274 midwives or nurse-midwives). Two randomized controlled trials and 2 quasi-experimental studies were included. Interventions used VR in 3 studies and AR in one, simulating scenarios such as normal and complicated childbirth, neonatal resuscitation, and emergency obstetric procedures. Most interventions focused on developing technical competencies, while one also addressed teamwork and communication. Only 2 studies explicitly reported theoretical or pedagogical frameworks, and 2 mentioned alignment with the INACSL Healthcare Simulation Standards of Best Practice. Structured briefing or debriefing was rarely described. All studies assessed learning outcomes (Kirkpatrick level 2), and 3 included participants' satisfaction (level 1). Methodological quality, measured by MERSQI, ranged from 10.0 to 14.5, with a mean score of 12.9 (SD 2.0), indicating moderate to high quality.

Conclusions: Evidence suggests that VR and AR can enhance learning in obstetric nursing and midwifery education by providing safe, experiential, and interactive learning environments. However, the limited use of theoretical frameworks and structured instructional design highlights the need for more pedagogically grounded interventions. Future research should prioritize the integration of learning theories, validated assessment tools, and best-practice simulation standards to strengthen the educational impact and curricular integration of immersive technologies in nursing and midwifery education.

Trial Registration: PROSPERO CRD42025635292; https://www.crd.york.ac.uk/PROSPERO/view/CRD42025635292

(JMIR Nursing 2026;9:e90252) doi:10.2196/90252

KEYWORDS

midwifery; virtual reality; augmented reality; obstetric nursing; nursing education; nurse midwives; simulation-based learning

XSL•FO
RenderX

Introduction

Background

The integration of virtual reality (VR) and augmented reality (AR) has demonstrated promising effectiveness in health education, particularly in the development of cognitive, psychomotor, and behavioral competencies [1-3]. These technologies enable immersion in interactive 3D environments, allowing for the safe practice of complex procedures and the visualization of anatomical and physiological phenomena that are difficult to observe in real clinical settings [1,4]. When applied in a planned manner and supported by clear educational objectives, they foster clinical reasoning, decision-making [5], and student engagement [6].

Nursing education has explored the potential of these technologies across different clinical areas [2], and obstetric nursing and midwifery, in particular, has benefited from the ability to simulate complex clinical scenarios, contributing to the development of technical and nontechnical competencies essential for safe care of women throughout the perinatal period [7,8].

Despite advances in research, there remains significant variability in how VR and AR are implemented within nursing education programs, reflecting the diversity of pedagogical approaches and levels of curricular integration reported in the literature [5,9,10]. Given the inherent complexity of immersive environments, challenges related to students' cognitive load may be exacerbated, particularly when the degree of immersiveness and the instructional design are not aligned with the learning objectives [11,12]. Therefore, it is relevant to understand how immersive technologies have been applied in obstetric nursing and midwifery education and to what extent their implementation is grounded in solid pedagogical and theoretical principles [13,14].

In this context, a systematic synthesis of the literature is essential to identify how VR and AR have been integrated into obstetric nursing and midwifery education, evaluating the presence of theoretical foundations, instructional models, and assessment strategies that guide their pedagogical application. This analysis contributes to the development of evidence-based educational practices for the integration of these technologies into nursing and midwifery education.

Objective

This systematic review aimed to synthesize and critically evaluate the characteristics of educational programs that use VR and AR technologies to develop obstetric competencies among nursing and midwifery students and professionals.

Methodology

Study Design

This systematic review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Checklist 1) guidelines and the Cochrane Handbook for Systematic Reviews of Interventions [15,16]. The study protocol was developed by the research team and registered in PROSPERO (International Prospective Register of Systematic Reviews; registration number: CRD42025635292). No ethical approval was required.

Information Sources and Search Strategy

The search was conducted on August 14, 2024, in the following databases: PubMed, Web of Science, Scopus, and Cumulative Index to Nursing and Allied Health Literature (CINAHL).

The search strategy was structured based on the PIO framework (Population, Intervention, Outcome). The population included undergraduate and postgraduate students in nursing, obstetric nursing, or midwifery, as well as practicing midwives and nurse-midwives. The interventions comprised educational programs using VR and AR as pedagogical tools for developing clinical competencies in obstetrics. The outcomes of interest encompassed theoretical and pedagogical models, instructional design, assessment strategies, and competencies addressed by the interventions.

The terms were organized into 3 conceptual blocks and adapted to the syntax of each database. No filters were applied regarding language, publication date, or document type. The complete search strategy is presented in Multimedia Appendix 1.

Eligibility Criteria

Eligible studies were primary research with experimental, quasi-experimental, or observational designs, published in peer-reviewed journals. The population included undergraduate nursing students, obstetric nursing and midwifery students, postgraduate obstetric nursing and midwifery students, and certified midwives and nurse-midwives.

Interventions were required to involve extended reality (XR) technologies, operationalized in this review as VR or AR applications used for clinical simulation with active learner interaction. VR interventions were eligible when they involved head-mounted or headset-based immersive delivery and first-person interaction within a simulated clinical environment, such as manipulating virtual objects, selecting or performing clinical actions, or responding to scenario-based clinical cues. AR interventions were eligible when they involved digital overlays integrated into the real-world environment, accessed through mobile or wearable devices, and required interaction with virtual elements as part of the simulated clinical activity. Passive 360-degree video, noninteractive video-based learning, desktop-based simulation, standard e-learning, mobile applications without AR overlays or immersive delivery, physical simulations without VR or AR integration, training conducted outside educational contexts, and studies in general nursing without a specific focus on obstetric or immediate neonatal care were excluded (Multimedia Appendix 1).

Study Selection Process

The search results were imported into Rayyan (Qatar Computing Research Institute Hamad Bin Khalifa University) [17]. After duplicate removal, 2 reviewers (MLF and APM) independently and blindly screened the studies across 3 successive phases: titles, abstracts, and full texts. Interrater reliability was assessed using the final subset of 33 records evaluated at the full-text stage. Agreement was reached for 28 of the 33 decisions

XSL•FO
RenderX

(84.8%), yielding a Cohen kappa coefficient of 0.52, which indicates moderate agreement. Each reviewer systematically recorded the reasons for exclusion using a structured format based on the eligibility criteria. Discrepant decisions were initially discussed between the 2 reviewers to reach consensus. When consensus was not possible, the studies were referred to 2 additional reviewers (AC and CSC), who independently reassessed them using the same criteria. Final inclusion or exclusion decisions were made after discussion among the reviewers, considering the predefined methodological criteria and the additional assessments provided.

Most disagreements involved borderline eligibility decisions related to the operational definition of immersive technology and clinical simulation. In particular, discrepancies concerned the distinction between fully immersive VR and 360-degree video or other noninteractive formats, the minimum level of user interaction required for inclusion, and whether the intervention involved scenario-driven simulation within an obstetric or immediate neonatal care context. These issues were resolved by applying the predefined eligibility criteria, which required immersive VR or AR with active learner interaction and a specific focus on obstetric or immediate neonatal care education.

Data Extraction

Data extraction was conducted independently by 2 reviewers (MLF and APM) using a standardized form, specifically developed for this review based on the study objectives. The form was pilot-tested on 2 included studies, allowing adjustments to ensure clarity and consistency in data collection.

Information was extracted regarding study characteristics, population, and educational intervention. These data included: study title, authors, year of publication, and country; study design; participants' professional background and educational level; learning theory adopted; instructional design models; the Bloom Revised Taxonomy domains targeted by the intervention [18]; Kirkpatrick levels used for evaluating the interventions [19]; type of clinical environment and simulated context; technical and nontechnical skills addressed; data collection methods and assessment instruments (type, content, validity, and timing); and data analysis methods. Regarding technological features, we extracted system-level characteristics related to immersive delivery, including device type and interaction modality with virtual elements. Additional data included supporting educational materials, use of complementary teaching strategies, duration of the simulation, reported implementation challenges, funding sources, and outcomes related to knowledge, skills, attitudes, and learning experience. Discrepancies were resolved by consensus or by consulting a third reviewer (CSC). The data were organized using Microsoft Excel spreadsheets.

Risk-of-Bias and Methodological Quality Assessment

Risk of bias was assessed independently by two reviewers (MLF and APM) using the Joanna Briggs Institute (JBI) critical appraisal tools [15], applied according to the methodological design of each included study.

For randomized controlled trials (RCTs), the revised JBI checklist comprising 13 items was used [20]. These are organized into 5 core domains: participant selection and allocation, delivery of the intervention or exposure, assessment, detection and measurement of outcomes, participant retention, and statistical conclusion validity.

For quasi-experimental studies, the revised JBI checklist with 9 items was applied [21], covering the following domains: temporal-procedural bias, selection and allocation bias, confounding bias, bias in the delivery of the intervention or exposure, bias in the assessment, detection and measurement of outcomes, and bias related to participant retention.

Each item in the checklists was classified as "yes," "no," "unclear," or "not applicable," in accordance with the guidance provided in the respective tools. Disagreements between reviewers were resolved by consensus.

In addition, the Medical Education Research Study Quality Instrument (MERSQI) [22] was used to assess the methodological quality of studies from the perspective of health professions education research. The instrument comprises 6 domains: study design, sampling, type of data, validity of instruments, statistical analysis, and educational outcomes. Each domain has a maximum score of 3 points, with the total score ranging from 5 to 18 points.

Uncertainties and disagreements were addressed and resolved through consensus. Studies were not blinded with regard to authorship, institutional affiliation, or publishing journal.

Data Synthesis Strategy

In line with the review objectives, a narrative synthesis approach was planned a priori and conducted following the Synthesis Without Meta-Analysis (SWiM) reporting guidelines [23]. A meta-analysis was not performed due to substantial heterogeneity across the included studies, including differences in intervention characteristics (VR vs AR), study designs, outcome measures, assessment instruments, and timing of evaluation. In addition, quantitative data were reported inconsistently, with some studies presenting means and SDs, while others reported percentages or did not provide sufficient data, further precluding meaningful statistical pooling.

The extracted information was organized into predefined thematic categories, including study characteristics, educational interventions, theoretical and pedagogical foundations, technologies used, assessment instruments, and main findings. The results were presented descriptively and analytically, allowing comparisons based on shared methodological elements. Supporting tables and summary charts were developed to facilitate comparative analysis, and quantitative evidence was summarized where appropriate. Findings were synthesized by grouping results according to outcome domains (knowledge, skills, attitudes, and learning experience) and by examining the direction and consistency of results across studies.

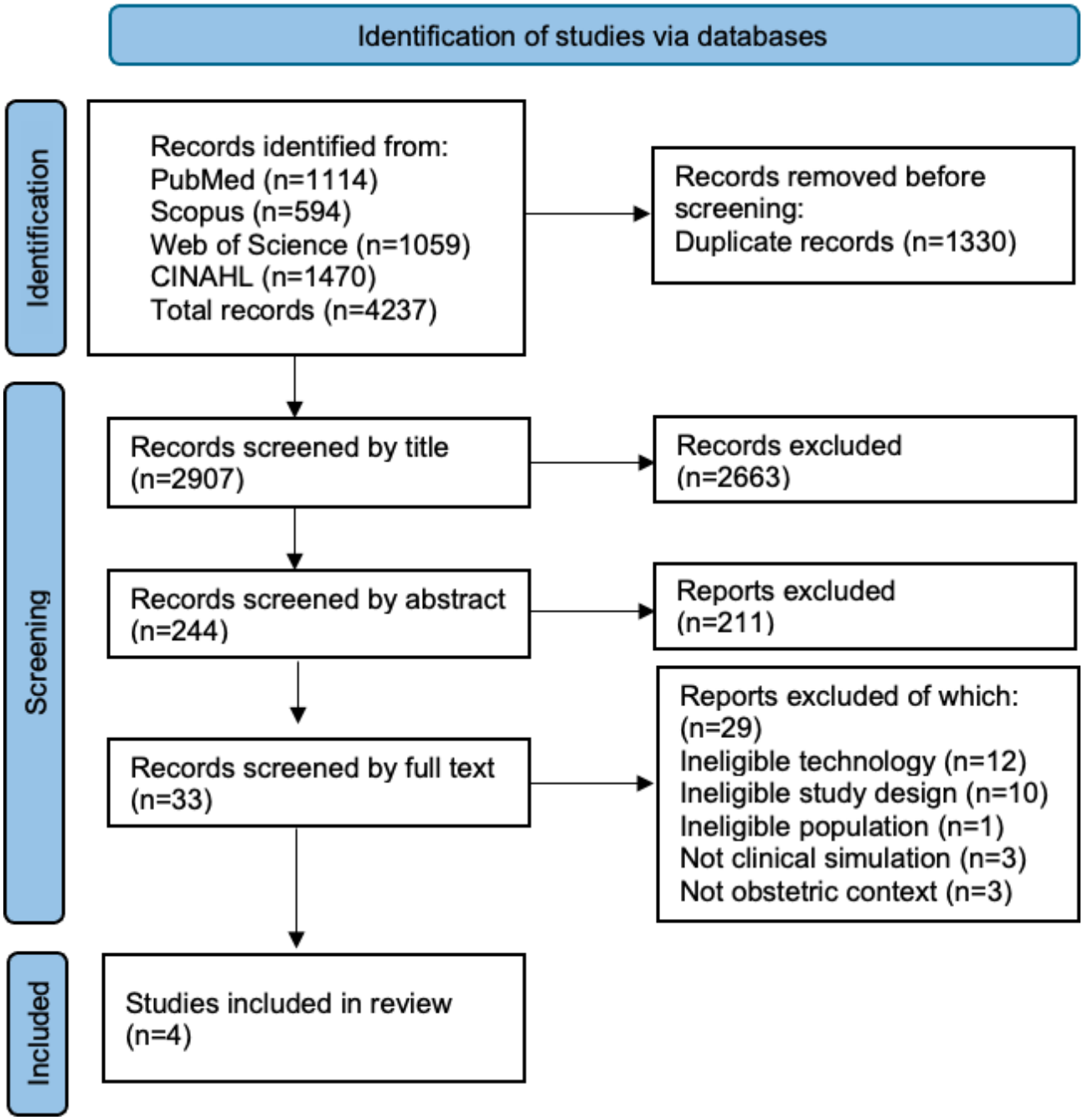
Results

The database search retrieved 4237 records. After the removal of 1330 duplicates, 2907 titles were screened, resulting in 244 records selected for abstract assessment. Of these, 33 full-text articles were evaluated, and 4 studies were included in the

XSL•FO
RenderX

review. Figure 1 presents the PRISMA flow diagram of the selection process.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the present study.



Characteristics of the Studies

The 4 included studies [24-27] were published between 2021 and 2024 and represented diverse geographical and educational contexts. Three studies involved undergraduate midwifery students [24-26], while one multicenter study included certified midwives and nurse-midwives working in maternity settings [27]. Overall, the review comprised 634 participants, including 360 midwifery students and 274 practicing midwives or nurse-midwives.

The methodological designs included 2 RCTs: one with 6-month follow-up comparing the use of VR with instructional video or digital guide [27], and another comparing VR with low-fidelity clinical simulation [25]. In addition, one quasi-experimental study with a control group compared VR with traditional teaching methods [26], and another quasi-experimental study without a control group used pre- and postintervention measures to evaluate the impact of AR across different simulated clinical scenarios [24]. Detailed study characteristics are presented in Table 1.

XSL•FO
RenderX

Table . Study characteristics.

Study (year)	Country	Type of study	Population and sample	Study objectives	XR[a] Type	Simulation scenario
Vogel et al (2024) [24]	Germany	Quasi-experimental, without a control group, (pre- and posttest design)	Midwifery students, n=133	To investigate whether AR[b] simulation training influences midwifery students' subjective perceptions of knowledge, confidence, and practical skills in emergency situations.	AR	Emergency tocolysis, preparing a pregnant woman for cesarian section, neonatal resuscitation
Öner and Turfan (2024) [25]	Türkiye	RCT[c], (posttest only design)	Midwifery students, n=92	To compare the impact of VR[d] simulation with low-fidelity simulation in teaching newborn care to obstetric nursing students.	VR	Initial care of the newborn
Zhao et al (2024) [26]	China	Quasi-experimental, with a control group, (posttest only design)	Midwifery students, n=135	To evaluate the effect of case-based learning method with VR simulation technology on midwifery courses.	VR	Eutocia (normal birth), dystocia, umbilical cord prolapse, and neonatal asphyxia and resuscitation
Umoren et al (2021) [27]	Nigeria and Kenya	RCT, 6 months follow-up	Midwives and nurse-midwives, n=274	To assess the impact of VR simulations using electronic Helping Babies Breathe or video for the maintenance of neonatal resuscitation skills in health care workers in resource-scarce settings.	VR	Routine newborn care, initial resuscitation, and prolonged resuscitation with positive pressure ventilation

[a]XR: extended reality.

[b]AR: augmented reality.

[c]RCT: randomized controlled trial.

[d]VR: virtual reality.

Description of the Educational Intervention

All 4 studies [24-27] investigated educational interventions that used clinical simulation mediated by XR technologies, with the aim of developing clinical competencies in obstetric or immediate neonatal care.

The simulated clinical contexts included: preparation of emergency tocolytics and preparation of the pregnant woman for cesarean section [24]; eutocic delivery, dystocia, and umbilical cord prolapse [26]; immediate newborn care [25]; and neonatal asphyxia and resuscitation [24,26,27]. Neonatal care was a common component in all 4 programs analyzed, consistently simulated in the context of immediate newborn care, immediately after birth.

Of the 4 studies included, 3 presented multiple simulation scenarios within the same intervention using VR [26,27] or AR [24]. None of the included studies provided detailed information about the physical simulated environment used in the interventions.

Technology Used

The included interventions used technically distinct XR configurations, with VR being more common than AR. Based on the operational definition used in this review, 3 studies were classified as VR because they used headset-based immersive delivery and first-person task-based interaction within simulated clinical scenarios [25-27]. Two of these used head-mounted display (HMD)–based VR with handheld controllers [25,26], while one used a low-cost smartphone-based HMD with a VR application [27]. One study was classified as AR because it used a smartphone-based marker system to overlay and manipulate virtual elements within the real-world environment [24]. The main technical characteristics of the interventions are summarized in Table 2.

Across studies, interaction was generally described as first-person and task-based, requiring participants to perform

XSL•FO
RenderX

clinical actions or manipulate virtual elements within simulated obstetric or neonatal scenarios [24-27]. However, the level of interactivity was not consistently detailed. In particular, information on free navigation, branched decision-making, customization of the virtual environment, and learner control was absent or limited.

Reporting of exposure and technical design features was also incomplete. Although some studies reported session duration or total usage time, the number of sessions and degree of learner autonomy were often unclear. Similarly, visual fidelity, spatial realism, software licensing, reusability, and open-access availability were not described. This limited reporting restricts comparison of the technological complexity and pedagogical affordances of the XR systems used across studies.

Table . Technical characteristics of XR[a] interventions.

Study	XR type	Device (hardware)	Software	Interaction	Duration and exposure	Presence and usability	Adverse effects
Vogel et al (2024) [24]	AR[b]	Smartphone +marker	Heb@AR	Active task-based interaction through manipulation of the virtual elements	Not reported	Not reported	Not reported
Öner and Turfan (2024) [25]	VR[c]	HMD[d] (Oculus Quest)+handheld controllers	Unity-based VR simulator	Active task-based interaction through manipulation of the virtual environment	Up to 10 minutes per session	Presence assessed; Usability not reported.	Dizziness and nausea
Zhao et al (2024) [26]	VR	HMD+handheld controllers	CBL-VR[e] platform	Active task-based interaction through manipulation of the virtual environment and feedback	Not reported	Not reported	Mild cybersickness (not quantified)
Umoren et al (2021) [27]	VR	Low-cost HMD+smartphone	eHBB[f] VR app	Active task-based interaction through manipulation of the virtual environment	101 minutes (IQR 81 - 120) over 6-month follow-up	Usability and acceptability assessed; presence not reported.	Not reported

[a]XR: extended reality.

[b]AR: augmented reality.

[c]VR: virtual reality.

[d]HMD: head-mounted display.

[e]CBL-VR: case-based learning virtual reality.

[f]eHBB: electronic Helping Babies Breathe.

Theoretical and Pedagogical Frameworks

Two studies [25,26] reported an explicit pedagogical or theoretical basis for their immersive simulation interventions. Zhao et al [26] integrated case-based learning (CBL) with VR simulation to support clinical reasoning, while Öner and Turfan [25] referred to the Jeffries Simulation Theory [28] as a framework for the design, implementation, and evaluation of the simulation. The remaining studies did not report an explicit theoretical model or instructional framework [24,27].

Across the included studies, the interventions primarily targeted the development of technical competencies related to obstetric or neonatal care, particularly procedures associated with childbirth, obstetric emergencies, immediate newborn care, and neonatal resuscitation [24-27]. Nontechnical or behavioral competencies were less frequently addressed; only Zhao et al [26] explicitly included communication, teamwork, and clinical decision-making.

Although Bloom's Taxonomy [18] was not explicitly used by the primary studies, the described activities could be mapped mainly to the cognitive and psychomotor domains. All studies involved the application of knowledge in simulated clinical contexts and the performance of obstetric or neonatal skills [24-27]. Higher-order cognitive processes, such as case analysis and decision-making, as well as affective elements related to teamwork, were identified only in Zhao et al [26].

Similarly, none of the studies explicitly adopted the Kirkpatrick model [19] as an evaluative framework. However, based on the reported outcomes, all studies assessed level 2 outcomes, including knowledge, practical skills, and/or attitudes. Three studies also assessed level 1 outcomes, such as satisfaction or perceptions of the learning experience [24-26]. No study evaluated level 3 outcomes or level 4 outcomes.

Reporting of instructional design and simulation best-practice components was limited. None of the studies described the use of a structured instructional design model, although 2 cited the

XSL•FO
RenderX

Healthcare Simulation Standards of Best Practice [29] as a guiding reference for simulation development and implementation [24,25]. Structured pedagogical briefing was not clearly reported. Instead, most studies described technical orientation to the technology, such as guidance on the use of VR or AR devices [24,26,27].

Debriefing was also inconsistently described. Only Öner and Turfan [25] reported a structured debriefing process, supported by the Jeffries Simulation Theory [28] and based on audiovisual review of the simulation. Other studies described less structured forms of feedback, including automated feedback provided by the application [24], feedback during the activity [26,27], or facilitator-mediated group feedback after the activity [26].

In all included studies, simulation was conducted individually. In one study, the facilitator was described as playing an active role providing guided instruction during the simulation [25]; in another, the activity was conducted without supervision, characterizing autonomous simulation [27]. In the remaining studies, no information was provided about the facilitator's role during the simulation [24,26].

Formal validation of scenario design was not reported in any study. One study used a validated clinical algorithm to structure the expected sequence of actions [27], while others reported that scenarios were based on clinical practice situations and reviewed by faculty members or experts [25,26].

Curricular integration was explicitly described only by Vogel et al [24], who used AR simulation as a complement to theoretical and practical obstetric teaching. Other studies did not specify whether the interventions were embedded in formal curricula. Across studies, immersive simulation was combined with additional teaching strategies, including preparatory reading, digital guide, lectures, mannequin-based teaching, clinical skills training, team-based role-play, and digital learning materials [24-27].

Instruments and Assessment

The included studies used different assessment instruments, which were distributed across the domains of knowledge, skills, and attitudes, in addition to addressing aspects related to the learning experience.

Theoretical knowledge was assessed using instruments constructed by the authors themselves based on the objectives of the intervention [24] or developed within the context of international educational programs [27].

Clinical skills were evaluated through checklists applied during the simulation [25,26], in Objective Structured Clinical Examination (OSCE) stations [27], and through online questionnaires of self-perceived practical performance [24].

Attitude was assessed in 2 studies, through validated self-administered scales that measured self-confidence [25] and self-directed learning ability [26].

Learning experiences and perceptions of the simulated environment were addressed differently across studies. Only Öner and Turfan [25] used validated scales to measure satisfaction, immersion perception, scenario design, and pedagogical practices. In the others, nonvalidated questionnaires were applied to assess aspects such as satisfaction [26], usability, realism, and clinical applicability [27].

Outcomes of the Educational Interventions

The included studies reported outcomes related to knowledge, practical skills, attitudes, and learning experience. Overall, findings suggested positive effects of VR- and AR-based interventions, particularly for practical skills, self-confidence, satisfaction, and perceived learning. However, the heterogeneity of outcome measures, assessment instruments, and follow-up periods limited direct comparison across studies.

Knowledge outcomes were assessed in 2 studies [24,27], with mixed findings. Vogel et al [24] reported significant improvements in perceived knowledge after AR simulation across all simulated clinical scenarios. In contrast, Umoren et al [27] found no statistically significant differences between groups in theoretical knowledge following an in-person course, although participants had only been exposed to the VR application for technical familiarization at that stage.

Practical skills were assessed in all 4 studies. Three studies reported significant improvements or higher performance associated with immersive simulation [24-26]. The strongest differences were observed when VR was combined with CBL, particularly for team-based skills [26], and when VR was compared with low-fidelity simulation for newborn care training [25]. In contrast, Umoren et al [27] did not identify statistically significant differences between groups in immediate or long-term skills performance, although the VR group showed a trend toward better retention at follow-up. Detailed findings for practical skills are presented in Table 3. As Vogel et al [24] did not include a control group, their results were not included in this analysis.

Attitudinal outcomes were assessed in 3 studies. Improvements were reported in self-confidence after AR simulation [24] and VR simulation [25], while Zhao et al [26] found gains in self-directed learning, including self-management, information processing, and collaboration, compared with traditional lecture-based teaching.

Learning experience was generally evaluated positively across studies. Participants reported favorable perceptions of satisfaction, usability, realism, clinical usefulness, immersion, and educational value [24-27]. However, validated instruments were used inconsistently, limiting comparison across interventions. Reported implementation challenges included mobile device compatibility issues, AR marker detection failures, dependence on project-provided smartphones, low adherence to recommended practice, and symptoms such as dizziness, nausea, or mild cybersickness [24-27].

XSL•FO
RenderX

Table . Instruments for data collection and observed outcomes related to the practical skills domain in the studies included in the systematic review.

Study and measurement instrument	Outcome of interest	Control group	Intervention group	Mean difference[a]	*P* value	Effect size (Cohen *d*)
Zhao et al [26]						
Academic performance evaluation form	Individual skill: normal birth, dystocia, umbilical cord prolapse, and neonatal resuscitation	89.24 (3.15)[b]	90.88 (2.14)[b]	+1.64	<.001	0.63
Academic performance evaluation form	Team-based skill: normal birth, dystocia, umbilical cord prolapse, and neonatal resuscitation	81.28 (5.45)[b]	90.97 (2.33)[b]	+9.69	<.001	2.42
Academic performance evaluation form	Case analysis: normal birth, dystocia, umbilical cord prolapse, and neonatal resuscitation	86.70 (2.56)[b]	88.64 (3.19)[b]	+1.94	<.001	0.66
Öner and Turfan [25]						
Practice skills list	Initial newborn care	74.17 (13.55)[b]	83.28 (19.67)[b]	+9.11	.01	0.54
Umoren et al [27]						
OSCE[c] A	Routine newborn care and initial resuscitation	72%[d]	76%[d]	+3.8%	.61	—[e]
OSCE B	Prolonged neonatal resuscitation	49%[d]	62%[d]	+13.2%	.09	—[e]
BMV[f] Skill check	Bag-mask ventilation in the newborn	22%[d]	28%[d]	+5.5%	.48	—[e]

[a]Mean difference represents the difference between the intervention and control groups.

[b]Mean (SD) of students' scores (scores ranging from 0 - 100).

[c]OSCE: Objective Structured Clinical Examination.

[d]Pass rate (%). Percentages are reported as presented in the original studies.

[e]Effect size not reported in the original study.

[f]BMV: bag-and-mask ventilation.

Methodological Quality and Risk of Bias

Overall, the included studies presented low to moderate risk of bias, with methodological limitations mainly related to blinding, allocation concealment, outcome measurement, and control of confounding factors. These limitations are common in educational intervention research but should be considered when interpreting the findings.

The 2 randomized controlled trials reported randomization, baseline comparability between groups, appropriate outcome assessment methods, and adequate statistical analyses [25,27]. However, allocation concealment was not described, and blinding of participants, intervention providers, or outcome assessors was not implemented. In addition, Öner and Turfan [25] did not report an intention-to-treat analysis, and the use of self-assessment measures may have increased the risk of measurement bias.

The 2 quasi-experimental studies reported a clear temporal relationship between intervention and outcomes, complete follow-up, and appropriate statistical analysis [24,26]. However, only Zhao et al [26] included a control group and reported baseline similarity between groups. In both quasi-experimental studies, limited information was provided on strategies to standardize data collection or ensure consistency among assessors. Vogel et al [24], which used a single-group pre-post design, did not report procedures to control for potential cointerventions.

MERSQI scores ranged from 10.0 to 14.5, with a mean score of 12.9 (SD 2.0), indicating moderate methodological quality overall. Higher scores were associated with stronger study designs, objective outcome measures, and appropriate statistical analyses. Lower scores mainly reflected limited evidence of instrument validity, single-institution sampling in most studies, and the absence of higher-level educational outcomes. Detailed MERSQI scores are presented in Multimedia Appendix 2.

XSL•FO
RenderX

Given the small number of included studies and the heterogeneity in interventions, outcome domains, assessment instruments, and timing of evaluation, a formal GRADE (Grading of Recommendations Assessment, Development, and Evaluation) assessment was not conducted. Instead, the overall confidence in the evidence base was assessed narratively, considering study design, risk of bias (JBI), methodological quality (MERSQI), consistency and directness of findings, and the use of validated outcome measures. Reporting bias could not be formally assessed using statistical methods, such as funnel plots, due to the limited number of studies and the absence of pooled effect estimates.

Discussion

Overview

This systematic review examined how VR and AR have been integrated into obstetric nursing and midwifery education, with particular attention to theoretical grounding, instructional design, technological characteristics, and educational outcomes. The findings indicate that immersive technologies have been used mainly to support the development of technical competencies in obstetric and neonatal care, particularly in childbirth procedures, emergency scenarios, immediate newborn care, and neonatal resuscitation. However, the evidence base remains limited, consisting of only 4 heterogeneous studies, and the pedagogical and technological reporting of the interventions was often incomplete.

In general, the included studies suggest that VR- and AR-based interventions may improve immediate learning outcomes, particularly practical skills, self-confidence, satisfaction, and perceived learning. However, the incorporation of theoretical frameworks and instructional design was limited, and these findings should therefore be interpreted with caution.

Theoretical and Pedagogical Grounding

A central finding of this review was the limited explicit use of theoretical or pedagogical frameworks.

Only one study referred to the Jeffries Simulation Theory [28], and another integrated CBL as a pedagogical strategy [26]. The remaining studies did not clearly describe the theoretical assumptions guiding the design, implementation, or evaluation of the XR-based interventions. This is relevant because, in technology-enhanced simulation-based learning, conceptual clarity is essential to strengthen pedagogical coherence, methodological rigor, and reproducibility [5]. Theories provide the conceptual foundation for selecting instructional strategies and organizing learning experiences according to learners' prior knowledge, motivation, and educational needs [30,31], while the instructional models translate these theoretical foundations into practical teaching strategies.

This finding is consistent with previous reviews showing that the absence or superficial use of theoretical frameworks remains a recurrent limitation in simulation and immersive technology research in nursing education [9,10,32]. Although constructivist and experiential approaches, such as Kolb Experiential Learning Theory [33], are frequently cited in VR simulation and technology-mediated nursing education [10,34], the studies included in this review rarely made explicit how learning theory informed scenario design, facilitation, feedback, or assessment. As a result, the interventions often appeared primarily technology, or procedure, driven rather than theory-informed.

Recent models specific to immersive learning, including the Cognitive Affective Model of Immersive Learning (CAMIL) [35], the Immersive Virtual Reality Pedagogical Model (iVRPM) [36], and embodied approaches to immersive extended reality [37], are useful for interpreting this gap. These models highlight constructs such as presence, agency, interactivity, feedback, embodiment, and cognitive engagement as mechanisms through which XR may support learning. However, in the included studies, these mechanisms were seldom assessed or clearly operationalized. For example, presence was measured in only one study [25], while constructs such as agency, cognitive load, usability, embodiment, and learner control were either not evaluated or insufficiently reported. Therefore, although these models offer relevant conceptual guidance, their application in obstetric nursing and midwifery XR education remains limited and requires further empirical validation.

The review also identified a lack of nursing or midwifery theoretical models oriented toward clinical practice. None of the included studies incorporated frameworks addressing comprehensive, relational, woman- and family-centered care, despite the relevance of these dimensions to nurse-midwives' professional practice [38]. This omission may contribute to a predominantly technicist approach to XR-based education, focused mainly on procedural performance and emergency management, while giving less attention to communication, emotional support, shared decision-making, and professional values [39]. Continuous exposure to the theoretical foundations of the profession can help students recognize how these principles apply in real care contexts [40] and may support patient-centered decision-making and therapeutic interventions [41,42]. Nevertheless, future XR interventions in obstetric nursing and midwifery should more explicitly integrate disciplinary theories to ensure that immersive simulation supports not only technical competence but also holistic and person-centered care.

Instructional Design Models

Instructional design was another area with limited explicit reporting across the included studies. Although some interventions incorporated pedagogical strategies such as CBL, repeated practice, preparatory activities, and feedback, none clearly described the instructional principles or models guiding the organization, sequencing, or progression of the learning experience. This limits the interpretation of how VR or AR was integrated into the educational process and reduces the replicability of the interventions in other contexts [30].

This gap is relevant because immersive technologies should not function as isolated or decontextualized learning activities, but as components of a broader instructional plan aligned with learning objectives, learner characteristics, and expected outcomes [30,43] as their effectiveness depends on the quality of instructional design [44]. Instructional design models such as Gagné's 9 events of instruction [45], Analysis, Design, Development, Implementation, and Evaluation (ADDIE)[46],

XSL•FO
RenderX

and Four-Component Instructional Design (4C/ID) [47] have been successfully applied in health and nursing education, including in obstetric emergency training and complex educational programs [48-50]. However, none of the studies included in this review explicitly adopted these or other structured instructional design models.

Although 2 studies [24,25] reported using the International Nursing Association for Clinical Simulation and Learning (INACSL) Healthcare Simulation Standards of Best Practice [29], their operationalization was only partially described, with key components such as learning objectives, briefing, facilitation, scenario design, feedback, and debriefing inconsistently reported.

The mapping of the interventions to Bloom's Taxonomy [18] showed that most studies addressed cognitive and psychomotor domains, mainly through the application of knowledge and performance of clinical procedures. However, the taxonomy was not explicitly used in instructional planning. This limits understanding of whether the degree of immersiveness and interactivity was intentionally aligned with different levels of cognitive complexity, as recommended in technology-mediated learning environments [36,51]. Similarly, although the Kirkpatrick model [19] was useful for interpreting evaluation levels, it was not explicitly adopted by the primary studies. All studies assessed learning outcomes, and some assessed satisfaction or perceptions, but none evaluated transfer to clinical practice or organizational or patient-related outcomes, reflecting a recurrent limitation in health professions education research [52].

The limited description of preparatory activities, briefing, and debriefing is particularly important in XR-based simulation. From the perspective of Cognitive Load Theory [53], these components may help manage the cognitive demands of immersive environments by preparing learners for the task, reducing unnecessary ambiguity, and supporting reflection after the experience. All studies described some form of preparation, such as theoretical classes, preparatory reading, technical instructions, practical training, or mannequin-based activities [24,26]. These strategies may have helped students build prior knowledge and reduce task complexity [31,54]. However, structured pedagogical briefing was rarely reported; most studies focused on technical familiarization with VR or AR devices rather than on broader briefing elements such as learning objectives, roles, expectations, psychological safety, confidentiality, and the fiction contract [10,55].

Debriefing was also insufficiently described. Only one study reported a structured debriefing process supported by a theoretical framework [25]. Another combined automated feedback with facilitator-guided discussion, suggesting a potentially useful hybrid approach [26]. Evidence indicates that debriefing and immediate feedback can strengthen clinical competence in VR simulation by supporting reflection, repetition, and individualized learning [56]. Automated feedback may therefore have pedagogical value when it is intentionally designed and aligned with learning objectives, but it should not be considered a substitute for facilitator-led reflective debriefing [57].

Pedagogical Implications of Immersiveness and Interactivity in Learning

The reviewed interventions used immersive and interactive technologies to support experiential learning in obstetric and neonatal scenarios. VR interventions generally relied on HMDs to create 3D environments, while the AR intervention used smartphone-based interaction. These features may enhance presence, realism, active participation, and knowledge retention when they are pedagogically aligned with the learning objectives [35,36,58].

However, the included studies provided limited information on how the level and type of immersion were selected or justified. In simulation-based education, immersion should not be understood only as physical or technological realism. Rather, different dimensions of fidelity may support different learning goals. Physical fidelity refers to the extent to which the simulation resembles the real clinical environment in sensory or material terms, whereas functional fidelity concerns the extent to which the simulation reproduces the relevant clinical tasks, decision pathways, and responses. Psychological fidelity refers to the extent to which the scenario engages learners cognitively and emotionally, encouraging them to think and act as they would in clinical practice. In this sense, high physical realism may be useful for spatial orientation, procedural familiarization, and anatomical understanding, while functional and psychological fidelity may be more important for clinical reasoning, diagnostic decision-making, communication, and prioritization in complex obstetric situations [59-61].

This distinction is particularly relevant for XR-based obstetric education. For lower-level cognitive objectives, such as remembering or understanding, highly immersive or highly interactive environments may not be necessary. Greater interactivity and psychological engagement may be more appropriate when learners are expected to apply knowledge, analyze clinical cues, make decisions, or coordinate care in emergency scenarios [38]. Thus, the educational value of VR and AR depends less on maximizing realism in all dimensions and more on aligning the type of fidelity with the intended competencies.

The findings also suggest that immersive design should be considered in relation to cognitive load. High levels of sensory stimulation, realism, or interaction may support engagement, but they may also increase extraneous load if learners are insufficiently prepared or if interface complexity distracts from the clinical task [11]. Although some studies described technical familiarization or preparatory activities, none explicitly reported how physical, functional, or psychological fidelity, interface design, or task complexity were adjusted to learner proficiency or cognitive demands.

Among the studies analyzed, only Öner and Turfan [25] assessed the sense of presence, reporting high perceived immersion and a positive relationship with students' performance and satisfaction. Zhao et al [26] also reported positive perceptions related to satisfaction and self-confidence. However, Umoren et al [27] found low adherence over time despite good usability, suggesting that ease of use and physical immersion alone may not ensure sustained engagement or motivation [58].

XSL•FO
RenderX

Although all studies described first-person interaction and the performance of clinical actions in 3D environments, none detailed the technical characteristics of the systems used, such as visual fidelity, navigation freedom, or decision-making mechanisms. Conversely, research indicates that these features of VR and AR are particularly effective for teaching complex anatomical and spatial content, with positive results also observed in obstetrics, especially due to the ability to visualize the dynamic relationship between the fetus and the maternal pelvis [1,3,8]. In addition, the incorporation of haptic resources and tactile, auditory, or visual feedback can enhance the sense of realism and encourage learners to reflect on their actions, thereby fostering the development of clinical reasoning. However, when poorly calibrated, these additional stimuli may cause distraction or sensory overload [11].

The technical diversity of the included interventions also indicates that the findings should not be interpreted as reflecting a homogeneous "XR effect." HMD-based VR, smartphone-based HMD VR, and smartphone-based AR differ in the type and degree of immersion they provide, as well as in their potential to support presence, spatial understanding, embodied interaction, and learner control. Therefore, as immersive technologies continue to expand in health professions education, there is a growing need to better understand the mechanisms through which learning occurs in these environments. Beyond the assessment of learning outcomes, future research should incorporate key constructs such as presence, agency, and cognitive load, assessed using validated instruments (eg, the Multidimensional Cognitive Load Scale for Virtual Environments [62], the Presence Questionnaire [63], and NASA-TLX [64]), as well as usability measures (eg, virtual reality system usability questionnaire [65]), to provide a more comprehensive understanding of XR-based learning processes.

Educational Effects of VR- and AR-Based Interventions

The included studies reported generally favorable educational outcomes for VR- and AR-based interventions, particularly in relation to practical skills, self-confidence, satisfaction, and perceived learning. However, these findings should be interpreted cautiously because the evidence base was small, heterogeneous, and mainly limited to short-term educational outcomes.

Assessment strategies varied considerably across studies, with a predominance of measures related to technical skills, learning experience, and perceived self-confidence. Most studies assessed outcomes immediately after the intervention, limiting conclusions about retention and transfer to clinical practice. Only Umoren et al [27] included follow-up assessment, suggesting a possible benefit of VR for the retention of psychomotor competencies. However, the absence of longitudinal assessment in the remaining studies prevents broader conclusions about the sustainability of learning effects.

The reviewed studies were concentrated at Kirkpatrick level 2, as they assessed learning outcomes such as knowledge, skills, and attitudes. Some studies also addressed level 1 outcomes, including satisfaction and perceptions of the learning experience. However, none evaluated level 3 (behavioral change in clinical practice) or level 4 outcomes (organizational, clinical, or patient-level impact) [19]. This distinction is important because improvements observed in simulated learning environments should not be interpreted as direct evidence of improved clinical performance or patient outcomes. Capturing higher-level Kirkpatrick outcomes is particularly challenging in obstetric emergency education. Events such as neonatal resuscitation, dystocia, umbilical cord prolapse, or emergency cesarean preparation may be high-risk but relatively infrequent, making it difficult to observe sufficient real-world cases after training. This limitation is consistent with previous reviews on virtual reality [66] and reflects methodological challenges inherent to clinical training, where assessing behavioral change in real practice and linking training to patient outcomes require longitudinal follow-up and are influenced by multiple confounding factors [52]. This may result in imprecise evidence, as observed in reviews of simulation in obstetric emergencies [67], reinforcing the need for caution when interpreting and measuring these outcomes.

Most interventions focused primarily on psychomotor skills and basic cognitive outcomes, especially the understanding and application of technical procedures. This focus is consistent with previous evidence suggesting that immersive technologies may support cognitive, psychomotor, and affective outcomes among nursing students and nurses [2,68], with promising applications in obstetric education [69]. However, higher-order cognitive processes were less consistently addressed. Zhao et al [26] was the only study that clearly stimulated clinical reasoning and decision-making through a CBL approach, integrating case analysis, diagnosis formulation, and procedural execution. This finding aligns with evidence associating VR with critical thinking and cognitive engagement among nursing students when it is embedded in intentional instructional design [12,70].

Despite the promising results in the cognitive and psychomotor domains, Zhao et al [26] was the sole study to address aspects of communication and teamwork, which were evaluated in the context of collaborative clinical case resolution. Recent evidence broadens this discussion by demonstrating that the use of immersive virtual environments can promote significant gains in communication, leadership, situational awareness, teamwork, and decision-making [71], particularly when there is a high degree of realism and interactivity [5,71]. These are fundamental competencies for safe clinical performance in obstetric contexts. Nevertheless, the literature indicates that the development and structured assessment of these competencies in VR and AR–based programs remain incipient, representing an important challenge for future research in obstetric and nursing education [32].

The included studies predominantly focus on procedural skills, with limited attention to the relational dimensions of care, revealing an important imbalance in midwifery training, which requires the integration of clinical, emotional, and interpersonal components [38].

XSL•FO
RenderX

Implications and Recommendations for Educational Practice

The integration of simulation mediated by VR and AR in obstetric education requires coherence among theoretical foundations, instructional design, technological features, and expected outcomes. Immersive technologies should be used as pedagogical tools within a structured educational program, rather than as isolated technological experiences. This recommendation is consistent with the recent Utstein-style consensus agenda for XR in health care simulation, which emphasizes evidence-informed adoption, faculty readiness, institutional investment, data-driven evaluation, and sustainable integration of XR technologies in health professions education [72].

Considering the context of obstetric nursing and midwifery education, nursing theories can guide the principles of comprehensive and person-centered care, reinforcing the ethical and disciplinary commitments of the nursing and midwifery professions. These values should permeate the construction and implementation of simulation scenarios, ensuring that students develop clinical judgment, ethical awareness, and decision-making grounded in evidence and in the needs expressed by women and their families.

In instructional planning, it is recommended that VR and AR be integrated into the curriculum as part of a continuum connecting theory, simulated practice, and clinical practice. To achieve this, it is suggested to use instructional design models that intentionally and systematically integrate different teaching strategies, including immersive technologies, grounded in the articulation between pedagogical theory and the expected learning outcomes.

The level of immersiveness and interactivity should be adjusted according to the cognitive levels involved and the learners' proficiency, balancing realism and cognitive load. Less immersive environments are indicated when the focus is on understanding concepts and initial procedural practice, whereas highly immersive simulations are recommended for learning objectives involving clinical reasoning and decision-making in complex obstetric contexts. Furthermore, obstetric education can leverage the potential of VR and AR to represent dynamic anatomical and physiological structures, such as the relationship between the fetus and the maternal pelvis during labor, thereby enhancing spatial understanding and clinical reasoning, provided that the simulation design avoids sensory distractions and maintains intentionality in all elements of the scenario.

Clinical simulation mediated by immersive technologies should follow international guidelines for best practices in clinical simulation, ensuring both pedagogical and ethical quality. Preliminary preparation is recommended to reduce intrinsic cognitive load through familiarization with the technology and the learning content. It is important to consider briefing as an essential stage for reducing extraneous cognitive load by aligning objectives and expectations, while facilitation should monitor and adjust cognitive load during the simulation, providing support compatible with the learner's proficiency level and encouraging autonomy.

Automated sensory feedback from virtual platforms can be useful for validating actions or signaling the need for adjustments in clinical reasoning, and when complemented by facilitator-mediated feedback, it promotes reflective learning and deeper understanding. During debriefing, conducted in a structured manner, it is important to include reflection on technical and behavioral dimensions related to woman- and family-centered care.

The evaluation of simulations mediated by VR or AR should encompass multiple dimensions and outcome levels, in alignment with the New World Kirkpatrick Model [19] and the cognitive, psychomotor, and affective domains. Evaluation should occur at different points in time, monitoring the retention and transfer of skills to obstetric care settings, and the results should provide feedback for the continuous improvement of the design, facilitation, and curricular integration of immersive technologies.

Limitations of the Review Process

In addition to the limitations of the included studies, some limitations of the review process should be acknowledged. The small number of included studies and their substantial heterogeneity limited direct comparisons across interventions. Although the included studies reported generally favorable educational outcomes, the evidence reflects considerable variation in technologies, pedagogical approaches, outcome measures, and assessment time points.

Furthermore, the synthesis was constrained by the information reported in the primary studies, precluding formal assessment of reporting bias and certainty of evidence. Consequently, the overall confidence in the evidence is limited. Most outcomes were restricted to satisfaction, perceived learning, knowledge, or skills, with no evidence on transfer to clinical practice or impact on organizational or patient outcomes.

Accordingly, these findings should be interpreted as preliminary, despite their potential implications for midwifery and obstetrics education.

Conclusion

Evidence suggests that VR and AR may support learning in obstetric nursing and midwifery education, particularly in relation to practical skills, self-confidence, satisfaction, and perceived learning. However, the overall certainty of the evidence remains limited because of the small number of heterogeneous studies.

Stronger integration of learning theories, instructional design models, simulation best-practice standards, and validated assessment strategies is needed. Future XR-based interventions should be pedagogically intentional, transparently reported, and aligned not only with technical competence but also with the relational, ethical, and woman-centered nature of obstetric and midwifery care.

XSL•FO
RenderX

Acknowledgments

The authors acknowledge the Instituto de Ciências Biomédicas Abel Salazar (ICBAS), University of Porto, for financial support toward the article publication fee. MLF also acknowledges the Federal University of Ceará (UFC) for the institutional support that facilitated the preparation of this manuscript.

During manuscript preparation, the authors used ChatGPT (OpenAI) to assist with language editing, translation, and improving the clarity of the manuscript. All AI-assisted content was critically reviewed and verified by the authors, who take full responsibility for the final manuscript.

Funding

No specific funding was received for the conduct of this study.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: MLF.
Data curation: MLF (lead), APM (supporting), VF (supporting).
Formal analysis: MLF.
Validation: AC (lead), CSC (equal), APM (supporting).
Supervision: AC (lead), CSC (supporting).
Writing – original draft: MLF.
Writing – review & editing: MLF (lead), CSC (supporting), AC (supporting).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Detailed search strategy and eligibility criteria used in the systematic review.
[PDF File, 57 KB - nursing_v9i1e90252_app1.pdf]

Multimedia Appendix 2
Risk-of-bias assessment of included studies.
[DOCX File, 80 KB - nursing_v9i1e90252_app2.docx]

Checklist 1
PRISMA 2020 checklist for the systematic review.
[PDF File, 82 KB - nursing_v9i1e90252_app3.pdf]

References

1. Aasekjær K, Bjørnås B, Skivenes HK, Vik ES. Immersive virtual reality (VR) when learning anatomy in midwifery education: a pre-post pilot study. Eur J Midwifery 2024;8:39206326. [doi: 10.18332/ejm/191364] [Medline: 39206326]
2. Efendi D, Apriliyasari RW, Prihartami Massie JGE, et al. The effect of virtual reality on cognitive, affective, and psychomotor outcomes in nursing staffs: systematic review and meta-analysis. BMC Nurs 2023 May 19;22(1):170. [doi: 10.1186/s12912-023-01312-x] [Medline: 37202768]
3. Jallad ST, Natsheh I, Helo LA, et al. Nursing student's perceptions, satisfaction, and knowledge toward utilizing immersive virtual reality application in human anatomy course: quasi-experimental. BMC Nurs 2024 Aug 28;23(1):601. [doi: 10.1186/s12912-024-02254-8] [Medline: 39198772]
4. Hamilton D, McKechnie J, Edgerton E, Wilson C. Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. J Comput Educ 2021 Mar;8(1):1-32. [doi: 10.1007/s40692-020-00169-2]
5. Stenseth HV, Steindal SA, Solberg MT, et al. Simulation-based learning supported by technology to enhance critical thinking in nursing students: scoping review. J Med Internet Res 2025 Feb 18;27:e58744. [doi: 10.2196/58744] [Medline: 39965203]
6. Lin XP, Li BB, Yao ZN, Yang Z, Zhang M. The impact of virtual reality on student engagement in the classroom-a critical review of the literature. Front Psychol 2024;15:1360574. [doi: 10.3389/fpsyg.2024.1360574] [Medline: 38659670]

XSL•FO
RenderX

7. Hasyim AR, Syarif S, Ahmad M, Niswar M, Nasrudin AM. Enhance midwifery student skills about active management third stage labor via learning media. Gac Sanit 2021;35 Suppl 2:S284-S287. [doi: 10.1016/j.gaceta.2021.10.035] [Medline: 34929833]
8. Ryan G, Callaghan S, Rafferty A, et al. Virtual reality in midwifery education: a mixed methods study to assess learning and understanding. Nurse Educ Today 2022 Dec;119:105573. [doi: 10.1016/j.nedt.2022.105573] [Medline: 36206631]
9. Hernon O, McSharry E, MacLaren I, Dunne R, Carr PJ. The use of educational technology in undergraduate and postgraduate nursing and midwifery education: a scoping review. Comput Inform Nurs 2023 Mar 1;41(3):162-171. [doi: 10.1097/CIN.0000000000000928] [Medline: 35562320]
10. Koskinen I, Stolt M, Widmer CT, et al. Methodological approaches and competence areas of nursing students in virtual reality simulation research - a scoping review. Nurse Educ Today 2024 Feb;133:106033. [doi: 10.1016/j.nedt.2023.106033] [Medline: 37988799]
11. Parong J, Mayer RE. Cognitive and affective processes for learning science in immersive virtual reality. Comput Assist Learn 2021 Feb;37(1):226-241. [doi: 10.1111/jcal.12482]
12. Sim JJM, Rusli KDB, Seah B, Levett-Jones T, Lau Y, Liaw SY. Virtual simulation to enhance clinical reasoning in nursing: a systematic review and meta-analysis. Clin Simul Nurs 2022 Aug;69:26-39. [doi: 10.1016/j.ecns.2022.05.006] [Medline: 35754937]
13. Asoodar M, Janesarvatan F, Yu H, de Jong N. Theoretical foundations and implications of augmented reality, virtual reality, and mixed reality for immersive learning in health professions education. Adv Simul (Lond) 2024 Sep 9;9(1):36. [doi: 10.1186/s41077-024-00311-5] [Medline: 39252139]
14. Bauce K, Kaylor MB, Staysniak G, Etcher L. Use of theory to guide integration of virtual reality technology in nursing education: a scoping study. J Prof Nurs 2023;44(1-7):1-7. [doi: 10.1016/j.profnurs.2022.10.003] [Medline: 36746594]
15. JBI Manual for evidence synthesis. JBI. 2024. URL: https://synthesismanual.jbi.global [accessed 2026-09-09]
16. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
17. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
18. Anderson LW, Krathwohl DR. A taxonomy for learning, teaching, and assessing: a revision of bloom's taxonomy of educational objectives: Addison Wesley Longman, Inc; 2001.
19. Kirkpatrick JD, Kirkpatrick WK. Kirkpatrick's Four Levels of Training Evaluation: Association for Talent Development 2016. URL: https://books.google.com/books?id=mo--DAAAQBAJ [accessed 2026-09-14]
20. Barker TH, Stone JC, Sears K, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for randomized controlled trials. JBI Evid Synth 2023 Mar 1;21(3):494-506. [doi: 10.11124/JBIES-22-00430] [Medline: 36727247]
21. Barker TH, Habibi N, Aromataris E, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for quasi-experimental studies. JBI Evid Synth 2024 Mar 1;22(3):378-388. [doi: 10.11124/JBIES-23-00268] [Medline: 38287725]
22. Cook DA, Reed DA. Appraising the quality of medical education research methods: the medical education research study quality instrument and the Newcastle-Ottawa Scale-education. Acad Med 2015 Aug;90(8):1067-1076. [doi: 10.1097/ACM.0000000000000786] [Medline: 26107881]
23. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020 Jan 16;368:l6890. [doi: 10.1136/bmj.l6890] [Medline: 31948937]
24. Vogel K, Bernloehr A, Willmeroth T, Blattgerste J, Hellmers C, Bauer NH. Augmented reality simulation-based training for midwifery students and its impact on perceived knowledge, confidence and skills for managing critical incidents. Midwifery 2024 Sep;136:104064. [doi: 10.1016/j.midw.2024.104064] [Medline: 38905862]
25. Öner S, Turfan E. Virtual reality versus low level simulation in newborn care teaching in Turkey. Afr J Reprod Health 2024 Jul 31;28(7):54-60. [doi: 10.29063/ajrh2024/v28i7.6] [Medline: 39097974]
26. Zhao L, Dai X, Chen S. Effect of the case-based learning method combined with virtual reality simulation technology on midwifery laboratory courses: a quasi-experimental study. Int J Nurs Sci 2024 Jan;11(1):76-82. [doi: 10.1016/j.ijnss.2023.12.009] [Medline: 38352279]
27. Umoren R, Bucher S, Hippe DS, et al. eHBB: a randomised controlled trial of virtual reality or video for neonatal resuscitation refresher training in healthcare workers in resource-scarce settings. BMJ Open 2021 Aug 25;11(8):e048506. [doi: 10.1136/bmjopen-2020-048506] [Medline: 34433598]
28. Jeffries PR. The NLN Jeffries Simulation Theory: Wolters Kluwer; 2016. URL: https://books.google.co.in/books?id=yafAsgEACAAJ&redir_esc=y [accessed 2026-09-14]
29. Watts PI, McDermott DS, Alinier G, et al. Healthcare simulation standards of best practice simulation design. Clin Simul Nurs 2021 Sep;58:14-21. [doi: 10.1016/j.ecns.2021.08.009]
30. Frèrejean J, Walker K, Symon B, Eppich W. Insights from 40 years of educational research: honoring Jeroen van Merriënboer. Adv Simul (Lond) 2025 Mar 13;10(1):11. [doi: 10.1186/s41077-025-00332-8] [Medline: 40083027]
31. O'Leary F. How to deliver effective paediatric simulation based education. Paediatr Respir Rev 2024 Sep;51:10-18. [doi: 10.1016/j.prrv.2024.05.001] [Medline: 39179445]

XSL•FO
RenderX

32. Ropponen P, Tomietto M, Pramila-Savukoski S, et al. Impacts of VR simulation on nursing students' competence, confidence, and satisfaction: a systematic review and meta-analysis of randomised controlled trials. Nurse Educ Today 2025 Sep;152:106756. [doi: 10.1016/j.nedt.2025.106756] [Medline: 40279771]
33. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development: Prentice-Hall; 1984. URL: https://books.google.com/books?id=ufnuAAAAMAAJ [accessed 2026-09-14]
34. O'Connor S, Kennedy S, Wang Y, Ali A, Cooke S, Booth RG. Theories informing technology enhanced learning in nursing and midwifery education: a systematic review and typological classification. Nurse Educ Today 2022 Nov;118:105518. [doi: 10.1016/j.nedt.2022.105518] [Medline: 36030581]
35. Makransky G, Petersen GB. The Cognitive Affective Model of Immersive Learning (CAMIL): a theoretical research-based model of learning in immersive virtual reality. Educ Psychol Rev 2021 Sep;33(3):937-958. [doi: 10.1007/s10648-020-09586-2]
36. Bicalho DR, Piedade J, Matos JF. iVRPM: Conceptual proposal of an immersive virtual reality pedagogical model. Appl Sci 2025;15(4):2162. [doi: 10.3390/app15042162]
37. Fugate JMB, Tonsager MJ, Macrine SL. Immersive extended reality (I-XR) in medical and nursing for skill competency and knowledge acquisition: a systematic review and implications for pedagogical practices. Behav Sci (Basel) 2025 Apr 4;15(4):468. [doi: 10.3390/bs15040468] [Medline: 40282089]
38. Transitioning the midwifery models of care: global position paper. : World Health Organization; 2024.
39. Cruz MDS, Denucci MAM. Teaching principles and values in nursing from a philosophical perspective of care [Article in Portugese]. Rev Foco 2025;18(5):e8402. [doi: 10.54751/revistafoco.v18n5-113]
40. Norris Waller M, Newsome Wicks M. Is there still value in teaching nursing theory? J Nurs Educ 2021 Nov;60(11):603-604. [doi: 10.3928/01484834-20211007-02] [Medline: 34723738]
41. Bilgiç F, Bozkurt G. Effect of online infant care training and postpartum counseling based on Meleis' transition theory on mothers' readiness for care and breastfeeding: a randomized controlled trial. Clin Exp Pediatr 2024 Oct;67(10):521-530. [doi: 10.3345/cep.2024.00423] [Medline: 39327684]
42. Estakhri RK, Ezbarami ZT, Peyrovi H, et al. Effectiveness of transition theory-based ointerventions on outcomes of nursing care: a scoping review. Iran J Nurs Midwifery Res 2025;30(4):461-467. [doi: 10.4103/ijnmr.ijnmr_233_24] [Medline: 40832496]
43. Mallek F, Mazhar T, Faisal Abbas Shah S, Ghadi YY, Hamam H. A review on cultivating effective learning: synthesizing educational theories and virtual reality for enhanced educational experiences. PeerJ Comput Sci 2024;10:e2000. [doi: 10.7717/peerj-cs.2000] [Medline: 38855256]
44. Cook DA, Hamstra SJ, Brydges R, et al. Comparative effectiveness of instructional design features in simulation-based education: systematic review and meta-analysis. Med Teach 2013;35(1):e867-e898. [doi: 10.3109/0142159X.2012.714886] [Medline: 22938677]
45. Gagné RM. The Conditions of Learning and Theory of Instruction 1985. URL: https://books.google.com/books?id=c1MmAQAAIAAJ [accessed 2026-09-14]
46. Branch RM. Instructional Design: The ADDIE Approach: Springer; 2009. [doi: 10.1007/978-0-387-09506-6]
47. van Merriënboer JJG, Kirschner PA, Frèrejean J. Ten Steps to Complex Learning: A Systematic Approach to Four-Component Instructional Design: Taylor & Francis; 2024. URL: https://books.google.co.in/books?id=G4gYEQAAQBAJ&redir_esc=y [accessed 2026-09-14]
48. Daniel A, Nagandla K, Daniel S. Application of Gagné's instructional design in clinical simulation: enhancing learning in obstetric emergency management. Cureus 2025 Jan;17(1):e76845. [doi: 10.7759/cureus.76845] [Medline: 39758862]
49. Sim WH. Development and evaluation of a multifaceted intervention program for preventing medication administration errors by nurses. J Patient Saf 2025 Jun 1;21(4):226-239. [doi: 10.1097/PTS.0000000000001345] [Medline: 40214292]
50. Musharyanti L, Haryanti F, Claramita M. Improving nursing students' medication safety knowledge and skills on using the 4C/ID learning model. J Multidiscip Healthc 2021;14:287-295. [doi: 10.2147/JMDH.S293917] [Medline: 33574673]
51. Barari N, RezaeiZadeh M, Khorasani A, Alami F. Designing and validating educational standards for e-teaching in virtual learning environments (VLEs), based on revised Bloom's taxonomy. Interactive Learning Environments 2022 Oct 3;30(9):1640-1652. [doi: 10.1080/10494820.2020.1739078]
52. Miranda FD, Santos BD, Kristman VL, Mininel VA. Employing Kirkpatrick's framework to evaluate nurse training: an integrative review. Rev Latino-Am Enferm 2025;33. [doi: 10.1590/1518-8345.7250.4431]
53. Sweller J. Cognitive load theory and educational technology. Educ Technol Res Dev 2020 Feb;68(1):1-16. [doi: 10.1007/s11423-019-09701-3]
54. Lapierre A, Arbour C, Maheu-Cadotte MA, Vinette B, Fontaine G, Lavoie P. Association between clinical simulation design features and novice healthcare professionals' cognitive load: a systematic review and meta-analysis. Simul Gaming 2022 Oct;53(5):538-563. [doi: 10.1177/10468781221120599]
55. Persico L, Ramakrishnan S, Wilson-Keates B, et al. Healthcare simulation standard of best practice prebriefing: preparation and briefing. Clin Simul Nurs 2025 Aug;105:101777. [doi: 10.1016/j.ecns.2025.101777]
56. Cho MK, Kim MY. Enhancing nursing competency through virtual reality simulation among nursing students: a systematic review and meta-analysis. Front Med (Lausanne) 2024;11:1351300. [doi: 10.3389/fmed.2024.1351300] [Medline: 38774395]

XSL•FO
RenderX

57. Escobar-Castillejos D, Noguez J, Neri L, Magana A, Benes B. A review of simulators with haptic devices for medical training. J Med Syst 2016 Apr;40(4):104. [doi: 10.1007/s10916-016-0459-8] [Medline: 26888655]
58. Lee H. A unified conceptual model of immersive experience in extended reality. Comput Hum Behav Rep 2025 May;18:100663. [doi: 10.1016/j.chbr.2025.100663]
59. Davies AJ, Krame G. Measuring the level of fidelity required for transfer of learning in simulation-based learning exercises for novice and experienced practitioners. Simul Gaming 2024 Aug;55(4):685-715. [doi: 10.1177/10468781241241548]
60. Lavoie P, Deschênes MF, Nolin R, et al. Beyond technology: a scoping review of features that promote fidelity and authenticity in simulation-based health professional education. Clin Simul Nurs 2020 May;42:22-41. [doi: 10.1016/j.ecns.2020.02.001]
61. Hamstra SJ, Brydges R, Hatala R, Zendejas B, Cook DA. Reconsidering fidelity in simulation-based training. Acad Med 2014 Mar;89(3):387-392. [doi: 10.1097/ACM.0000000000000130] [Medline: 24448038]
62. Andersen MS, Makransky G. The validation and further development of a multidimensional cognitive load scale for virtual environments. J Comput Assist Learn 2021 Feb;37(1):183-196. [doi: 10.1111/jcal.12478]
63. Witmer BG, Singer MJ. Measuring presence in virtual environments: a presence questionnaire. Presence 1998 Jun;7(3):225-240. [doi: 10.1162/105474698565686]
64. Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. Adv Psychol 1988:139-183. [doi: 10.1016/S0166-4115(08)62386-9]
65. Kim YM, Rhiu I. Development of a virtual reality system usability questionnaire (VRSUQ). Appl Ergon 2024 Sep;119:104319. [doi: 10.1016/j.apergo.2024.104319] [Medline: 38797014]
66. Foronda CL, Gonzalez L, Meese MM, et al. A comparison of virtual reality to traditional simulation in health professions education: a systematic review. Simul Healthc 2024 Jan 1;19(1S):S90-S97. [doi: 10.1097/SIH.0000000000000745] [Medline: 37651101]
67. Fransen AF, van de Ven J, Banga FR, Mol BWJ, Oei SG. Multi-professional simulation-based team training in obstetric emergencies for improving patient outcomes and trainees' performance. Cochrane Database Syst Rev 2020 Dec 16;12(12):CD011545. [doi: 10.1002/14651858.CD011545.pub2] [Medline: 33325570]
68. Ryan GV, Callaghan S, Rafferty A, Higgins MF, Mangina E, McAuliffe F. Learning outcomes of immersive technologies in health care student education: systematic review of the literature. J Med Internet Res 2022 Feb 1;24(2):e30082. [doi: 10.2196/30082] [Medline: 35103607]
69. Visan IG, Toma CV, Petca R, Petrescu GED, Noditi AR, Petca A. Exploring the efficacy of virtual reality training in obstetric procedures and patient care—a systematic review. Health Care (Don Mills) 2025;13(7):784. [doi: 10.3390/healthcare13070784]
70. Gabr MAM, Sleem WF, El-Wkeel NS. Effect of virtual reality educational program on critical thinking disposition among nursing students in Egypt: a quasi-experimental pretest-posttest design. BMC Nurs 2025 Jul 7;24(1):874. [doi: 10.1186/s12912-025-03488-w] [Medline: 40624646]
71. Chan K, Kor PPK, Liu JYW, Cheung K, Lai T, Kwan RYC. The use of immersive virtual reality training for developing nontechnical skills among nursing students: multimethods study. Asian Pac Isl Nurs J 2024 Jul 10;8:e58818. [doi: 10.2196/58818] [Medline: 38986130]
72. Issenberg B, Ostergaard D, Matos F, et al. Setting a research agenda for the use of extended reality in healthcare simulation: an Utstein style meeting. Adv Simul (Lond) 2026 Mar 3;11(1):16. [doi: 10.1186/s41077-026-00409-y] [Medline: 41776623]

Abbreviations

4C/ID: Four-Component Instructional Design
ADDIE: Analysis, Design, Development, Implementation, and Evaluation
AR: augmented reality
CAMIL: Cognitive affective model of immersive learning
CBL: case-based learning
CINAHL: Cumulative Index to Nursing and Allied Health Literature
GRADE : Grading of Recommendations Assessment, Development, and Evaluation
HMD: head-mounted display
INACSL: International Nursing Association for Clinical Simulation and Learning
iVRPM: Immersive Virtual Reality Pedagogical Model
JBI: Joanna Briggs Institute
MERSQI: Medical Education Research Study Quality Instrument
OSCE: Objective Structured Clinical Examination
PIO : population, intervention, outcome
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO : International Prospective Register of Systematic Reviews
RCT: randomized controlled trial

XSL•FO
RenderX

SWiM : Synthesis Without Meta-Analysis
VR: virtual reality
XR: extended reality

Edited by E Borycki; submitted 25.Dec.2025; peer-reviewed by JF Costa, T Davidson; revised version received 15.Jun.2026; accepted 17.Jun.2026; published 05.Oct.2026.

Please cite as:
Firmiano ML, Moreira APA, Vieira FVM, Sa-Couto C, Cardoso A
Integration of Virtual and Augmented Reality Into Obstetric Nursing and Midwifery Education: Systematic Review
JMIR Nursing 2026;9:e90252
URL: https://nursing.jmir.org/2026/1/e90252
doi:10.2196/90252

© Mariana Luisa Firmiano, Ana Paula Assunção Moreira, Flaviana Vely Mendonça Vieira, Carla Sa-Couto, Alexandrina Cardoso. Originally published in JMIR Nursing (https://nursing.jmir.org), 5.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Perceptions and Intentions of Nursing Students Regarding Digital Health: Cross-Sectional Study

Alexandre Castonguay[1], PhD; Sandrine Hegg-Deloye[1], PhD; Guy Paré[2,3], PhD; Faustin Armel Etindele Sosso[2,4], DBA

[1]Faculté des sciences infirmières, Université de Montréal, 2375 chemin de la Côte-Sainte-Catherine, Montreal, QC, Canada

[2]Département de technologie de l'information, HEC Montréal, Montreal, QC, Canada

[3]Research Chair in Digital Health, HEC Montréal, Montreal, QC, Canada

[4]Université de Montréal, Montreal, QC, Canada

Corresponding Author:
Alexandre Castonguay, PhD
Faculté des sciences infirmières, Université de Montréal, 2375 chemin de la Côte-Sainte-Catherine, Montreal, QC, Canada

Related Article:

Correction of: https://nursing.jmir.org/2026/1/e77051

(JMIR Nursing 2026;9:e98552) doi:10.2196/98552

In "Perceptions and Intentions of Nursing Students Regarding Digital Health: Cross-Sectional Study" [1], the authors made one correction.

The file originally published as Multimedia Appendix 2 has been replaced with the correct questionnaire corresponding to this study, which is attached here as Multimedia Appendix 1.

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Multimedia Appendix 1
Study questionnaire assessing nursing students' perceptions, experiences, and attitudes toward digital health technologies (revised Multimedia Appendix 2).
[DOCX File, 50 KB - nursing_v9i1e98552_app1.docx]

Reference

1. Castonguay A, Hegg-Deloye S, Paré G, Etindele Sosso FA. Perceptions and intentions of nursing students regarding digital health: cross-sectional study. JMIR Nurs 2026 Mar 5;9:e77051. [doi: 10.2196/77051] [Medline: 41813428]

Submitted 16.Apr.2026; this is a non–peer-reviewed article;accepted 16.Apr.2026; published 30.Apr.2026.

Please cite as:
Castonguay A, Hegg-Deloye S, Paré G, Etindele Sosso FA
Correction: Perceptions and Intentions of Nursing Students Regarding Digital Health: Cross-Sectional Study
JMIR Nursing 2026;9:e98552
URL: https://nursing.jmir.org/2026/1/e98552
doi:10.2196/98552

© Alexandre Castonguay, Sandrine Hegg-Deloye, Guy Paré, Faustin Armel Etindele Sosso. Originally published in JMIR Nursing (https://nursing.jmir.org), 30.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Nursing Informaticians in Spain: Scoping Review and Expert-Validated Gap Analysis

Adrián Marco-Moyano[1], RN, BSN, MSN; Patricia Verdú Rodríguez[1], BSBE, MSBI; Manuel Lillo-Crespo[2], RN, BSN, MBAExec, MSN, MSCAnthro, PhD

[1]University of Alicante, San Vicente del Raspeig, Alicante, Spain

[2]Faculty of Health Sciences, University of Alicante, Carretera de San Vicente del Raspeig, s/n, San Vicente del Raspeig, Alicante, Spain

Corresponding Author:
Manuel Lillo-Crespo, RN, BSN, MBAExec, MSN, MSCAnthro, PhD
Faculty of Health Sciences, University of Alicante, Carretera de San Vicente del Raspeig, s/n, San Vicente del Raspeig, Alicante, Spain

Related Article:

Correction of: https://nursing.jmir.org/2026/1/e83373/

(JMIR Nursing 2026;9:e100481) doi:10.2196/100481

In "Nursing Informaticians in Spain: Scoping Review and Expert-Validated Gap Analysis" [1], the authors noted one error.

The corresponding author information has been changed from:

Adrián Marco-Moyano, RN, BSN, MSN

Nursing, Faculty of Health Sciences, University of Alicante

Carr. de San Vicente del Raspeig, s/n 965 90 34 00

San Vicent del Raspeig, Alicante 03690

Spain

Phone: 34 681281259

Email: adrian.marco@grupohla.com

It now appears as follows:

Manuel Lillo-Crespo, RN, BSN, MBAExec, MSN, MSCAnthro, PhD

Faculty of Health Sciences, University of Alicante

Carretera de San Vicente del Raspeig, s/n

San Vicente del Raspeig, Alicante 03690

Spain

Phone: 34 965903400 ext 3518

Email: manuel.lillo@ua.es

This information was added to the affiliations list as affiliation 2.

In addition, affiliation 1 was changed from:

Nursing, Faculty of Health Sciences, University of Alicante, San Vicent del Raspeig, Alicante, Spain

It now appears as:

University of Alicante, San Vicent del Raspeig, Alicante, Spain

The correction will appear in the online version of the paper on the JMIR Publications website together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Marco-Moyano A, Verdú Rodríguez P, Lillo-Crespo M. Nursing informaticians in Spain: scoping review and expert-validated Gap Analysis. JMIR Nurs 2026 Apr 27;9:e83373. [doi: 10.2196/83373] [Medline: 42044327]

Submitted 06.May.2026; this is a non–peer-reviewed article;accepted 19.May.2026; published 11.Aug.2026.

Please cite as:
Marco-Moyano A, Verdú Rodríguez P, Lillo-Crespo M
Correction: Nursing Informaticians in Spain: Scoping Review and Expert-Validated Gap Analysis
JMIR Nursing 2026;9:e100481
URL: https://nursing.jmir.org/2026/1/e100481
doi:10.2196/100481

XSL•FO
RenderX

© Patricia Verdú Rodríguez, Manuel Lillo-Crespo. Originally published in JMIR Nursing (https://nursing.jmir.org), 11.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Nurses' Experiences of Interprofessional Collaboration in Digitally Supported Hospital Discharge Planning: Qualitative Study

Mera Delima[1,2*], MN; Musheer A Aljaberi[3,4*], PhD; Regidor III Dioso[1*], PhD

[1]Faculty of Nursing, Lincoln University, Malaysia, Selangor, Malaysia

[2]Faculty of Health Sciences, Universitas Perintis Indonesia, Jl. Adinegoro, Simpang Kalumpang, Lubuk Buaya, Padang, West Sumatra, Indonesia

[3]Franciscus Academy, Franciscus Gasthuis & Vlietland, Rotterdam, The Netherlands

[4]Research Centre Innovations in Care, Rotterdam University of Applied Sciences, Rotterdam, The Netherlands

[*]all authors contributed equally

Corresponding Author:
Mera Delima, MN
Faculty of Nursing, Lincoln University, Malaysia, Selangor, Malaysia

Abstract

Background: Effective interprofessional collaboration (IPC) in patient discharge planning is essential for ensuring continuity of care, improving patient outcomes, and strengthening coordination among health care professionals. Nurses often serve as primary coordinators due to their continuous engagement in patient care. However, the implementation of IPC continues to face barriers at the individual, team, and organizational levels. Many hospitals have adopted digital tools, such as integrated patient progress notes (IPPNs), to facilitate information sharing. Nevertheless, the use of these tools to support IPC remains suboptimal and has been insufficiently explored, particularly within the Indonesian digital health context.

Objective: This study aimed to explore how IPPNs support IPC during patient discharge planning, particularly from the nursing perspective.

Methods: A qualitative phenomenological study was conducted at a hospital in Bukittinggi, West Sumatra. Data were collected through in-depth interviews and a focus group discussion involving 9 purposively selected health care professionals. Thematic analysis was used to identify key patterns related to IPC practices and communication dynamics involving the use of IPPNs.

Results: The findings revealed 3 main themes: (1) individual understanding and motivation in IPC, encompassing motivation, role expectations, personality style, and professional strengths; (2) team dynamics, including leadership, management, communication, and social support; and (3) organizational support for IPC, comprising collaborative culture, institutional goals, organizational structures, and the organizational environment. Participants perceived IPC as essential yet inefficiently utilized for coordinating patient care across disciplines, with limitations in standardization, accessibility, and clarity of digital documentation hindering effective collaboration.

Conclusions: This study demonstrated that IPC practices were shaped by individual, team, and organizational factors, with digital communication holding a potentially transformative role in facilitating collaboration. These findings contribute to existing knowledge by highlighting context-specific challenges in Indonesian digital health settings, including digital literacy, system usability, and institutional support, which influence IPC and discharge planning outcomes. Integrating digital optimization within IPC frameworks may represent a valuable strategy for advancing digital health practices.

(JMIR Nursing 2026;9:e81961) doi:10.2196/81961

KEYWORDS

digital health; digital communication; discharge planning; electronic health records; health information systems; interprofessional collaboration; integrated patient progress notes; qualitative study

Introduction

Patient discharge from hospital care represents a critical stage in the care continuum, particularly for individuals with chronic or complex conditions [1]. Effective discharge planning requires interprofessional collaboration (IPC) to ensure continuity of care and minimize adverse health outcomes [2]. Nurses play a central role as patient and family educators and function as key coordinators of interprofessional communication. However, IPC remains suboptimal in many health care systems, particularly in low- and middle-income countries (LMICs), contributing to prolonged hospital stays, medication errors, fragmented care, and reduced patient satisfaction [1,3].

Technological innovations have been identified as potential strategies for enhancing IPC. Web-based platforms, such as *Discharge Today*, have demonstrated improvements in team

XSL•FO
RenderX

coordination without increasing workload, with most users perceiving the information as accurate and useful [3]. In situations with limited face-to-face interaction, such as shift handovers, these tools support continuity of information exchange among team members [4]. Virtual simulation–based interprofessional education (IPE) has also been shown to strengthen collaborative competencies among health care professionals, particularly nursing students [5]. However, the successful implementation of such interventions is highly dependent on health care professionals' digital readiness, motivation, and technological competence.

Barriers to effective IPC persist at both institutional and individual levels. Organizational constraints, including weak leadership and inadequate managerial support, hinder technology-enabled collaboration and contribute to poor communication and reduced team performance [6,7]. At the individual level, resistance to change, heavy workload, limited digital literacy, and unclear professional roles impede collaboration, particularly during discharge transitions [2,6].

Digital documentation tools, including electronic health records, shared care plans, and standardized discharge forms, are designed to support IPC by improving clarity and consistency of clinical information [3]. However, practical challenges, including documentation fatigue, limited training, and system incompatibility, often limit their optimal use in clinical practice [4].

In Indonesia, hospitals have adopted an integrated patient progress notes (IPPNs), locally known as *Catatan Perkembangan Pasien Terpadu*, as part of the national electronic health records system to support multidisciplinary documentation [8]. Although nurses are required to routinely document patient progress using this system, studies have reported implementation inconsistencies, including incomplete entries, lack of standardization, and suboptimal use during discharge planning [9], compounded by unclear protocols, administrative workload, and limited digital literacy training [10].

Similar implementation gaps are evident in a type B teaching referral hospital in West Sumatra that has adopted the IPPNs system as part of its digital health strategy. However, despite being policy-mandated, no formal evaluation has examined the use of IPPNs in supporting IPC during discharge planning, particularly from nurses' perspectives. Existing studies have primarily focused on technical or administrative aspects of documentation, with limited examination of IPPNs as a digital collaboration tool within clinical practice.

This hospital was selected as the study site because it functions as a regional referral center and teaching institution that actively implements national digital health integration and manages multidisciplinary, high-complexity cases requiring intensive interprofessional coordination. These characteristics make it representative of mid-level health care facilities transitioning toward integrated digital health systems. Therefore, this study aimed to explore how IPPNs are used as a digital collaboration tool to support IPC during patient discharge planning, with a particular focus on nurses' perspectives and interprofessional practice.

Methods

Study Design

This study used a qualitative phenomenological design to explore health care professionals' experiences of IPC in patient discharge planning supported by digital documentation in a hospital setting. The phenomenological approach was selected because it enables an in-depth exploration of participants' subjective experiences and the meanings they ascribe to the phenomenon under investigation [11-15].

Setting and Samples

This study was conducted between January and February 2025 at a type B teaching referral hospital in West Sumatra, Indonesia. The hospital has implemented IPPNs, locally known as *Catatan Perkembangan Pasien Terpadu*, as part of its digital service standards. The hospital was selected because it involves multiple health care professions in the collaborative process of patient discharge planning.

Participants were recruited using purposive sampling and comprised 9 participants: 3 physicians, 2 nurses, 2 pharmacists, and 2 nutritionists. This interprofessional composition was intended to reflect collaborative dynamics among health care team members. Data saturation was achieved, as no new themes emerged after the eighth interview, and the ninth interview confirmed thematic repetition. In phenomenological studies, data saturation is typically achieved with 6 to 10 participants who have direct experience with the phenomenon under investigation [12,13].

Although the study was nursing-oriented, 2 nurses were included because the primary focus was IPC in discharge planning rather than nursing practice alone. The selected nurses had direct clinical and coordinative experience in the discharge process, providing sufficiently rich insights to represent nursing perspectives. In phenomenological research, depth and richness of meaning are prioritized over sample size, making this sample appropriate for addressing the research objectives [13,14].

Participants were selected based on the following inclusion criteria: (1) Physicians must hold a valid practice license (Surat Izin Praktik, SIP) and a clinical competency certificate (Surat Keterangan Kewenangan Klinik, RKK), have at least 2 years of work experience at the hospital, and consent to participate; (2) Clinical nurses (level III) must possess an SIP and a signed RKK, have at least 2 years of work experience at the hospital, and consent to participate; (3) Pharmacists must hold an SIP and an RKK, have at least 2 years of work experience at the hospital, and consent to participate; and (4) Nutritionists must hold an SIP and an RKK, have at least 2 years of work experience at the hospital, and consent to participate.

All participants were anonymized and assigned codes to ensure confidentiality and protect personal information throughout the research process.

Data Collection Procedures

Data were collected through in-depth interviews and face-to-face focus group discussions (FGDs) conducted at the hospital. The interview guide was developed based on the World Health

XSL•FO
RenderX

Organization (WHO) IPC framework and King's Interpersonal System Theory, focusing on 4 main domains: (1) perceptions of IPC in discharge planning; (2) roles and responsibilities of each professional group within the team; (3) barriers and facilitating factors influencing IPC implementation in the hospital; and (4) utilization of the IPPNs as a digital communication tool for coordination among professionals.

The following are the example interview questions:

1. How do you perceive IPC in the discharge planning process?
2. What challenges have you encountered when using IPPNs to coordinate with other professionals?
3. In your view, what factors could strengthen interprofessional teamwork in this hospital?

Each interview lasted approximately 45 to 60 minutes, was audio-recorded with participants' consent, and was transcribed verbatim.

Following the individual interviews and initial thematic analysis, an FGD was conducted to validate and refine the themes that emerged from the interviews. The FGD served as a member-checking process to confirm the credibility of the findings and to provide cross-professional reflection on shared experiences. Importantly, the FGD was not intended to generate new themes but to enhance trustworthiness through data triangulation and ensure consistency of meaning across professional perspectives.

All participants received a clear explanation of the study objectives and procedures, and informed consent was obtained before data collection.

Data Analysis

Data collection and analysis were conducted in 2 stages. The first stage involved 9 in-depth interviews using a semistructured interview guide consisting of open-ended questions based on the WHO framework for IPC and the core IPC competencies as outlined by the Institute of Medicine in the expert panel report. The information obtained from these face-to-face interviews was recorded with consent and transcribed verbatim. As the interviews and FGD were conducted in Bahasa Indonesia, data analysis was performed in the original language. Quotations included in the manuscript were translated into English carefully to retain conceptual accuracy. Interview data were analyzed thematically following the approaches proposed by Braun and Clarke [12,13], including open coding, theme development, and interpretation. The coding and theme categorization process was supported by NVivo 15 software (Lumivero) to ensure analytic consistency and traceability.

The reporting of this study adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist to ensure transparency and methodological rigor. The coding process followed the 6 phases of thematic analysis proposed by Braun and Clarke, including data familiarization, generation of initial codes, searching for themes, reviewing, defining, and naming themes. A preliminary codebook was developed inductively from the data, then refined through iterative discussions among the research team to ensure conceptual clarity. Coding discrepancies were resolved through peer debriefing and consensus meetings to achieve analytic consistency.

The second stage of data collection was an FGD conducted only once, including 4 participants selected from the initial participants based on the richness and relevance of the information provided during the previous analysis. FGD data were analyzed using the same thematic method, and the results from both stages were integrated using triangulation.

Potential bias was addressed in the FGD by ensuring equal opportunities for all participants to speak and by avoiding dominant voices that overshadowed others. The discussion was guided neutrally, with questions designed to elicit further reflection rather than primary responses.

Trustworthiness and Rigor

To maintain the trustworthiness and rigor of qualitative data, the strategies outlined by Nowell et al [14,15] were applied as follows: credibility was ensured through data saturation, verbatim transcription, and triangulation of in-depth interviews and FGDs; transferability was supported by a detailed description of the study context, including participants' profiles and data collection procedures; dependability was ensured through detailed documentation of the data collection and analysis processes; and confirmability was established through triangulation and discussions with study team members to ensure consistent and objective interpretations [14,15].

Ethical Considerations

This study underwent ethical review and was approved by DR Drs M Hatta Brain Hospital, Bukittinggi, Indonesia (002555/KEP.RSOMH Bukittinggi/2024). Before data collection, participants were provided with a study information sheet explaining the purpose, benefits, methods, and participant rights. Participation was voluntary, and participants could withdraw at any time without consequences. Written consent was obtained from each participant after they read and understood the information provided. Confidentiality and anonymity were maintained by excluding individual identities from the report. Data were used solely for this study and stored securely in accordance with privacy policies. The research was conducted in accordance with the Helsinki Declaration. Participants did not receive any financial or material compensation for their participation in this study.

Results

Descriptive Results

Before presenting the main findings, an overview of participant characteristics is provided in Table 1, which summarizes the characteristics of 9 participants, all within the productive age range (25 - 64 y) as defined by the WHO. Most participants were female (n=7, 78%) and held a master's degree (n=7, 78%). The majority were married (n=7, 78%) and had more than 10 years of work experience (n=6, 66.7%).

XSL•FO
RenderX

Table . Participants' characteristics.

Characteristics	Frequency, n (%)
Age (y)	
25 - 64: adult (productive age range) according to WHO[a]	9 (100)
Sex	
Male	2 (22)
Female	7 (78)
Education	
Bachelor's degree	2 (22)
Master's degree	7 (78)
Marital status	
Married	7 (78)
Not married	2 (22)
Profession	
Doctor	3 (33.3)
Nurse	2 (22.2)
Pharmacist	2 (22.2)
Nutritionist	2 (22.2)
Work period (y)	
<5	0 (0)
5 - 10	3 (33.3)
>10	6 (66.7)

[a]WHO: World Health Organization.

Participants represented 4 professional backgrounds: doctors, nurses, pharmacists, and nutritionists, reflecting multidisciplinary involvement in patient discharge planning.

Health Care Professionals' Experiences in Implementing IPC in Hospital Settings

Overview of Emergent Themes

Three main themes and 12 subthemes emerged from the analysis, describing health care professionals' experiences in implementing IPC during patient discharge planning. The main themes and subthemes are summarized in Table 2.

Table . Themes of health professionals' experiences in IPC[a] practice at a type B teaching referral hospital in West Sumatra, Indonesia.

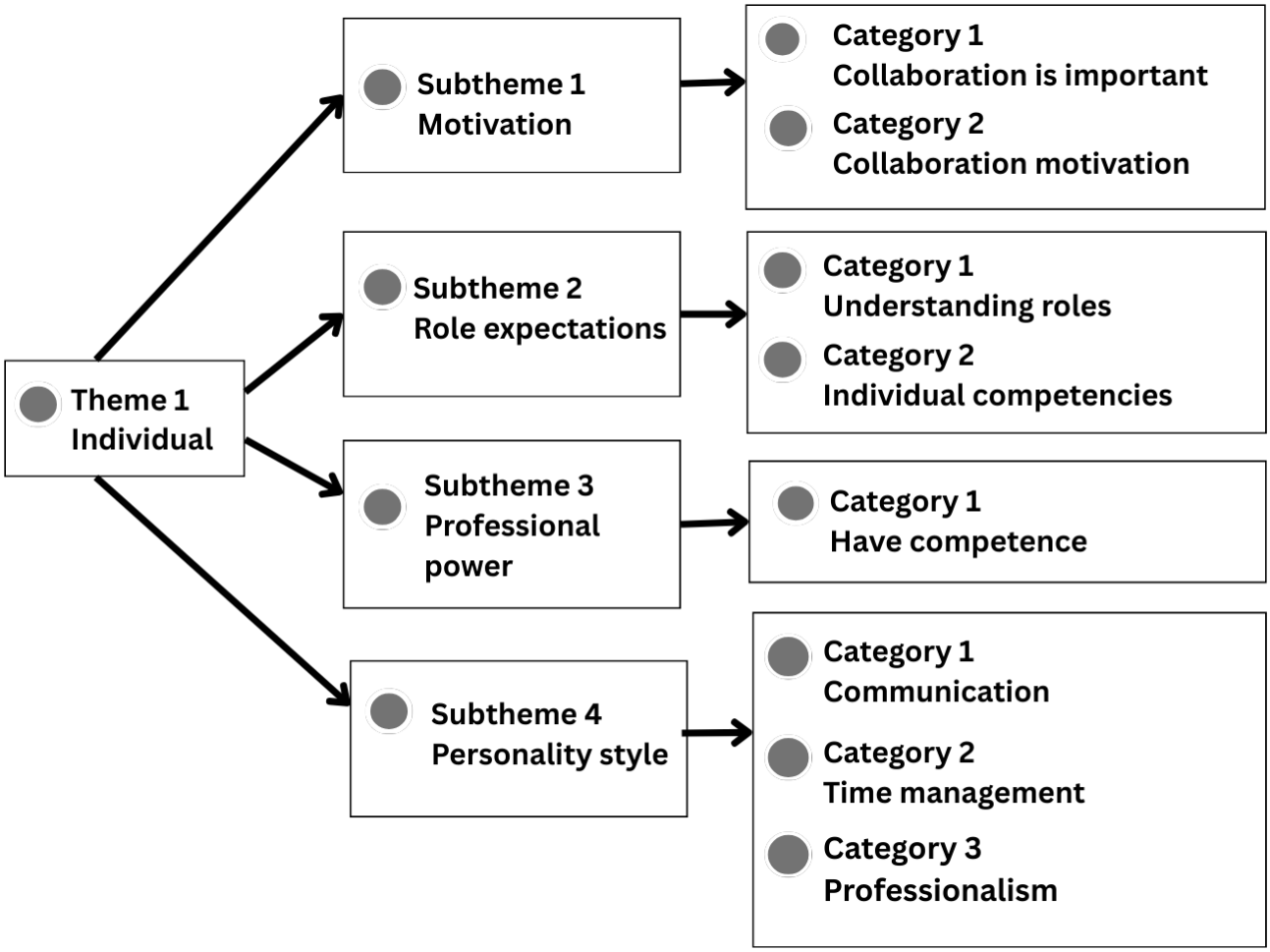
Code	Themes	Subthemes
1	Individual understanding and motivation in IPC	• Motivation • Role expectations • Professional power • Personality style
2	Team interaction dynamics in the discharge planning process	• Group leadership • Coping • Communication • Social support
3	Organizational support for IPC	• Organizational culture • Organizational goals • Organizational domain • The organizational environment

[a]IPC: interprofessional collaboration.

XSL•FO
RenderX

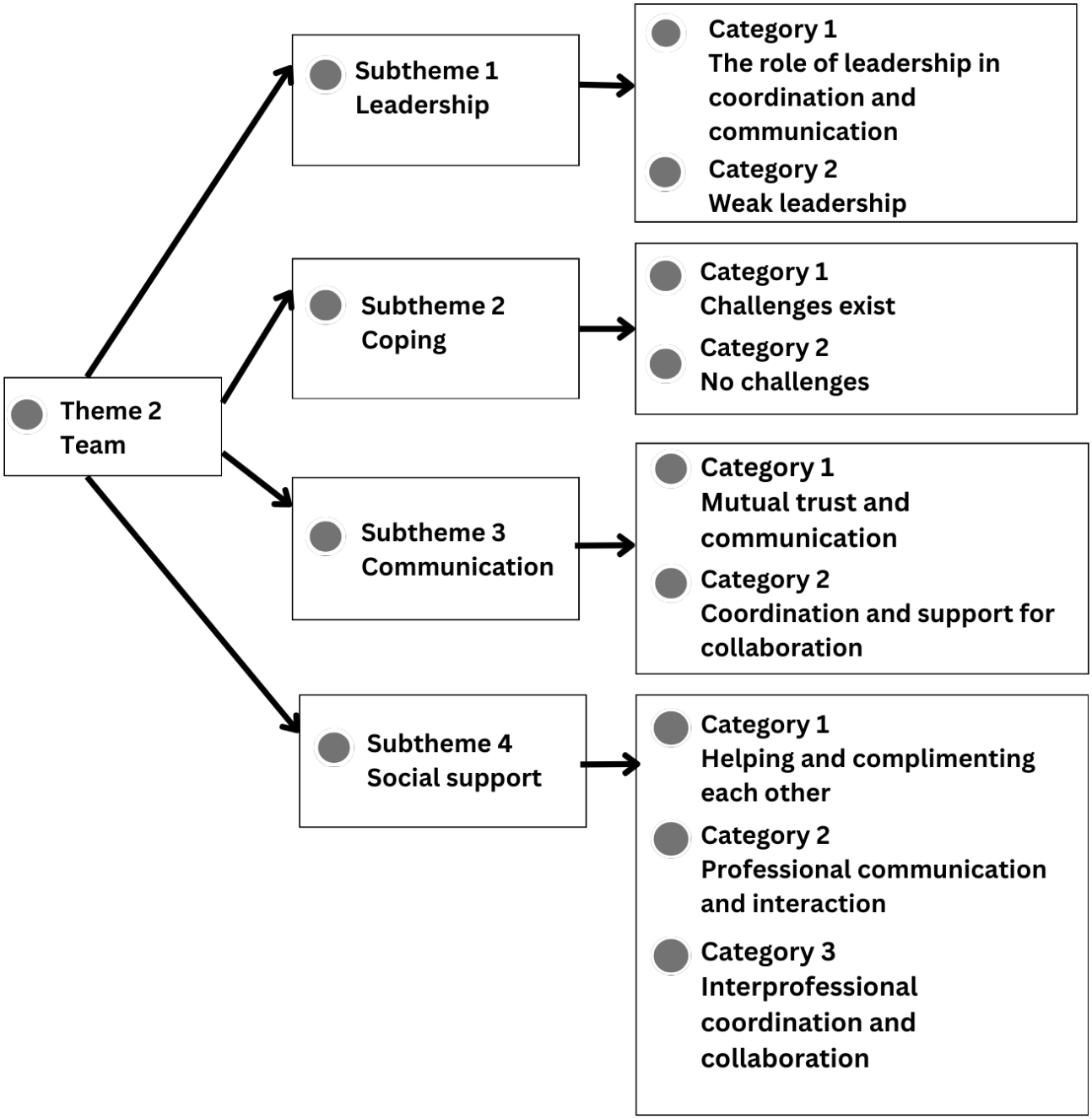
Further details of each theme, including supporting categories and illustrative participant quotations, are presented in Figures 1-3, which were generated using NVivo 15 to visualize the thematic structure, and described in the following sections.

Figure 1. Thematic analysis of theme 1: individual.



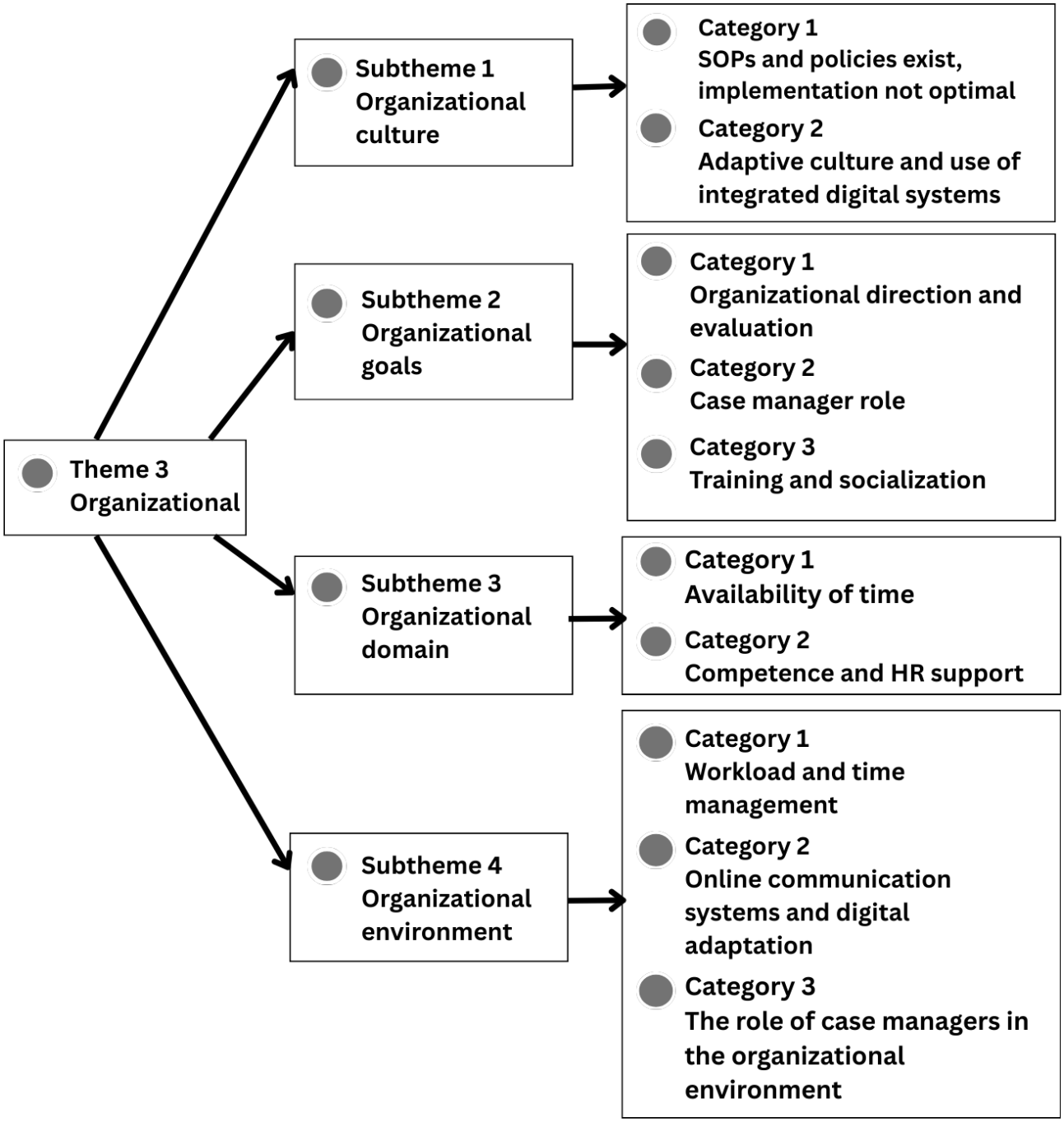
XSL•FO
RenderX

Figure 2. Thematic analysis of theme 2: team



XSL•FO
RenderX

Figure 3. Thematic analysis of theme 3: organizational support. HR: human resource; SOP: standard operating procedure.



Theme 1: Individual Understanding and Motivation in IPC

Overview of Theme 1

Qualitative findings suggest that individual factors play a crucial role in the implementation of IPC, particularly in the context of patient discharge planning. This theme comprises 4 subthemes: motivation, role expectations, professional authority, and personality styles, as presented in Figure 1.

Figure 1 illustrates theme 1, "individual factors", which includes 4 subthemes with associated categories: (1) motivation, including the importance of collaboration and collaborative motivation; (2) role expectations, including understanding professional roles and individual competencies; (3) professional authority, including possession of professional competence; and (4) personality style, including communication, time management, and professionalism.

The following sections elaborate on each subtheme and present supporting quotations from the 9 participants.

Subtheme 1: Motivation

Internal motivation emerged as a key driver encouraging nurses to participate actively in collaborative discharge planning. Several participants recognized teamwork as essential to ensuring patient safety and continuity of care.

> *Collaboration is essential, especially in planning patient discharge to avoid mistakes.* [Participant 2]

XSL•FO
RenderX

Nurses also expressed motivation stemming from a desire to provide comprehensive patient education before discharge.

> *I'm driven to collaborate so patients receive the right information before going home.* [Participant 4]

This intrinsic drive was reinforced by shared professional goals and a sense of collective responsibility.

> *I feel more enthusiastic when working with a cross-professional team because we share the same goal—the patient's recovery.* [Participant 5.]

Some participants noted that motivation diminished when their contributions were not equally recognized.

Subtheme 2: Role Expectations

Role clarity influenced participation in discharge planning. Participants emphasized that understanding their professional responsibilities guided their involvement in team discussions.

> *I understand that as a nurse, my responsibility is to educate the patient before discharge.* [Participant 2]

Several participants identified interprofessional communication as part of their professional role.

> *I report the patient's condition to the team before they are discharged.* [Participant 3]

Some nurses reported frustration when their roles were perceived as limited to routine tasks.

> *Sometimes we're seen only as medication providers, but we also educate patients about side effects.* [Participant 9]

This perception gap between professions can act as both a facilitator and a barrier to collaboration, suggesting the need for clearer role delineation and mutual respect in team settings.

Subtheme 3: Professional Power

Professional hierarchy influenced collaborative dynamics. Some nurses expressed confidence in their clinical competence and independence.

> *As a nurse, I feel capable of carrying out discharge planning.* [Participant 3]

Others reported discomfort when expressing opinions in the presence of physicians.

> *Sometimes it's difficult to share opinions because we're seen as subordinates.* [Participant 5]

Training and institutional support were described as contributing to professional confidence.

> *After attending training, I felt more confident managing discharge planning.* [Participant 6]

Conversely, some participants reported that dominant attitudes from certain professionals affected collaboration.

> *Sometimes doctors act like they know everything, which makes it hard for others to express their views.* [Participant 7]

These findings illustrate the dual nature of professional power; it can empower individuals through competence and training or hinder collaboration through entrenched hierarchies.

Subtheme 4: Personality Styles

Interpersonal communication and personal traits influenced teamwork. Participants associated effective collaboration with openness, confidence, and adaptability.

> *I'm used to coordinating and listening to team members' input.*
>
> *[Participant 7]*
>
> *I prefer being open if there's a different view, I take it as constructive feedback.* [Participant 8]

Adaptability was identified as helpful when working with professionals from different disciplines.

> *I'm adaptable and used to working across different professions.* [Participant 9]

Nevertheless, communication gaps were occasionally reported, especially when individuals lacked confidence or were hesitant to speak up. This subtheme underscores that successful IPC depends not only on technical skills but also on interpersonal awareness and flexibility.

Summary of Theme 1

Individual factors influenced participation in IPC during discharge planning. Participants described motivation, role understanding, professional hierarchy, and personal communication styles as shaping their collaborative experiences.

Theme 2: Team Interaction Dynamics in the Discharge Planning Process

Overview of Theme 2

Team interaction dynamics played an important role in the discharge planning process and influenced the implementation of IPC. Participation and coordination among different professional groups were described as central to daily collaborative practice. Participants from various professions, including doctors, nurses, pharmacists, and nutritionists, described diverse interaction patterns that shaped how collaboration was formed, maintained, or hindered during discharge planning. These interaction dynamics encompassed leadership, communication, coordination, and social relationships within the health care team. The identified patterns illustrate how team members worked together to support patient discharge planning in routine practice. The structure of the themes, subthemes, and categories related to team interaction dynamics is presented in Figure 2.

Figure 2 presents Theme 2, "team interaction dynamics", which includes 4 subthemes with associated categories: (1) group leadership, including leadership roles in coordination and communication and weak leadership; (2) coping, including collectively managed challenges and absence of challenges; (3) communication, including mutual trust in team communication and coordination and support for collaboration; and (4) social support, including helping and complementing each other, professional communication and interaction, and interprofessional coordination and collaboration.

XSL•FO
RenderX

Subtheme 1: Leadership

Group leadership influenced the effectiveness of collaboration in discharge planning. Participants described leaders as facilitators who initiated meetings, allocated responsibilities, and guided discussions.

> *Our leader always starts the discharge planning meeting and divides the tasks.* [P1]
>
> *The team leader always guides the discharge planning discussion firmly but openly.* [P2]

Participants described leadership styles characterized by fairness and openness: "Our leader encourages all members to express their opinions" (P3), and "Our leader is wise and willing to listen to all input" (P5). Some participants highlighted emotional control and role modeling by leaders: "Our team leader can maintain harmony, not authoritarianism" (P6), and "Our leader sets an example and supports joint decisions" (P7). Collectively, these narratives underscore leadership that balances authority with inclusivity.

Subtheme 2: Coping/Handling

Participants described both the presence and absence of challenges in IPC. Reported barriers included inconsistent evaluation, limited family engagement, and uneven participation across professions: "The challenges exist, especially because there has been no proper evaluation" (P1); "Coordination is a challenge because not all professions are present during joint visits" (P5). Differences in patient contact among professionals were also noted (P3). Some participants reported minimal challenges due to clear standard operating procedures (SOPs) and defined roles: "There are no challenges because all professions already have clear SOPs" (P6).

Subtheme 3: Communication

Participants described communication as occurring primarily through digital platforms, particularly WhatsApp groups: "Communication is carried out via WA groups" (P2, P7, P8, P9). Verbal discussions and documentation were also used to share patient updates: "We routinely inform each other about patient progress" (P1); "I routinely report patient progress to the team via medical records" (P3). Some participants reported communication challenges among new staff unfamiliar with interprofessional communication practices: "There are minor challenges if there are new employees who do not yet understand the ethics or interprofessional communication" (P6). This indicates the need for structured orientation and communication training within multidisciplinary teams.

Subtheme 4: Social Support

Participants described social support as contributing to team cohesion. They reported mutual trust, respect, and shared responsibility. As participant 1 stated, "I feel supported by other team members, so I am more confident," while participant 3 added, "We respect each other; no one feels more important." Participants also described encouragement and practical support among team members: "My team members always give encouragement and moral support" (P4). Moreover, collective accountability was evident when colleagues substituted for absent members without complaint: "If someone is absent, the team still covers for each other without protesting" (P6); "If someone is absent, the others are ready to change shifts without any problems" (P8).

Summary of Theme 2

Team interaction dynamics influenced IPC during discharge planning. Participants described leadership, communication, coping with challenges, and social support as shaping their collaborative experiences.

Theme 3: Organizational Support for IPC

Overview of Theme 3

Organizational support influenced the implementation of IPC in patient discharge planning. Participants described the role of policies, management, resources, information systems, and work culture in shaping collaborative practice. The structure of the themes, subthemes, and categories for organizational support is presented in Figure 3.

Figure 3 presents Theme 3, "organizational support", consisting of 4 subthemes with associated categories: (1) organizational culture, including existing SOPs and policies with suboptimal implementation and adaptive culture through integrated digital systems; (2) organizational goals, including organizational direction and evaluation, case manager roles, and training and socialization; (3) organizational structure, including availability of time and competence supported by human resources; and (4) organizational environment, including workload coverage, time management, online communication systems, and the role of case managers in supporting IPC.

Subtheme 1: Organizational Culture

Participants reported the presence of SOPs and policies supporting discharge planning. However, limited dissemination and inconsistent implementation were noted.

> *There are SOPs, but access is uneven across units, which can hinder implementation.* [Participant 4]
>
> *[A]lthough collaboration is mandated, the execution remains suboptimal.* [Participant 9]

Some participants described the integration of electronic medical records as supporting collaboration. For example, a participant shared, "The new H-1 discharge policy is linked to electronic medical records, making it more transparent" (Participants 3, 5).

Others noted that implementation at the unit level remained inconsistent: "Although management has established procedures, their execution at the ward level still needs improvement" (Participants 2, 5). A formal structure supporting IPC exists, while cultivating a shared culture that translates into consistent, collaborative behavior in daily clinical work remains a challenge.

Subtheme 2: Organizational Goals

Participants described management support for collaboration through direction, case management, training, and evaluation.

> *There are directions from management, but no overarching policy; however, patient satisfaction remains a central focus.* [Participant 2]

XSL•FO
RenderX

> [The presence of a case manager facilitates cross-professional decisions. Participant 6]
>
> *Periodic training and evaluation are carried out to adjust discharge planning practices.* [Participant 8]

Some participants highlighted the role of digital systems in monitoring collaboration:

> *The iKame system allows us to track multidisciplinary participation.* [Participant 7]

Subtheme 3: Organizational Structure

Participants reported that the availability of discharge planning teams and digital tools supported collaboration. Several participants described sufficient time allocation and improved efficiency through electronic systems. A participant shared, "Time is sufficient as long as roles are clear" (Participant 4), while another added, "Using the electronic system saves time compared to manual charting" (Participant 9). However, human resource limitations were also reported: "Time remains a challenge, but we hope things improve with more staff" (Participants 3, 5).

Facilities and digital access were described as adequate: "computers are available" and "nutrition leaflets and counseling are provided" (Participant 2). Participants also described the hospital's electronic system as structured and role-based: "the IT system supports collaboration" and "access is structured based on role and profession" (Participants 6, 7). In summary, structural elements such as time, facilities, and information systems play a critical role in enabling IPC, although workload distribution and continuity of IPE still require improvement.

Subtheme 4: Organizational Environment

Participants generally perceived the work environment as supportive of collaboration: "The work environment supports collaboration and discussion" (Participant 2). Scheduling and coordination challenges were commonly reported. For example, "lack of adherence to schedules affects plans like physiotherapy" and "conflicting rounds and outpatient duties disrupt communication" (Participant 6).

Some participants suggested integrated scheduling and dedicated interprofessional meetings. "More integrated scheduling and dedicated interprofessional meeting time" (Participant 8). While digital systems were viewed as helpful, some participants preferred direct communication: "IT helps, but direct coordination is still more effective" (Participant 7).

Participants also mentioned the need for regular evaluations: "There should be routine meetings and comprehensive evaluations beyond administrative aspects" (Participant 2). Importantly, individual commitment was stated: "Even with full organizational support, without individual commitment, collaboration won't work" (Participant 9).

Few participants consider IPC to be running optimally. However, the results suggest that organizational improvements, particularly in scheduling, evaluation, and communication, are still needed to sustain collaborative efforts.

Summary of Theme 3

Participants described organizational policies, digital systems, managerial support, training, case managers, and work environment as influencing IPC during discharge planning. They also reported challenges related to policy implementation, scheduling, workload, and communication.

Discussion

Principal Findings

The findings of this study indicate that the success of IPC in patient discharge planning is influenced by 3 main factors: individual factors, team dynamics, and organizational support. Individual factors include motivation, role understanding, professional competence, and personality styles; team dynamics involve leadership, communication, problem-solving abilities, and social support; and organizational support encompasses work culture, service goals, resource availability, and a supportive work environment. These findings underscore that IPC effectiveness depends on the interaction between personal capabilities, effective team processes, and organizational structures.

Before discussing the findings in detail, participant characteristics should be considered, as they influence experiences and perceptions of IPC. All participants (n=9) were within the productive age range of 25 to 64 years, reflecting readiness for IPC and adaptation to digital systems [16]. Most participants were female (78.0%), consistent with the gender distribution in nursing and public health professions, which may influence collaboration dynamics through communicative and empathetic approaches [17]. Most participants held a master's degree (78%), indicating strong academic capacity for understanding IPC concepts and applying digital communication in clinical practice [18,19]. The majority had more than 10 years of work experience (66.7%), contributing to professional maturity and confidence in interprofessional communication [17].

Although only 2 nurses participated, they were selected due to their strategic role as primary coordinators of discharge planning, making them representative of nursing competencies in IPC [16].

Exploring the Professional Experiences of Nursing Professionals

Theme 1: Individual Understanding and Motivation in IPC

Overview of Individual Factors Influencing IPC

The findings demonstrate the critical role of individual factors in IPC implementation during discharge planning. Motivation, role expectations, professional power, and personality traits shaped health care professionals' engagement and readiness. These findings align with international IPC literature, highlighting individual awareness and intrinsic motivation as foundational to effective collaboration.

XSL•FO
RenderX

Motivation

Motivation emerged as a strong internal driver rooted in responsibility for patient safety and care continuity. This aligns with Smith et al [5], who showed that interprofessional simulation strengthened motivation and collaborative readiness, and with Reinders et al [20], who emphasized intrinsic motivation in sustaining discharge coordination. In Indonesian settings, motivation is influenced by workload, resource constraints, and institutional incentives, indicating that it is both intrinsic and contextually shaped [20].

This study further shows that digital platforms such as IPPNs and WhatsApp function as external motivators by facilitating timely information exchange and shared accountability, consistent with global evidence on digital engagement in teamwork [16]. However, infrastructure limitations, technical barriers, and workload concerns remain as challenges [21]. Thus, technology-supported motivation requires organizational training, leadership, and incentives, particularly in LMIC contexts where motivation is more relational than structurally reinforced.

Role Expectations

Clear role understanding enhanced IPC participation, especially in patient education and communication. Conversely, undervalued or ambiguous roles—particularly among nurses—limited engagement. These findings support Tong et al [22], who found that role clarity improves collaboration and mutual respect. In Indonesia, hierarchical traditions and inconsistent policy enforcement continue to constrain role clarity [9,10]. Compared with high-income settings where roles are codified through standardized policies and digital protocols, local adaptation of IPC policies remains essential [16].

Professional Power

Professional hierarchy remained a persistent barrier. Despite clinical competence, some nurses hesitated to voice opinions due to perceived lower status. This mirrors Tan et al [23] and Nie et al [4], who reported that digital tools alone do not flatten hierarchies without inclusive leadership and policy support. In LMIC contexts, unequal authority and weak institutional enforcement limit the transformative impact of digitalization on power relations [4].

Personality Style

Personality traits such as openness, adaptability, and empathy facilitated IPC, consistent with prior studies linking emotional intelligence to effective handoffs and reduced information loss [10,24,25]. In the Indonesian context, strong family involvement added complexity to discharge planning, requiring sensitivity to sociocultural dynamics [26]. Comparable findings in high-income settings also highlight emotional intelligence as a key IPC enabler [27]. These results indicate that IPC effectiveness is shaped not only by systems but also by personal and interpersonal competencies [28].

Critical Reflection

These findings reinforce that IPC is shaped by individual attitudes and professional culture in addition to formal systems. Digital platforms can enhance coordination, but their effectiveness depends on leadership, organizational readiness, and equitable governance. Persistent hierarchies and uneven policy implementation in LMICs underscore the need for capacity building that integrates reflective leadership, interprofessional mentoring, and psychological safety.

Theme 2: Team Dynamics in IPC for Discharge Planning

Overview of Team Interaction Dynamics

Team dynamics revealed the importance of leadership, communication, coping strategies, and social support. While digital platforms facilitated coordination, collaboration quality ultimately depended on leadership consistency, institutional accountability, and cultural readiness.

Group Leadership and Participatory Team Culture

Inclusive leadership fostered trust, shared accountability, and psychological safety, consistent with Bornman and Louw [29] and Keniston et al [30]. Conversely, inconsistent leadership resulted in fragmented communication and unclear roles, reflecting uneven policy implementation and limited leadership development common in LMICs [5].

Coping and Operational Challenges in Coordination

Teams faced logistical challenges such as inconsistent participation and fragmented workflows, echoing findings by Nie et al [4]. Clear SOPs and predefined roles mitigated some challenges, aligning with Buljac-Samardzic et al [31]. Persistent hierarchies and uneven enforcement reflected deeper governance barriers, while family involvement occasionally delayed decision-making when authority was unclear [31].

Communication Infrastructure and Team Transparency

Hybrid communication using WhatsApp, face-to-face meetings, and documentation supported coordination, consistent with Keniston et al [30] and Teuwen et al [32]. However, message overload and unequal participation persisted without shared digital literacy and leadership oversight [21,33-35]. Compared with high-income settings, IPC in Indonesia remains dependent on interpersonal initiative rather than integrated systems [27,34]

Social Support and Emotional Climate

Mutual respect, encouragement, and role flexibility strengthened team resilience, consistent with Cadel et al [34]. These findings align with evidence from high-income contexts showing emotional cohesion as a universal driver of IPC effectiveness [35-42].

Critical Reflection

Digitalization facilitates coordination but does not eliminate hierarchies or policy inconsistencies. Compared with high-income settings, IPC success in Indonesia relies more heavily on interpersonal leadership and local initiative, highlighting the need for aligned policy, leadership development, and collaborative governance.

XSL•FO
RenderX

Theme 3: Organizational Support for IPC

Organizational-Level Support for IPC

Organizational support emerged as a prerequisite for sustainable IPC, encompassing culture, goals, structure, and work environment.

Organizational Culture

Although SOPs and policies exist, uneven dissemination and weak supervision limited consistent implementation. Hierarchical norms continued to shape communication despite digital tools. These findings align with Redzewsky et al [40] and Ishii et al [41], emphasizing that policy without localized integration fails to sustain IPC.

Organizational Goals

Leadership support through case managers, training, and evaluation facilitated IPC, although the absence of a comprehensive IPC policy remained a barrier. These findings align with Labrague et al [42], who linked supportive environments to improved safety outcomes.

Organizational Domain

Structural supports such as discharge teams, digital systems, and facilities enhanced IPC efficiency [43-46]. However, workload imbalance and limited IPE persisted [45]. Rawlinson et al [46] emphasized that the success of IPC was strongly influenced by team competencies, including effective communication, clearly defined roles, and consistent organizational support.

Adopting digital systems, such as electronic medical records and the iKame app, streamlines coordination and reduces time burden across professions. However, human resource shortages and compressed work schedules continue to present major barriers. Buljac-Samardzic et al [31] suggested that team structure optimization, including task clarity, adherence to SOPs, and structured time allocation, could alleviate operational barriers to IPC. Digital platforms such as IPPNs and WhatsApp improved responsiveness but remained informally governed, limiting institutional monitoring.

The Organizational Environment

The organizational environment plays a critical role in enabling effective IPC. Factors such as the availability of resources, dedicated time allocation, and structured workflows are essential for fostering a productive collaborative climate. Although digital technologies have been introduced to support interprofessional communication, several structural barriers persist, including noncompliance with schedules, overlapping clinical responsibilities, workload pressure, and a lack of integrated cross-professional scheduling systems. Ginting et al [47] emphasized that while digital tools could facilitate IPC, their effectiveness was limited in the absence of a supportive organizational culture and inclusive leadership.

Beyond systemic factors, individual commitment remains essential to ensure sustainable collaboration. A previous study found that interprofessional trust emerges from the quality and regularity of relationships, individual attitudes, and a supportive organizational culture, and that trust, in turn, enhances communication, coordination, teamwork, and ultimately collaborative outcomes [48]. Similarly, Reeves et al [49] showed that workplaces promoting team reflection and shared evaluation cultivated trust and accountability among professionals.

When compared to high-income settings, where IPC practices are often embedded in digitalized and standardized systems, hospitals in Indonesia face persistent challenges related to uneven digital infrastructure, informal communication channels, and resource limitations. Similar barriers have been observed in other low- and middle-income contexts such as Ethiopia and the Philippines, where digital and managerial gaps hinder consistent IPC outcomes. These contrasts highlight that while policy frameworks may be globally aligned, contextual and resource-based differences shape IPC implementation in diverse health systems [50-54].

These inconsistencies often stem from institutional hierarchies, variable managerial commitment, and uneven digital literacy among professionals, which prevent the full translation of collaborative policies into daily clinical practice [55].

Holistic Interdependence Between IPC and Organizational Support

Cadel et al [34] and Smith et al [56] reported that peer substitution and team solidarity were critical to sustaining IPC, particularly during workforce fluctuations. Interprofessional simulation-based learning was also shown to strengthen team relationships and preparedness for role substitution when necessary. These findings emphasize the importance of organizational support mechanisms, including structured training and adaptive team models that can withstand staffing dynamics.

Link to Regional Evidence

A systematic review found that satisfaction with organizational support and positive attitudes toward IPC were associated with collaboration effectiveness, although not often statistically significant [50]. This highlights the need for systemic approaches to improve working conditions, promote regular training, and implement inclusive interprofessional models in clinical practice.

Critical Reflection

Organizational support for IPC depends on alignment between structure, culture, and leadership. Policies and digital tools alone are insufficient without consistent supervision, organizational learning, and reflective practice. Strong leadership commitment and structured evaluation forums are essential to embed IPC as a sustainable norm.

Strengths and Limitations

This study's strengths include its qualitative phenomenological approach, inclusion of multiple professions, and focus on digital documentation (IPPNs) within a middle-income country context. Methodological rigor was ensured through triangulation and trustworthiness criteria, contributing to theory building in IPC within resource-constrained systems.

Limitations include the single-site setting and small sample size. The study did not fully examine system usability or interoperability, and cultural hierarchies may limit

XSL•FO
RenderX

transferability. Future research should adopt multisite and mixed-method designs to explore system-level and leadership influences on IPC outcomes.

Conclusion

In conclusion, this study revealed the complex dynamics that underlie IPC in hospital discharge planning. A total of 3 interrelated themes emerged: individual understanding and motivation, team interaction dynamics, and organizational support for IPC. The findings demonstrated that intrinsic motivation and clear role expectations were fundamental in fostering individual readiness for collaboration. IPE played a significant role in strengthening professional identity and enhancing collaborative competence among nursing professionals.

Professional hierarchies and leadership styles influenced collaboration outcomes. Power imbalances hindered open communication and participation, whereas inclusive leadership fostered psychological safety and active engagement. Organizational support encompassing culture, structure, and systems proved essential, although challenges persisted in coordination, communication, and policy implementation.

Effective IPC in discharge planning requires a holistic approach that integrates individual preparedness, cohesive team dynamics, and strong organizational commitment. Actionable strategies, such as structured interprofessional orientation programs, regular digital literacy training, and consistent policy dissemination at the unit level, could strengthen collaborative practice and ensure its sustainability. These efforts can promote more integrated, patient-centered care and enhance continuity of care post discharge.

The findings provide implications for policymakers, hospital administrators, and nursing education institutions to incorporate IPC principles into professional development programs, digital health initiatives, and institutional governance frameworks, supporting collaboration as a sustained norm within health care systems.

Acknowledgments

The authors thank all nurses who participated in this study, as well as the Management and Ethics Committee of DR Drs M Hatta Brain Hospital, Bukittinggi, for their permission and support. The authors also extend their gratitude to academic supervisors and colleagues for valuable feedback and encouragement, and to their families for moral support throughout the research process.

Funding

The authors did not receive any funding or financial support for this work.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

References

1. De Regge M, De Pourcq K, Meijboom B, Trybou J, Mortier E, Eeckloo K. The role of hospitals in bridging the care continuum: a systematic review of coordination of care and follow-up for adults with chronic conditions. BMC Health Serv Res 2017 Aug 9;17(1):550. [doi: 10.1186/s12913-017-2500-0] [Medline: 28793893]
2. Steele Gray C, Tang T, Armas A, et al. Building a digital bridge to support patient-centered care transitions from hospital to home for older adults with complex care needs: protocol for a co-design, implementation, and evaluation study. JMIR Res Protoc 2020 Nov 25;9(11):e20220. [doi: 10.2196/20220] [Medline: 33237037]
3. Fuller TE, Pong DD, Piniella N, et al. Interactive digital health tools to engage patients and caregivers in discharge preparation: implementation study. J Med Internet Res 2020 Apr 28;22(4):e15573. [doi: 10.2196/15573] [Medline: 32343248]
4. Nie JX, Heidebrecht C, Zettler A, et al. The perceived ease of use and perceived usefulness of a web-based interprofessional communication and collaboration platform in the hospital setting: interview study with health care providers. JMIR Hum Factors 2023 Jan 23;10:e39051. [doi: 10.2196/39051] [Medline: 36689261]
5. Smith LM, Jacob J, Prush N, et al. Virtual interprofessional education: team collaboration in discharge planning simulation. Prof Case Manag 2024;29(5):206-217. [doi: 10.1097/NCM.0000000000000717] [Medline: 38421733]
6. Nabelsi V, Plouffe V, Leclerc MC. Barriers to and facilitators of implementing overnight nursing teleconsultation in small, rural long-term care facilities: qualitative interview study. JMIR Aging 2025 May 7;8:e71950. [doi: 10.2196/71950] [Medline: 40334266]
7. van Sluisveld N, Oerlemans A, Westert G, van der Hoeven JG, Wollersheim H, Zegers M. Barriers and facilitators to improve safety and efficiency of the ICU discharge process: a mixed methods study. BMC Health Serv Res 2017 Apr 4;17(1):251. [doi: 10.1186/s12913-017-2139-x] [Medline: 28376872]
8. Kusumaningrum PR, Dharmana E, Sulisno M. The implementation of integrated patient progress notes in interprofessional collaborative practice. J Ners dan Kebidanan Indones 2019;6(1):32. [doi: 10.21927/jnki.2018.6(1).32-41]

XSL•FO
RenderX

9. Sukawan A, Meilany L, Rahma AN. Literature review: peran CPPT dalam meningkatkan komunikasi efektif pada pelaksanaan kolaborasi interprofesional di rumah sakit [Article in Indonesian]. Indones Health Inf Manag J 2021;9(1):30-37. [doi: 10.47007/inohim.v9i1.239]
10. Sule AA, Caputo D, Gohal J, Dascenzo D. Desirable features of an interdisciplinary handoff. JMIR Nurs 2020;3(1):e18914. [doi: 10.2196/18914] [Medline: 34345786]
11. Purwanza SW. Metodologi Penelitian Kuantitatif, Kualitatif Dan Kombinasi [Book in Indonesian]: Media Sains Indoneisa; 2022. URL: https://openlibrary.telkomuniversity.ac.id/pustaka/179643/metodologi-penelitian-kuantitatif-kualitatif-dan-kombinasi.html [accessed 2026-01-19]
12. Braun V, Clarke V. Reporting guidelines for qualitative research: a values-based approach. Qual Res Psychol 2025;22(2):399-438. [doi: 10.1080/14780887.2024.2382244]
13. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2021 Jul 3;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
14. Nowell L, Johnston S, Dolan S, Jacobsen M, Lorenzetti DL, Oddone Paolucci E. Exploring educators' perceptions and experiences of online teaching to foster caring profession students' development of virtual caring skills: sequential explanatory mixed methods study. JMIR Nurs 2025 Jan 15;8:e64548. [doi: 10.2196/64548] [Medline: 39608377]
15. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1):1-13. [doi: 10.1177/1609406917733847]
16. Nordmann K, Sauter S, Möbius-Lerch P, Redlich MC, Schaller M, Fischer F. Conceptualizing interprofessional digital communication and collaboration in health care: protocol for a scoping review. JMIR Res Protoc 2023 Jun 26;12:e45179. [doi: 10.2196/45179] [Medline: 37358886]
17. Daskalopoulou A, Pirani D, Ostberg J, editors. Sexuality in Marketing and Consumption: Queer Theory, Feminist Research, and Intersectionality: Routledge; 2025, Vol. 17. [doi: 10.4324/9781003454533-1]
18. Franco CM, Bula D, Toso BRGO, Giovanella L, Hämel K. Nurse physician interprofessional collaboration in the context of expanding nursing practice in primary health care: a comparative qualitative study of Brazil and Germany. BMC Nurs 2025 Mar 7;24(1):259. [doi: 10.1186/s12912-025-02885-5] [Medline: 40055672]
19. Huang L, Chang H, Peng X, Zhang F, Mo B, Liu Y. Formally reporting incidents of workplace violence among nurses: a scoping review. J Nurs Manag 2022 Sep;30(6):1677-1687. [doi: 10.1111/jonm.13567] [Medline: 35213934]
20. Reinders JJ, Krijnen W. Interprofessional identity and motivation towards interprofessional collaboration. Med Educ 2023 Nov;57(11):1068-1078. [doi: 10.1111/medu.15096] [Medline: 37073763]
21. Borges do Nascimento IJ, Abdulazeem H, Vasanthan LT, et al. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. NPJ Digit Med 2023 Sep 18;6(1):161. [doi: 10.1038/s41746-023-00899-4] [Medline: 37723240]
22. Tong R, Brewer M, Flavell H, Roberts LD. Professional and interprofessional identities: a scoping review. J Interprof Care 2025;39(5):848-856. [doi: 10.1080/13561820.2020.1713063] [Medline: 32053408]
23. Tan AJQ, Chua WL, McKenna L, Tan LLC, Lim YJ, Liaw SY. Interprofessional collaboration in telemedicine for long-term care: an exploratory qualitative study. J Nurs Scholarsh 2023 Nov;55(6):1227-1237. [doi: 10.1111/jnu.12925] [Medline: 37482951]
24. Louwen C, Reidlinger D, Milne N. Profiling health professionals' personality traits, behaviour styles and emotional intelligence: a systematic review. BMC Med Educ 2023 Feb 18;23(1):120. [doi: 10.1186/s12909-023-04003-y] [Medline: 36803372]
25. McNulty JP, Politis Y. Empathy, emotional intelligence and interprofessional skills in healthcare education. J Med Imaging Radiat Sci 2023 Jun;54(2):238-246. [doi: 10.1016/j.jmir.2023.02.014] [Medline: 37032263]
26. Rayan-Gharra N, Tonkikh O, Gur-Yaish N. Family caregiver role in care recipient preparedness for hospital discharge: a parallel mediation model. PLoS ONE 2025;20(7):e0327758. [doi: 10.1371/journal.pone.0327758] [Medline: 40674334]
27. Dietl JE, Derksen C, Keller FM, Lippke S. Interdisciplinary and interprofessional communication intervention: how psychological safety fosters communication and increases patient safety. Front Psychol 2023;14:1164288. [doi: 10.3389/fpsyg.2023.1164288] [Medline: 37397302]
28. McLaney E, Morassaei S, Hughes L, Davies R, Campbell M, Di Prospero L. A framework for interprofessional team collaboration in a hospital setting: advancing team competencies and behaviours. Healthc Manage Forum 2022 Mar;35(2):112-117. [doi: 10.1177/08404704211063584] [Medline: 35057649]
29. Bornman J, Louw B. Leadership development strategies in interprofessional healthcare collaboration: a rapid review. J Healthc Leadersh 2023;15:175-192. [doi: 10.2147/JHL.S405983] [Medline: 37641632]
30. Keniston A, McBeth L, Pell Sr J Sr, et al. Development and implementation of a multidisciplinary electronic discharge readiness tool: user-centered design approach. JMIR Hum Factors 2021 Apr 23;8(2):e24038. [doi: 10.2196/24038] [Medline: 33890860]
31. Buljac-Samardzic M, Doekhie KD, van Wijngaarden JDH. Interventions to improve team effectiveness within health care: a systematic review of the past decade. Hum Resour Health 2020 Jan 8;18(1):2. [doi: 10.1186/s12960-019-0411-3] [Medline: 31915007]

XSL•FO
RenderX

32. Teuwen CJ, Kusurkar RA, Schreurs H, Daelmans HEM, Peerdeman SM. Interprofessional collaboration skills and motivation one year after an interprofessional educational intervention for undergraduate medical and nursing students. BMC Med Educ 2024 Mar 11;24(1):269. [doi: 10.1186/s12909-024-05262-z] [Medline: 38468246]
33. Corbaz-Kurth S, Weissbrodt R, Juvet TM, et al. How does trust emerge in interprofessional collaboration? A qualitative study of the significance, importance, and dynamics of trust in healthcare teams and networks. J Interprof Care 2025 May 15:1-10. [doi: 10.1080/13561820.2025.2495013] [Medline: 40370201]
34. Cadel L, Sandercock J, Marcinow M, Guilcher SJT, Kuluski K. A qualitative study exploring hospital-based team dynamics in discharge planning for patients experiencing delayed care transitions in Ontario, Canada. BMC Health Serv Res 2022 Dec 3;22(1):1472. [doi: 10.1186/s12913-022-08807-4] [Medline: 36463159]
35. Smith LM, Keiser M, Turkelson C, Yorke AM, Sachs B, Berg K. Simulated interprofessional education discharge planning meeting to improve skills necessary for effective interprofessional practice. Prof Case Manag 2018;23(2):75-83. [doi: 10.1097/NCM.0000000000000250] [Medline: 29381672]
36. Wener P, Leclair L, Fricke M, Brown C. Interprofessional collaborative relationship-building model in action in primary care: a secondary analysis. Front Rehabil Sci 2022;3(May):890001. [doi: 10.3389/fresc.2022.890001] [Medline: 36189054]
37. Johnson JM, Hermosura BJ, Price SL, Gougeon L. Factors influencing interprofessional team collaboration when delivering care to community-dwelling seniors: a metasynthesis of Canadian interventions. J Interprof Care 2021;35(3):376-382. [doi: 10.1080/13561820.2020.1758641] [Medline: 32400226]
38. Davidson AR, Kelly J, Ball L, Morgan M, Reidlinger DP. What do patients experience? Interprofessional collaborative practice for chronic conditions in primary care: an integrative review. BMC Prim Care 2022 Jan 14;23(1):8. [doi: 10.1186/s12875-021-01595-6] [Medline: 35172731]
39. Bentley M, Freeman T, Baum F, Javanparast S. Interprofessional teamwork in comprehensive primary healthcare services: findings from a mixed methods study. J Interprof Care 2018 May;32(3):274-283. [doi: 10.1080/13561820.2017.1401986] [Medline: 29182411]
40. Redzewsky L, Thielhorn U, Wallesch C, Körner M. The meaning of organizational culture and structure for interprofessional collaboration. Eur J Public Health 2024 Nov 1;34(Supplement_3). [doi: 10.1093/eurpub/ckae144.1624]
41. Ishii K, Fujitani K, Matsushita H. Interprofessional collaboration mediates the relationship between perceived organizational learning and safety climate in hospitals: a cross-sectional study. Int J Risk Saf Med 2024;35(3):217-232. [doi: 10.3233/JRS-230026] [Medline: 38759025]
42. Labrague LJ, Al Sabei S, Al Rawajfah O, AbuAlRub R, Burney I. Interprofessional collaboration as a mediator in the relationship between nurse work environment, patient safety outcomes and job satisfaction among nurses. J Nurs Manag 2022 Jan;30(1):268-278. [doi: 10.1111/jonm.13491] [Medline: 34601772]
43. Zanetoni TC, Cucolo DF, Perroca MG. Interprofessional actions in responsible discharge: contributions to transition and continuity of care. Rev Esc Enferm USP 2023;57:e20220452. [doi: 10.1590/1980-220X-REEUSP-2022-0452en] [Medline: 38047743]
44. Kutz A, Koch D, Haubitz S, et al. Association of interprofessional discharge planning using an electronic health record tool with hospital length of stay among patients with multimorbidity: a nonrandomized controlled trial. JAMA Netw Open 2022 Sep 1;5(9):e2233667. [doi: 10.1001/jamanetworkopen.2022.33667] [Medline: 36169957]
45. Dib K, Belrhiti Z. Unpacking the black box of interprofessional collaboration within healthcare networks: a scoping review. BMJ Open 2025 Jun 13;15(6):e101702. [doi: 10.1136/bmjopen-2025-101702] [Medline: 40514237]
46. Rawlinson C, Carron T, Cohidon C, et al. An overview of reviews on interprofessional collaboration in primary care: barriers and facilitators. Int J Integr Care 2021 Jun 22;21(2):32. [doi: 10.5334/ijic.5589] [Medline: 34220396]
47. Ginting AB, Made N, Santi D, Widjaja YR. Strategi membangun budaya organisasi yang inovatif dalam menghadapi transformasi digital di rumah sakit [Article in Indonesian]. MANEGGIO J Ilm Magister Manaj 2025;8(1):36-41. [doi: 10.30596/maneggio.v8i1.24840]
48. Palanisamy R, Taskin N, Verville J. Impact of trust and technology on interprofessional collaboration in healthcare settings: an empirical analysis.. Int J e-Collab 2017 Apr;13(2):10-44. [doi: 10.4018/IJeC.2017040102]
49. Reeves S, Xyrichis A, Zwarenstein M. Teamwork, collaboration, coordination, and networking: why we need to distinguish between different types of interprofessional practice. J Interprof Care 2018 Jan;32(1):1-3. [doi: 10.1080/13561820.2017.1400150] [Medline: 29131697]
50. Moloro AH, Sabo KG, Lahole BK, Wengoro BF, Mare KU. Prevalence of interprofessional collaboration towards patient care and associated factors among nurses and physician in Ethiopia, 2024: a systematic review and meta-analysis. BMC Nurs 2025 Feb 25;24(1):210. [doi: 10.1186/s12912-025-02847-x] [Medline: 40001025]
51. Labrague LJ, De Los Santos JAA. COVID-19 anxiety among front-line nurses: predictive role of organisational support, personal resilience and social support. J Nurs Manag 2020 Oct;28(7):1653-1661. [doi: 10.1111/jonm.13121] [Medline: 32770780]
52. Sentika R, Setiawan T, et al. The importance of interprofessional collaboration (IPC) guidelines in stunting management in Indonesia: a systematic review. Healthcare (Basel) 2024 Nov 7;12(22):1-16. [doi: 10.3390/healthcare12222226] [Medline: 39595424]

XSL•FO
RenderX

53. Degu T, Amsalu E, Kebede A, Adal O. Inter-professional collaboration and associated factors among nurses and physicians in specialized public hospitals, the Northwest, Ethiopia: mixed method multi-centered cross-sectional study. BMC Health Serv Res 2023 Mar 27;23(1):286. [doi: 10.1186/s12913-023-09200-5] [Medline: 36973734]
54. Yi S, Yam ELY, Cheruvettolil K, et al. Perspectives of digital health innovations in low- and middle-income health care systems from South and Southeast Asia. J Med Internet Res 2024 Nov 25;26:e57612. [doi: 10.2196/57612] [Medline: 39586089]
55. Han SP, Jumat MR, Cleland JA. Interprofessional collaboration (or lack thereof) between faculty and learning technologists in the creation of digital learning. BMC Med Educ 2023 Oct 4;23(1):727. [doi: 10.1186/s12909-023-04728-w] [Medline: 37794342]
56. Smith H, Harvey C, Portela A. Discharge preparation and readiness after birth: a scoping review of global policies, guidelines and literature. BMC Pregnancy Childbirth 2022 Apr 5;22(1):281. [doi: 10.1186/s12884-022-04577-3] [Medline: 35382773]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
FGD: focus group discussion
IPC: interprofessional collaboration
IPE: interprofessional education
IPPN: integrated patient progress notes
RKK: Surat Keterangan Kewenangan Klinik
SIP: Surat Izin Praktik
SOP: standard operating procedure
WHO: World Health Organization

Edited by E Borycki; submitted 11.Aug.2025; peer-reviewed by A Ganefianty, MA Zoubi; revised version received 29.Nov.2025; accepted 28.Dec.2025; published 09.Feb.2026.

Please cite as:
Delima M, Aljaberi MA, Dioso RIII
Nurses' Experiences of Interprofessional Collaboration in Digitally Supported Hospital Discharge Planning: Qualitative Study
JMIR Nursing 2026;9:e81961
URL: https://nursing.jmir.org/2026/1/e81961
doi:10.2196/81961

© Mera Delima, Musheer A Aljaberi, Regidor III Dioso. Originally published in JMIR Nursing (https://nursing.jmir.org), 9.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Explainable AI for Equitable Nurse Scheduling: Pragmatic Pre-Post Implementation Study

Ben-Chang Shia[1*], Prof Dr, PhD; Szu-Ming Peng[1*], PhD; Qui-Yang Zhang[1*], PGDM; Chiung-Yun Lo[2], BSN, MHA; Sheng-Ru Wang[3*], MD, EMBA

[1]Graduate Institute of Business Administration, College of Management, Fu Jen Catholic University, New Taipei City, Taiwan

[2]Deputy Director of Teaching and Research Department, St. Paul's Hospital, Taoyuan City, Taiwan

[3]Department of Pediatric Emergency, Fu Jen Catholic University Hospital, No. 69, Sec. 1, Gui-zi Road, Taishan Dist., New Taipei City, Taiwan

[*]these authors contributed equally

Corresponding Author:
Sheng-Ru Wang, MD, EMBA
Department of Pediatric Emergency, Fu Jen Catholic University Hospital, No. 69, Sec. 1, Gui-zi Road, Taishan Dist., New Taipei City, Taiwan

Abstract

Background: Inequitable and time-consuming shift scheduling contributes to nurse burnout, dissatisfaction, and turnover. In Taiwan, annual nurse turnover reaches 11.6%, with rigid 3-shift systems and unfair workload distribution frequently cited as key drivers. Although artificial intelligence (AI) scheduling tools exist, most lack transparency and do not formally address algorithmic bias, limiting clinical adoption.

Objective: This study aimed to design, deploy, and evaluate a transparent, fairness-audited, explainable AI–enabled nurse scheduling decision support system (XAI-NSDSS) to reduce administrative burden, eliminate experience-based algorithmic bias, and enhance staff acceptance in a real-world hospital setting.

Methods: A pragmatic before-after implementation study was conducted at a 671-bed teaching hospital in Taiwan (January-December 2023), involving 8 departments and 156 nurses (42 novice, 78 midlevel, and 36 experienced). A 6-month manual scheduling baseline (January-June 2023) was compared with a 6-month AI-assisted period (July-December 2023). The XAI-NSDSS integrates a random forest workload prediction model (R^2=0.887), Shapley Additive Explanations–based explainability, a hybrid integer programming and binary differential evolution (IP+ BDE) optimizer, and a multidimensional fairness monitoring dashboard. A formal weight sensitivity analysis (WSA) was conducted across 7 prespecified weight configurations using full-factorial repeated-measures ANOVA to assess outcome robustness. Primary outcomes were scheduling time, error rate, and user satisfaction. Statistical analyses used linear mixed effects models (LMMs) and generalized estimating equations (GEE) with department as a random effect.

Results: Monthly scheduling time decreased by 81.2% (mean 32.0, SD 8.0-mean 6.0, SD 2.0) hours; P<.001; Cohen d=4.33) and error rate decreased by 73.8% (mean 18.3, SD 4.3%-mean 4.8, SD 1.2%; P<.001; Cohen d=4.12). Nurse satisfaction improved from a mean of 3.2 (SD 0.8) to a mean of 4.4 (SD 0.6; P<.001), with 148 out of 156 nurses (94.9%) adopting the system by Month 3. Preexisting experience-based bias was fully eliminated: workload coefficient of variation (CV) decreased 50% (0.18-0.09; P<.001), disparate impact ratios normalized from 1.35 - 1.56 to 1.01 - 1.04, and preference satisfaction equity was achieved across experience tiers (ANOVA P=.38). Among 156 nurses, 82 (52.6%) regularly engaged with Shapley Additive Explanations; this engagement was positively associated with satisfaction (Pearson r=0.456; P<.001). The WSA across 7 configurations confirmed that the consensus-derived default weights achieved the highest composite quality score (mean 82.1, SD 3.2) and that disparate impact ratios remained within the 0.80 - 1.25 fairness threshold across all configurations (P=.12), demonstrating structural robustness of the fairness-auditing module.

Conclusions: This study presents the first longitudinally validated explainable AI implementation framework for nurse scheduling with formal algorithmic fairness auditing and WSA. The XAI-NSDSS framework is replicable, scalable, and provides a practical blueprint for responsible AI adoption in health care workforce governance, with fairness guarantees that are robust to institutional customization of optimization priorities.

(JMIR Nursing 2026;9:e94450) doi:10.2196/94450

KEYWORDS

nurse scheduling; workload equity; explainable artificial intelligence; algorithmic fairness; implementation science; decision support systems; workforce management

XSL•FO
RenderX

Introduction

Background and Rationale

The global nursing workforce crisis is one of the most pressing challenges in modern health care. The World Health Organization (WHO) projects a shortage of 5.9 million nurses by 2030, with high-income countries experiencing turnover exceeding 15% annually [1]. In Taiwan, the annual nursing turnover rate reached a 10-year high of 11.6%, with scheduling-related dissatisfaction identified as a leading contributor to burnout and early career exit [2,3]. Inefficient scheduling creates a vicious cycle; unfair shift distribution increases fatigue, which accelerates attrition, which further strains remaining staff.

Nurse scheduling is a complex combinatorial optimization problem requiring simultaneous consideration of regulatory constraints (labor laws and union agreements), operational requirements (shift coverage and skill mix), and individual staff preferences (time-off requests and work-life balance) [4]. Traditional manual scheduling is time-intensive, error-prone, and chronically fails to achieve fair workload distribution, particularly disadvantaging novice nurses who lack institutional leverage to negotiate favorable assignments [5].

Gaps in Existing AI Scheduling Systems

Artificial intelligence (AI) and operations research approaches have shown promise in nurse scheduling. Recent implementations include mixed integer programming heuristics [6], genetic algorithms [7], and machine learning–based prediction models [8]. However, a systematic review of the literature reveals four persistent, unresolved gaps:

1. Opacity and lack of explainability: most systems operate as "black boxes," generating schedules without interpretable rationale, a critical barrier to trust in high-accountability health care environments [9,10].
2. Absence of formal fairness auditing: while scheduling optimization papers frequently mention equity as an objective, formal auditing of algorithmic bias using established fairness metrics (eg, disparate impact ratios) remains virtually absent from the literature [11].
3. Insufficient real-world longitudinal validation: the majority of published systems are evaluated on simulated or retrospective datasets. Prospective, hospital-based evaluations are rare, limiting evidence for practical adoption [12,13].
4. Neglect of user trust mechanisms: system design rarely accounts for how transparency features influence actual adoption behavior, creating a gap between algorithmic performance and clinical uptake [14,15].

Study Objectives and Contributions

This study directly addresses all 4 gaps through the design, deployment, and 12-month evaluation (6-mo pre- and 6-mo postimplementation) of the explainable AI–enabled nurse scheduling decision support system (XAI-NSDSS) at a 671-bed tertiary referral hospital in Taiwan. The aim of this study was to demonstrate that a single integrated framework can simultaneously achieve operational efficiency, provable algorithmic equity, and clinician trust through explainability. Specific contributions are as follows:

1. Framework contribution: among the first end-to-end XAI implementation frameworks for nurse scheduling integrating Shapley Additive Explanations (SHAP)–based explainability, formal optimization, and a multidimensional algorithmic fairness monitoring system.
2. Empirical contribution: prospective real-world evidence from 156 nurses across 8 departments over 12 months, among the largest single-site longitudinal evaluations of an explainable artificial intelligence (XAI) nurse scheduling system reported to date.
3. Fairness contribution: formal documentation and complete elimination of preexisting experience-based algorithmic bias using disparate impact ratios, Gini coefficients, and ANOVA-based equity testing.
4. Trust mechanism contribution: empirical evidence that SHAP engagement (82/156, 52.6% nurses) is positively associated with satisfaction (r=0.456; P<.001), positioning explainability as a functional organizational intervention.
5. Implementation science contribution: actionable guidelines for change management, governance, and phased rollout based on observed adoption patterns, learning curves, and organizational barriers.

Methods

Study Setting and Design

This pragmatic before-after implementation study was conducted at Fu Jen Catholic University Hospital, a 671-bed tertiary referral hospital in New Taipei City, Taiwan, from January to December 2023. The XAI-NSDSS was deployed across 8 nursing departments—medical, surgical, intensive care, emergency, pediatrics, obstetrics, oncology, and geriatrics—serving 156 nurses (42 novice [<3 y experience], 78 midlevel [3 - 10 y], and 36 experienced [>10 y]).

We used a mixed methods evaluation design with a 6-month preimplementation baseline (manual scheduling; January-June 2023) compared to a 6-month postimplementation period (AI-assisted scheduling; July-December 2023). Randomization was not feasible due to the hospital-wide nature of the intervention and risk of cross-contamination; the large effect sizes observed (Cohen d>4) and consistency across all 8 departments strengthen confidence in attribution. This study is reported in accordance with the Standards for Quality Improvement Reporting Excellence (SQUIRE 2.0) guidelines.

Participants and Eligibility

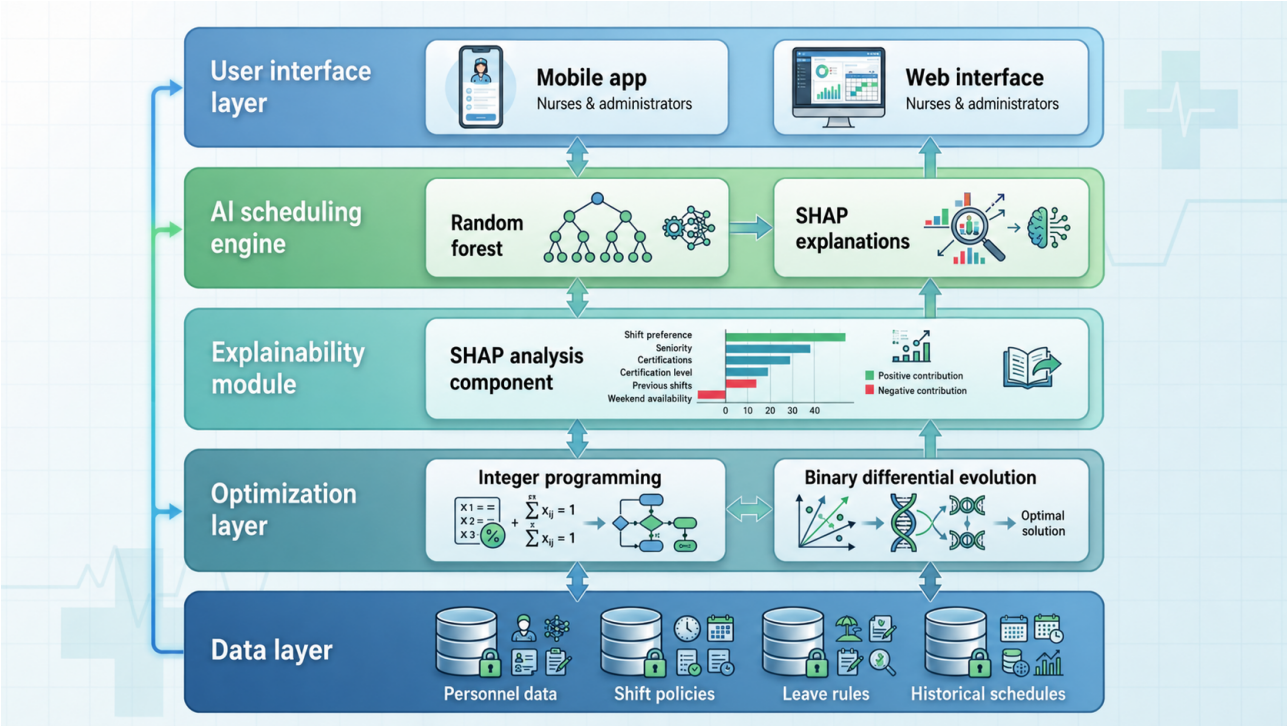
All registered nurses employed in the participating departments during the study period were eligible for inclusion. Exclusion criteria included (1) nurses on extended leave (>4 consecutive weeks), (2) temporary or agency staff without permanent employment contracts, and (3) nurses in administrative roles without direct patient care responsibilities. Of 162 eligible nurses, 156 (96.3%) participated in the full 12-month study period. Six nurses were excluded due to maternity leave (n=3), resignation (n=2), or transfer to nonparticipating departments (n=1).

XSL•FO
RenderX

System Architecture and Decision Pipeline

The XAI-NSDSS uses a modular 5-layer architecture (Figure 1; detailed in Section S2.1 in): (1) user interface layer—mobile-responsive dashboards (React.js 18.2.0; Meta Platforms, Inc); (2) AI engine layer—random forest regressors (scikit-learn 1.3.0; scikit-learn developers) for workload capacity prediction; (3) explainability module—SHAP TreeExplainer (shap v0.42.1; Scott Lundberg); (4) optimization layer—hybrid integer programming (IP; Gurobi 10.0.3; Gurobi Optimization, LLC)+binary differential evolution (BDE); and (5) data layer—PostgreSQL (15.3; PostgreSQL Global Development Group) with TimescaleDB (Tiger Data) and advanced encryption standard with a 256-bit key encrypted storage.

Figure 1. System architecture diagram. AI: Artificial Intelligence; SHAP: Shapley Additive Explanations.



XAI Components

SHAP-Enhanced Workload Prediction

Random forest regressors (200 trees, max_depth=15, min_samples_split=10, max_features="sqrt") predicted nurse workload capacity scores based on 18 input features spanning demographic characteristics, workload indicators, fatigue metrics, and preference alignment (Table S1 in). All continuous features were standardized using min-max normalization prior to model training. Models were trained on 1872 nurse-quarter observations (2021 - 2023) with 70/15/15 train/validation/test splits. Test-set performance: R^2=0.887, mean absolute error (MAE)=5.1, root-mean-square error (RMSE)=7.4. SHAP TreeExplainer values provided both global feature importance rankings and local, assignment-level explanations displayed to each nurse [16].

2D Assignment Logic

A 2D assignment framework balances professionalism scores (weighted composite of performance evaluations, certifications, and experience; S2.2 in) with fatigue indicators (consecutive shifts, overtime hours, night shift frequency, and recovery time). Exponential penalty functions capture the nonlinear clinical risk of accumulated fatigue. The key terms are defined as:

Professionalism score = 0.35 × Performance_Evaluation + 0.25 × Experience_Normalized + 0.20 × Certifications + 0.15 × Patient_Satisfaction + 0.05 × Peer_Collaboration

Fatigue score = 0.30 × Consecutive_Shifts_Penalty + 0.25 × Overtime_Penalty + 0.20 × Night_Shift_Penalty + 0.15 × Recovery_Deficit + 0.10 × Weekend_Penalty

where penalty functions follow exponential forms (eg, Consecutive_Shifts_Penalty=min[100, 10× e^(0.3× consecutive_shifts)]) to reflect the nonlinear accumulation of clinical risk.

Optimization Methodology

Hybrid IP+BDE Approach

Let x_ijs ∈ {0,1} denote whether nurse i (i ∈ N, |N|=156) is assigned to shift s (s ∈ S={day, evening, night, off}) on day j (j ∈ D={1,...,30}). The IP model enforces the following hard constraints, which carry zero tolerance for violation:

- Coverage: Σ x_ijs≥ R_js for all j ∈ D, s ∈ S (minimum staffing requirements per shift)
- One shift per day: Σs x_ijs≤1 for all i ∈ N, j ∈ D
- Maximum consecutive shifts: Σ_{k=j}^{j+5} Σ_{s≠off} x_iks≤5 for all i ∈ N, j ∈ D
- Minimum rest: x_{i,j,night}+ x_{i,j+1,day}≤1 for all i, j (ensuring ≥11 h intershift rest)

XSL•FO
RenderX

- Skill mix: Σ (q_ik× x_ijs)≥ Q_jsk for all j, s, k (minimum qualified staff per specialty)

These constraints are implemented as strict mathematical constraints within the IP model (Gurobi 10.0.3), guaranteeing 100% compliance in all generated schedules. The BDE component then minimizes the following multiobjective penalty-based fitness function after IP feasibility repair:

$$F_schedule=\lambda 1C_hard+\lambda 2C_soft+\lambda 3P_unmet+\lambda 4W_imbalance+\lambda 5F_total$$

where C_hard is the count of IP hard-constraint violations in the BDE trial vector, used only as a barrier function; C_soft is the sum of weekend and night-shift fairness violations across all nurse pairs; P_unmet = $\Sigma\ \Sigma_j\ \Sigma_s(1 - match_ijs) \times preference_strength_ijs$; W_imbalance = Var(monthly_hours) + Var(weekend_shifts) + Var(night_shifts) across all nurses; and F_total = $\Sigma\ \Sigma_j\ \Sigma_s$ fatigue_score(i,j) × x_ijs.

The penalty-scale coefficients were set as λ_1 =1000, λ_2 =100, λ_3 =50, λ_4 =30, λ =20, calibrated through stakeholder consultation and validated via Pareto frontier analysis.

Importantly, the λ_1 =1000 penalty in the BDE fitness function serves exclusively as a barrier function to prevent BDE from exploring infeasible regions, not as a relaxation of hard constraints. All BDE trial solutions with C_hard>0 are repaired to feasibility by the IP solver prior to fitness evaluation, ensuring hard constraints remain inviolable. The W_imbalance term acts as a proxy for demographic equity by minimizing variance in shift allocation across all nurses; post hoc disparate impact ratio analysis confirmed that this proxy effectively eliminated experience-based disparities (see the Fairness and Equity Analysis section and Table 1).

Table . Undesirable shift distribution by experience level. Expected proportion of novice nurses: 26.9% (42/156 nurses). Acceptable disparate impact range: 0.80 - 1.20 [11].

Shift type	Preimplementation: novice actual, % (n/N)[a]	Preimplementation: DIR[b]	Postimplementation: novice actual, % (n/N)[a]	Postimplementation: DIR
Night shifts	42.1 (69/164)	1.56[c]	27.4 (49/179)	1.02
Weekend shifts	38.5 (69/179)	1.43[c]	28.0 (49/175)	1.04
Overtime shifts	37.1 (69/186)	1.38[c]	27.4 (49/179)	1.02
Consecutive night shifts	36.3 (69/190)	1.35[c]	27.2 (49/180)	1.01

[a]n/N denotes novice-assigned undesirable shift events divided by total events in each category; denominators therefore vary across categories.

[b]DIR: disparate impact ratio.

[c]Preimplementation violation of four-fifths fairness threshold (P<.05 for all 4 categories).

The hybrid strategy exploits complementary strengths; IP guarantees hard constraint satisfaction (coverage, rest periods, consecutive shift limits, and skill mix), while BDE explores the feasible solution space to maximize preference satisfaction and workload equity. Computation averaged 12.7 minutes per monthly schedule. Full pseudocode is provided in S2.3 in Multimedia Appendix 1.

Multiobjective Optimization Module

The scheduling optimizer balances five competing objectives: (1) workload equity (minimizing coefficient of variation [CV] in shift assignments), (2) preference satisfaction (maximizing alignment with stated shift preferences), (3) skill-mix adequacy (ensuring appropriate distribution of experience levels across shifts), (4) regulatory compliance (satisfying labor law constraints on consecutive shifts and rest periods), and (5) operational efficiency (minimizing understaffing and overstaffing events).

The composite objective function is formulated as a weighted linear scalarization:

$$F=W1\cdot f1+W2\cdot f2+W3\cdot f3+W4\cdot f4+W5\cdot f5$$

where each subobjective is defined as follows:

1. f_1 (Workload equity): CV of monthly shift assignments across all nurses; minimized to promote equitable workload distribution.
2. f_2 (Preference satisfaction): proportion of fulfilled shift-preference requests; maximized (entered as 1, fulfillment rate to convert to a minimization problem).
3. f_3 (Skill-mix adequacy): deviation from target experience-level ratios per shift slot; minimized to ensure appropriate distribution of junior and senior nurses.
4. f_4 (Regulatory compliance): count of constraint violations (consecutive-shift limits, mandatory rest periods, overtime caps); minimized to zero as a hard constraint.
5. f (Operational efficiency): sum of absolute understaffing and overstaffing events across all shifts in a scheduling cycle; minimized to maintain safe staffing levels.

All subobjectives are normalized to the (0 and 1) range using min-max scaling derived from the historical baseline period (mo 1 - 6) prior to optimization. The default weight configuration deployed in clinical practice was established through a structured consensus process involving 4 nurse managers and 2 clinical informaticists, yielding w_1 =0.30, w_2 =0.25, w_3 =0.20, w_4 =0.15, and w =0.10, reflecting the institutional priority of fairness and preference satisfaction over strict operational efficiency (weight sensitivity analysis [WSA] systematically varies the normalized weights w_1 -w $\in$ [0,1] that govern the multiobjective trade-off, which are distinct from the penalty-scale barrier weights λ_1 =1000, λ_2 =100 etc used in the BDE fitness function). The complete formulation is available in S2.8 in Multimedia Appendix 1.

XSL•FO
RenderX

WSA Design

To formally assess the robustness of scheduling outcomes to variations in objective weights—and to address the inherent subjectivity of the consensus-derived default configuration—a prespecified WSA was conducted using a full-factorial grid design. Seven configurations were evaluated, spanning 3 clinically meaningful priority scenarios.

- Fairness-priority (FP) configurations (FP-Low, FP-Default, FP-High) systematically vary w_1 (workload equity) from 0.20 to 0.40 while proportionally redistributing the remaining weight budget across w_2 -w , holding their relative ratios constant. This isolates the marginal impact of fairness emphasis on all outcome domains.
- Preference-priority (PP) configurations (PP-Low and PP-High) vary w_2 (preference satisfaction) from 0.15 to 0.35, redistributing the residual weight proportionally across w_1 , w_3 -w . This tests whether increasing preference responsiveness compromises workload equity.
- Efficiency-priority (EP) configuration: sets w =0.30 as the dominant weight, reducing w_1 and w_2 to 0.20 each, to simulate a cost-minimization context where operational staffing coverage is paramount. An Equal-Weight (EW) configuration (all weights=0.20) serves as an unweighted baseline. For each configuration, the IP+BDE optimizer was rerun on the same 6-month intervention-period scheduling instances (n=48 monthly department schedules), and results were compared by one-way repeated-measures ANOVA with Bonferroni correction; the FP-Default configuration served as the reference group for all pairwise comparisons.

Data Collection and Outcome Measures

Outcome Measures

Primary outcomes included scheduling time (hours/schedule), error rate (%), and user satisfaction (5-point Likert scale, validated 8-item instrument; S2.4 in). Secondary outcomes included workload CV, Gini coefficient, disparate impact ratios by experience tier, preference satisfaction equity, and SHAP engagement rate (proportion of nurses viewing SHAP explanations at least once per scheduling cycle).

Statistical Analysis

Continuous outcomes were analyzed using linear mixed effects models (LMM) with time (pre/post) and month as fixed effects, and department as a random intercept, to account for clustering of nurses within departments. Count outcomes (error counts and constraint violations) were analyzed using generalized estimating equations (GEE, Poisson family, and exchangeable correlation structure). Effect sizes were estimated using Cohen *d* for continuous outcomes and incidence rate ratios (IRRs) for count outcomes, with 95% CIs reported in Table S2 in S1.2 in Multimedia Appendix 1. Three prespecified sensitivity analyses were conducted: (1) Hawthorne effect assessment comparing Months 1 - 3 versus 4 - 6, (2) monthly learning curve analysis, and (3) seasonal adjustment for Taiwan's major holiday periods. Statistical significance was set at α=.05 (2-tailed). All analyses were performed in R (version 4.3.0; R Core Team); full code is provided in S2.5 in Multimedia Appendix 1.

For the WSA, one-way repeated-measures ANOVA was used to compare outcome metrics across the 7 weight configurations, with the FP-Default configuration as the reference. Pairwise post hoc comparisons were adjusted using the Bonferroni correction. Effect sizes were reported as partial η^2. A configuration was deemed clinically noninferior to FP-Default if the 95% CI for the difference in workload equity CV was within ±0.02 and the preference fulfillment rate differed by less than 5 percentage points.

Qualitative Assessment

Semistructured interviews were conducted with a purposive sample of 24 nurses and 8 nurse managers, selected to ensure maximum variation across experience tiers (novice, midlevel, and experienced), departments (all 8 represented), and SHAP engagement levels (active users vs nonusers). Thematic saturation was achieved at interview 26, defined as no new codes emerging across 3 consecutive interviews. Interview guides addressed four domains: (1) perceived usability and system trust; (2) experience with SHAP explanations; (3) perceived fairness; and (4) barriers and facilitators to adoption. All 32 interviews were conducted in Mandarin Chinese, audio-recorded with written consent, and professionally transcribed verbatim (mean duration 38 min, range 28 - 52 min). Thematic analysis followed the 6-phase framework of Braun and Clarke [17]. Two authors (QYZ and CYL) independently coded all transcripts using NVivo 14. Interrater reliability was assessed on a randomly selected 20% subsample (κ=0.81, indicating strong agreement). Disagreements were resolved by consensus, with a third author (SRW) serving as arbitrator. Member checking was performed with 5 participants; negative cases were actively sought to challenge emerging themes. Participant characteristics are reported in Table S3 in S1.2 in Multimedia Appendix 1.

Fairness Monitoring Framework

Algorithmic fairness was audited quarterly across 3 dimensions: (1) demographic parity, proportional undesirable shift distribution across experience tiers; (2) equal opportunity, preference satisfaction equity (one-way ANOVA); and (3) disparate impact, selection rate ratios against the 0.80 - 1.20 threshold ("four-fifths rule") [11]. Gini coefficients provided complementary inequality measurement. Automated alerts upon threshold violations were reviewed monthly by the AI Scheduling Oversight Committee (ASOC), a multidisciplinary body comprising 8 nursing managers, 2 information technology staff, 1 clinical ethicist, 1 legal advisor, and 1 elected nurse representative (S2.6 in Multimedia Appendix 1 and dashboard: Figure S2 in S1.2 in Multimedia Appendix 1).

Governance and Data Privacy

System integrity was maintained through weekly automated model drift detection and quarterly retraining (threshold: R^2 drop >0.03 from baseline). SHAP explanations were scoped to individual nurse data only, preventing disclosure of colleagues' performance metrics. All 156 nurses completed a 1-hour domain-specific training session covering system usage and SHAP interpretation prior to deployment. All data were stored in databases encrypted using the advanced encryption standard

XSL•FO
RenderX

with a 256-bit key and protected by role-based access controls, in compliance with the Taiwan Personal Data Protection Act.

Ethical Considerations

The study was approved under an exemption waiver by the Institutional Review Board of Fu Jen Catholic University Hospital (IRB reference: FJUH-IRB-114‐459). All participating nurses and managers provided written informed consent prior to enrollment. Participation was voluntary, and participants were informed that they could withdraw from the study at any time without penalty or any effect on their employment status, work assignments, or institutional evaluation. All data were deidentified before analysis and reported only in aggregate form. Individual-level scheduling records, interview transcripts, SHAP engagement data, and satisfaction responses were stored in encrypted, access-controlled databases, and no personally identifiable information is disclosed in the manuscript or supplementary materials. No images or other materials containing identifiable participant information are included. No financial compensation or other incentives were provided to participants.

Results

Operational Performance Improvements

Table 2 summarizes key operational metrics. The XAI-NSDSS reduced monthly scheduling time by 81.2% (from mean 32.0, SD 8.0 to mean 6.0, SD 2.0 h; P<.001; Cohen d=4.33; 95% CI for difference −29.1 to −22.9 h), compressing a 3‐5 workday administrative burden to under 1 day. At the Emergency Department, reduction was most pronounced (45 → 8 h; 82.2%). Schedule error rates decreased by 73.8% (mean 18.3, SD 4.3% to 4.8 ± 1.2%; P<.001; d=4.12), with reductions in rest period violations (91.7% reduction, from 8.4 to 0.7 violations per schedule), hard constraint violations (86.9%), and skill mismatches (78.4%).

Table . Operational performance metrics: pre- versus postimplementation (n=156 nurses; 8 departments).

Metric	Preimplementation, mean (SD)	Postimplementation, mean (SD)	Change (%)	Cohen *d*	*P* value
Scheduling time (hours/month)	32.0 (8.0)	6.0 (2.0)	–81.2	4.33	<.001
Error rate (%)	18.3 (4.3)	4.8 (1.2)	–73.8	4.12	<.001
Constraint violations (per schedule)	12.7 (2.8)	1.3 (0.6)	–89.8	5.42	<.001
Preference satisfaction (%)	72.4 (5.1)	88.1 (2.8)	+21.7	3.87	<.001
Workload CV[a]	0.18 (0.03)	0.09 (0.02)	–50.0	3.74	<.001
Nurse satisfaction (1-5)	3.2 (0.8)	4.4 (0.6)	+37.5	1.71	<.001
Manager satisfaction (1-5)	3.5 (0.7)	4.7 (0.4)	+34.3	2.08	<.001

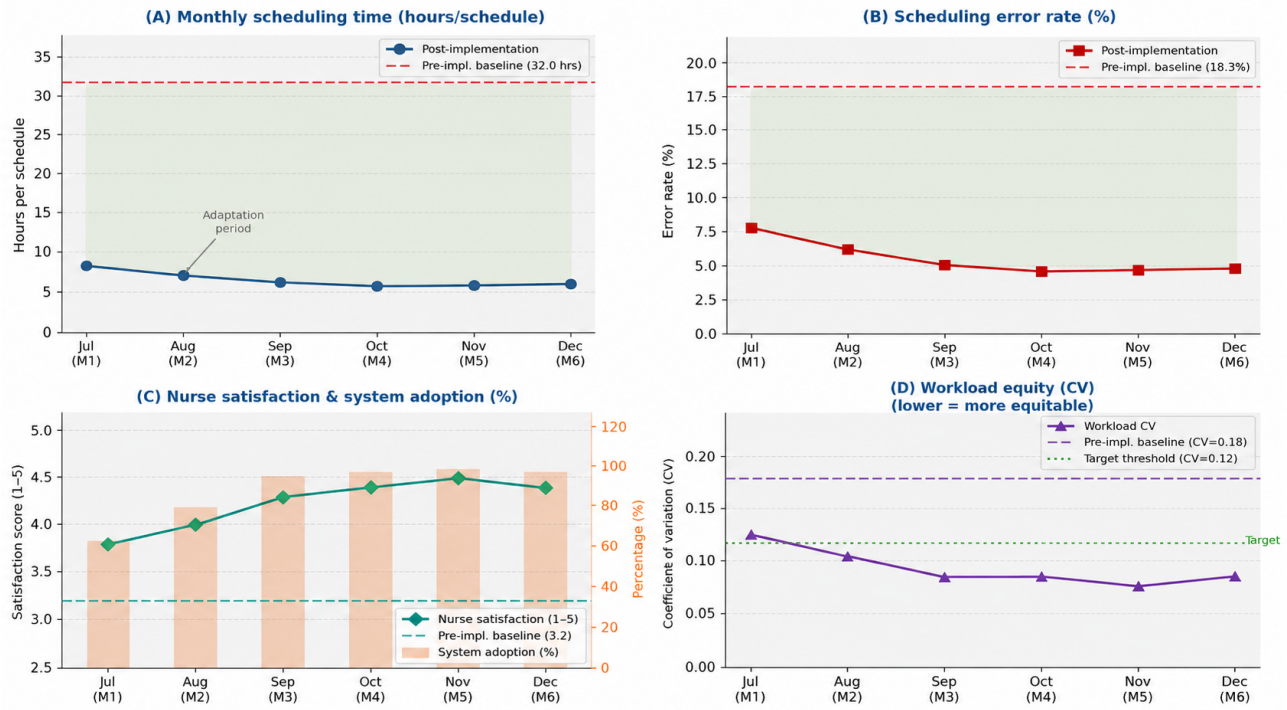
[a]CV: coefficient of variation

User Satisfaction and System Adoption

System adoption reached 148 out of 156 nurses (94.9%) by Month 3, with subsequent stability through Month 6 (Figure 2A-2D). The largest pre-to-post satisfaction gains were in transparency/explainability (mean 2.6, SD 1.0-mean 4.3, SD 0.6; P<.001; d=2.08) and fairness perception (mean 2.9, SD 1.0-mean 4.2, SD 0.7; P<.001; d=1.50). Trust in AI recommendations improved from a mean of 2.8 (SD) 0.9 to mean of 4.1 (SD 0.7; P<.001; d=1.62). Of 156 nurses, 82 (52.6%) regularly engaged with SHAP explanations, with a significant positive association between engagement frequency and overall satisfaction (Pearson r=0.456; P<.001; Figure S3 in). Full satisfaction subscale results are reported in Table S4 in Multimedia Appendix 1.

XSL•FO
RenderX

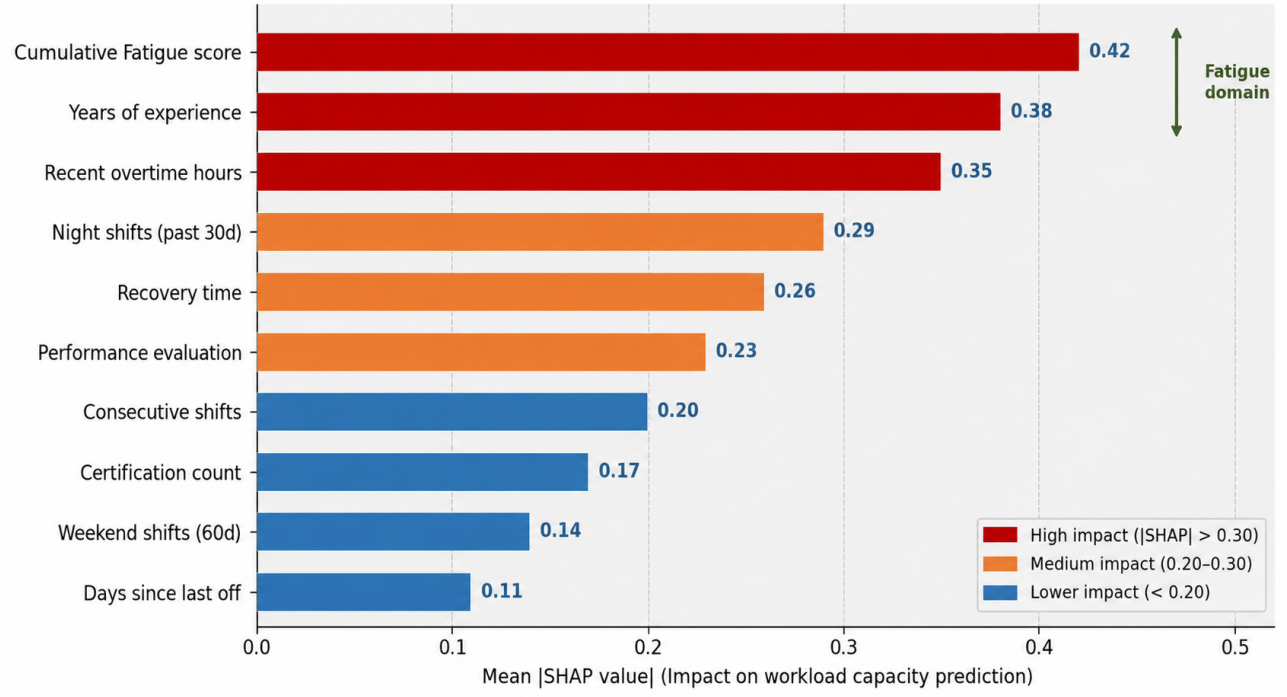
Figure 2. Monthly trend analysis: postimplementation performance (July-December 2023). (A) Scheduling time showing rapid decline from initial 8.2 hrs to stable 6.0 hrs, with preimplementation baseline (32.0 h) as dashed reference. (B) Error rate declining from 7.8% (Month 1) to stable 4.8%. (C) Nurse satisfaction improvement with system adoption reaching 94% by Month 3. (D) Workload CV sustained below the 0.12 target throughout, with preimplementation baseline (0.18) as reference.



SHAP Explainability Insights

Global SHAP analysis identified cumulative fatigue score (mean |SHAP|=0.42), years of experience (0.38), and recent overtime hours (0.35) as the 3 most influential scheduling determinants (Figure 3). Survey responses confirmed that SHAP explanations helped 122 out of 156 (78.2%) nurses understand shift assignments, increased trust for 111 out of 156 (71.2%) nurses, and identified actionable strategies to improve scheduling priority for 100 out of 156 (64.1%) nurses.

Figure 3. Global feature importance: mean absolute Shapley Additive Explanations (SHAP) values for the random forest workload capacity model (n=18 features). Bars colored by impact magnitude: high (red, |SHAP|>0.30), medium (amber, 0.20 - 0.30), lower (blue, <0.20). SHAP: Shapley Additive Explanations.



XSL•FO
RenderX

Algorithm Performance Comparison

The hybrid IP+BDE approach significantly outperformed either component in isolation (Table 3). BDE-only achieved only 95.2% hard constraint compliance versus IP's guaranteed 100%. IP-only achieved only 78.3% preference satisfaction versus the hybrid's 88.1%. Computation time averaged 12.7 minutes per monthly schedule versus 32.0 hours for manual scheduling (99.3% reduction). Emergency rescheduling resolved within 5 - 8 minutes.

Table . Algorithm performance comparison. Results are based on a 6-month postimplementation period (n=48 schedules).

Method	Hard constraint compliance (%)	Preference satisfaction (%)	Workload CV[a]	Computation time
Manual scheduling (baseline)	81.7	72.4	0.18	32.0 h
IP[b] only	100	78.3	0.14	8.2 min
BDE[c] only	95.2	85.6	0.11	15.4 min
Hybrid IP+BDE (proposed)	100	88.1	0.09	12.7 min

[a]CV: coefficient of variation.

[b]IP: integer programming.

[c]BDE: binary differential evolution.

Fairness and Equity Analysis

Workload Distribution Equity

Postimplementation workload equity improved substantially. Hospital-wide CV decreased 50.0% (mean 0.18, SD 0.03-mean o 0.09, SD 0.02; P<.001; d=3.74; 95% CI for mean difference: −0.10 to −0.08). The Gini coefficient improved 45.8% (0.24-0.13), corroborating CV findings. Post implementation, experience level no longer predicted workload variability ($F_{2,153}$=1.42; P=.24), and improvements were uniform across all 8 departments ($F_{7,148}$=0.89; P=.51). Monthly workload CV trend data are shown in Figure 2D.

Elimination of Systematic Bias in Undesirable Shift Allocation

As shown in Table 1, novice nurses were overrepresented in all 4 undesirable shift categories before implementation, accounting for 42.1% (69/164) of night shift assignments, 38.5% (69/179) of weekend shift assignments, 37.1% (69/186) of overtime shift assignments, and 36.3% (69/190) of consecutive-night events, despite representing 26.9% (42/156) of the nursing workforce. Postimplementation, these proportions decreased to 27.4% (49/179), 28.0% (49/175), 27.4% (49/179), and 27.2% (49/180), respectively, with disparate impact ratios normalized from 1.35-1.56 to 1.01-1.04. This represents a measurable, auditable transition from an algorithmically biased manual system to a provably equitable AI-assisted system.

Preference Satisfaction Equity

Preimplementation, experienced nurses received significantly higher preference satisfaction than novice nurses (76.2% vs 68.4%; gap=7.8 percentage points; P=.009). Postimplementation, this disparity was eliminated (ANOVA P=.38; all pairwise comparisons P>.30), with all 3 experience tiers achieving satisfaction rates ≥87%. Equity was achieved through a "leveling up" strategy—absolute gains >13 percentage points for all groups (P<.001)—rather than redistributing the burden from experienced to novice nurses.

Sensitivity Analyses

Hawthorne effect analysis (Months 1 - 3 vs 4 - 6) confirmed sustained benefits, scheduling time remained stable (mean 6.2, SD 2.1 vs mean 5.8, SD 1.9 h; P=.34); satisfaction continued its upward trend (mean 4.2, SD 0.7-mean 4.6, SD 0.5; P=.03). Workload CV remained below the preimplementation level across all 6 months (range 0.08 - 0.10), confirming sustained improvement rather than novelty effect. Seasonal adjustment revealed no significant interaction between time and holiday period (P=.41). Detailed results are provided in Table S5 in Multimedia Appendix 1.

WSA

Beyond the prespecified sensitivity analyses reported above, a formal WSA was conducted to assess whether scheduling outcomes depended on the choice of objective-function weights. WSA was conducted across all 7 prespecified configurations (Table 4). One-way repeated-measures ANOVA revealed a significant main effect of weight configuration on workload equity CV ($F_{,222}$=18.43; P<.001; η^2=0.28), preference fulfillment rate ($F_{,222}$=22.17; P<.001; η^2=0.32), and composite schedule quality score ($F_{,222}$=9.84; P<.001; η^2=0.17). No significant main effect was observed for the disparate impact ratio range ($F_{,222}$=1.72; P=.12), indicating that the fairness-auditing module maintained equitable shift distribution regardless of weight settings.

XSL•FO
RenderX

Table . Scheduling outcome metrics across weight sensitivity configurations (n=48 schedules per configuration).

Configuration	Workload equity CV[a], mean (SD)	Preference fulfillment (%)	Disparate impact ratio range	Composite quality score (0 - 100)
FP-Low[b]	0.13 (0.02)	81.4 (4.1)	0.98 - 1.06	74.2 (3.8)
FP-Default (Clinical)	0.09 (0.02)	84.6 (3.7)	1.01 - 1.04	82.1 (3.2)
FP-High	0.07 (0.01)[c]	79.3 (4.8)[c]	1.00 - 1.03	79.8 (4.1)
PP-Low[d]	0.11 (0.02)[c]	76.2 (5.2)[c]	1.01 - 1.05	76.4 (4.5)[c]
PP-High	0.12 (0.03)[c]	89.1 (3.2)[c]	1.02 - 1.07	78.3 (3.9)[c]
EP[e]	0.15 (0.03)[c]	74.8 (5.6)[c]	1.03 - 1.09	71.6 (4.7)[c]
EW[f]	0.11 (0.02)[c]	80.7 (4.3)[c]	1.00 - 1.06	77.5 (3.6)[c]

[a]CV: coefficient of variation.

[b]FP: fairness-priority.

[c]Significantly different from FP-Default (Bonferroni-corrected $P<.05$).

[d]PP: preference-priority.

[e]EP: efficiency-priority.

[f]EW: equal-weight.

Workload Equity

The FP-Default configuration achieved the best balance between equity (CV=0.09±0.02) and overall quality. FP-High yielded marginally lower CV (0.07±0.01; mean difference −0.02, 95% CI −0.03 to −0.01; $P=.008$), but at the cost of a significant reduction in preference fulfillment (79.3% vs 84.6%; mean difference −5.3 percentage points, 95% CI −7.8 to −2.8; $P<.001$). The EP configuration produced the highest CV (0.15±0.03; mean difference 0.06; 95% CI 0.04-0.08; $P<.001$), confirming that deprioritizing workload equity substantially degraded fairness outcomes.

Preference Satisfaction

PP-High maximized preference fulfillment (mean 89.1, SD 3.2%), exceeding FP-Default by mean difference 4.5 percentage points (95% CI 2.1 - 6.9; $P=.001$). However, this gain was accompanied by deterioration in workload equity (CV=0.12±0.03 vs 0.09±0.02; $P=.003$) and a reduction in composite quality score (78.3 vs 82.1; $P=.01$). Conversely, PP-Low reduced preference fulfillment to 76.2% (mean difference=−8.4% percentage points; $P<.001$) while offering no significant improvement in any other outcome domain, suggesting $w_2 = 0.15$ is below the minimum threshold for acceptable preference responsiveness in this institutional context.

Disparate Impact Ratios

Critically, all 7 weight configurations maintained disparate impact ratios within the 0.80 - 1.25 fairness threshold (range across all configurations: 0.98 - 1.09), and the ANOVA revealed no significant between-configuration differences ($P=.12$). This finding demonstrates that the fairness-auditing module provides a structural guarantee against experience-based bias that is robust to changes in optimization priorities.

Composite Schedule Quality

FP-Default achieved the highest composite quality score (mean 82.1, SD 3.2), significantly outperforming all alternative configurations (all Bonferroni-corrected $P<.05$) except FP-High (mean 79.8, SD 4.1; mean difference=−2.3, 95% CI −4.8 to +0.2; $P=.07$). Applying the prespecified noninferiority criteria (CV within ±0.02 and preference fulfillment within ±5 percentage points), only FP-High met both criteria simultaneously. All other configurations failed at least one criterion, supporting the robustness of the FP-Default configuration as the clinically optimal choice for this institution.

Monthly Trend Analysis

Figure 2 presents monthly trend data across all 4 primary performance domains during the postimplementation period. The 2-month adaptation period followed by sustained stability aligns with the "implementation cliff" described in implementation science literature [18] and supports the recommendation of a 3-month intensive support period for similar deployments. Detailed monthly data are provided in Table S6 in Multimedia Appendix 1.

Qualitative Findings

Semistructured interviews (n=24 nurses; n=8 managers; mean duration 38 min) yielded 5 major themes. A full participant characteristics table is provided in Table S3 in Multimedia Appendix 1.

Transparency as a Conflict-Resolution Mechanism (21 of 24 Nurses; 8 of 8 Managers)

SHAP-derived rationale for unfavorable assignments substantially reduced grievance escalations. One nursing manager stated, "There are fewer complaints about unfair schedules. When someone questions an assignment, I can show them the SHAP explanation."

Fairness as a Structural Achievement (18 of 24 Nurses; 7 of 8 Managers)

Nurses attributed equity improvements to objective system design rather than managerial goodwill, increasing systemic trust.

XSL•FO
RenderX

Administrative Burden Relief Enabling Clinical Refocus (8 of 8 Managers)

All managers described time savings as professionally transformative. Emergency department manager said, “I used to spend an entire week every month on scheduling. Now it’s done in a day.”

Learning-to-Trust Trajectory (19 of 24 Nurses)

Initial resistance in Months 1 - 2 diminished substantially by Month 3. A nurse with 7 years of experience stated, “At first I was skeptical, but once I learned how to input my preferences properly, it worked really well.”

Residual Tensions: Preference Override Acceptance (11 of 24 Nurses)

Explainability reduces but does not eliminate individual dissatisfaction when equity constraints conflict with strong personal preferences.

Discussion

Principal Findings

This study demonstrates that a fairness-audited, XAI scheduling system can simultaneously reduce administrative burden by 81.2%, eliminate preexisting experience-based algorithmic bias, and achieve 94.9% (148/156 nurses) system adoption in a real-world hospital setting. Three findings warrant particular emphasis: (1) preexisting experience-based bias was formally documented and completely eliminated; (2) SHAP engagement was significantly associated with user satisfaction; and (3) sustained improvements confirmed by sensitivity analyses ruled out novelty effects.

Algorithmic Fairness as a Primary Clinical Contribution

The most significant finding is the systematic identification and complete elimination of preexisting experience-based bias. Preimplementation, 69 out of 156 novice nurses were over-represented in all 4 undesirable shift categories by 26% - 46% (disparate impact ratios 1.35 - 1.56), exceeding the established four-fifths rule threshold for systemic discrimination [11]. This trend aligns with documented informal seniority hierarchies in nursing workplaces [19] but had not previously been quantified with formal fairness metrics in the AI scheduling literature.

The complete normalization of all disparate impact ratios postimplementation (1.01 - 1.04) demonstrates that embedding W_imbalance and F_total into the objective function can mitigate institutionalized bias. The W_imbalance term minimizes variance in shift allocation across all nurses, and post hoc disparate impact analysis confirmed this proxy effectively eliminated experience-based disparities, validating the chosen surrogate metric. The Gini coefficient improvement (0.24-0.13) further confirms equity advancements. Future implementations should treat preexisting bias characterization as a mandatory predeployment audit step.

SHAP Explainability as a Functional Trust Mechanism

A central hypothesis was that SHAP-based explainability would serve as an active trust-building mechanism. Our data show a significant positive association between SHAP engagement and satisfaction (r=0.456; P<.001). While causality cannot be established from this observational correlation—satisfied nurses may be more inclined to engage with explanations—the qualitative data consistently suggest that explanations functionally mitigate organizational conflict when used, consistent with a bidirectional reinforcing relationship. This extends work by Liu et al [20] on “integrability” as a determinant of AI acceptance, demonstrating that explanations must be actively used—not merely available—to generate trust benefits.

Critically, while 82 out of 156 nurses (52.6%) regularly viewed SHAP explanations, this study measured engagement frequency rather than comprehension accuracy. A formal assessment of whether nurses correctly interpreted feature importance rankings was not conducted, leaving a “comprehension gap” unquantified. Future research should use validated XAI literacy instruments to assess whether engagement translates to understanding.

The 52.6% rate of voluntary SHAP engagement surpasses engagement rates reported for passive transparency features in health care AI [21]. The 1-hour domain-specific training investment appears to be an effective and scalable strategy for facilitating XAI adoption.

Contextualizing Performance Against Prior Literature

Our 73.8% error rate reduction exceeds the 40% - 60% range reported in comparable scheduling optimization studies [22,23], attributable primarily to the human-in-the-loop verification enabled by SHAP transparency. The hybrid IP+BDE approach outperformed both single-method alternatives, consistent with the demonstration by Ben Said et al [24] that hybrid constraint-learning pipelines outperform single-method approaches in health care scheduling. Regarding the role of XAI versus IP optimization, XAI in this system is not merely a technical overlay but a functional organizational mechanism. Without SHAP explanations, the system would generate algorithmically optimal schedules that clinicians cannot audit, verify, or trust, making sustained adoption unlikely regardless of algorithmic performance. The 94.9% adoption rate at Month 3 versus typical rates for opaque AI systems substantiates this claim [9,10].

Weight Sensitivity and Optimization Robustness

The WSA addresses a fundamental challenge in multiobjective optimization for health care applications, the selection of objective weights is inherently value-laden and context-dependent; yet, the robustness of outcomes to weight perturbations has rarely been formally evaluated in prior nurse scheduling studies. Our findings contribute 3 key insights to this literature.

First, the consensus-derived FP-Default configuration (w_1 =0.30; w_2 =0.25) achieved the highest composite schedule quality score and was the only configuration meeting prespecified noninferiority criteria on both workload equity and

XSL•FO
RenderX

preference fulfillment simultaneously. This validates the stakeholder co-design process as an effective method for identifying near-optimal weight configurations in institutional contexts, consistent with findings from participatory AI scheduling design studies [10,19].

Second, the analysis revealed an inherent trade-off between workload equity and preference satisfaction that cannot be eliminated through algorithmic design alone. Increasing w_1 beyond 0.30 (FP-High) improved equity marginally but reduced preference fulfillment by 5.3 percentage points—a clinically meaningful reduction given that preference accommodation is a primary driver of nurse satisfaction and retention [5,22]. Conversely, maximizing preference satisfaction (PP-High) degraded workload equity, potentially reintroducing the experience-based disparities that the system was designed to eliminate. This trade-off surface highlights the irreducible nature of the fairness–preference tension and underscores the importance of transparent weight disclosure to affected stakeholders.

Third, and most critically for regulatory compliance, the disparate impact ratios remained within the 0.80 - 1.25 fairness threshold across all 7 weight configurations (P=.12 for between-configuration differences). This structural invariance demonstrates that the fairness-auditing module provides a hard constraint on algorithmic bias that operates independently of optimization priorities, a design principle with direct implications for responsible AI deployment in health care workforce governance [11,25]. Institutions adopting the XAI-NSDSS framework can therefore customize weight configurations to reflect local priorities without compromising the fairness guarantees central to equitable workforce governance.

The differential engagement patterns across experience levels suggest that explainability features are particularly valuable for less-experienced staff who may have less institutional power to challenge unfair scheduling decisions. This finding aligns with qualitative research by Gerlach et al [19] identifying transparency as a critical adoption factor for junior nurses, and extends it by quantifying the relationship between explanation engagement and satisfaction outcomes.

Implementation Science Perspective

The initial satisfaction dip (Month 1) followed by sustained improvement (Months 3 - 6) is consistent with the "implementation cliff" [18]. Our data suggest a 2 - 3 month stabilization window as the critical support period. The ASOC governance structure—with monthly fairness dashboard reviews and multidisciplinary oversight—addresses the requirements for organizational accountability in AI deployment reported by Saeed et al [25], distinguishing this implementation from most published systems that lack ongoing governance postlaunch.

Limitations

Single-site implementation in Taiwan limits generalizability to other health care systems, labor regulations, and cultural contexts; multisite validation is needed before broad adoption recommendations can be made. The 6-month postimplementation period does not address long-term system drift or multiyear organizational embedding. The quasiexperimental before-after design inherently limits causal inference: despite large effect sizes (Cohen d>3.7 across all primary outcomes), secular trends, regression to the mean, or co-occurring organizational changes cannot be fully excluded. SHAP comprehension accuracy was not formally assessed—only viewing frequency—leaving a "comprehension gap" unquantified. Patient care outcomes (eg, medication error rates and fall incidence) were not measured; connecting scheduling equity to downstream clinical outcomes remains an important research direction. The SHAP-satisfaction correlation is observational and directional causality cannot be established without experimental manipulation.

The WSA was conducted within a single institutional context; the consensus-derived FP-Default weight configuration (w_1 =0.30, w_2 =0.25) reflects this institution's scheduling culture, labor regulations, and workforce demographics. While the structural robustness of fairness metrics across all 7 configurations is likely to generalize, the absolute weight values may require site-specific calibration in institutions with different operational priorities. Future multisite replications should include systematic WSA to determine whether FP-Default represents a transferable standard or an institution-specific optimum.

Future Research Directions

Priority directions include (1) multisite validation across diverse health care systems, (2) randomized trials comparing explainable versus nonexplainable scheduling systems to establish the causal impact of XAI on trust and adoption, (3) integration with real-time patient acuity data, and (4) federated learning for privacy-preserving multi-institutional model development [26].

Regarding optimization robustness, adaptive weight optimization methods—such as Bayesian preference elicitation or reinforcement learning from stakeholder feedback—could automate the weight configuration process and reduce reliance on expert consensus, potentially improving outcomes in institutions lacking access to multidisciplinary co-design teams. The WSA framework developed in this study should be replicated in multisite contexts to determine whether the FP-Default weight configuration generalizes across institutions with different staffing cultures and regulatory environments or whether site-specific calibration is required.

Conclusions

This study presents a prospectively validated, fairness-audited XAI implementation framework for nurse scheduling delivering measurable operational improvements while maintaining interpretability and human oversight. Three key findings are noteworthy:

1. The formal documentation and complete elimination of preexisting experience-based algorithmic bias—disparate impact ratios normalized from 1.35 - 1.56 to 1.01 - 1.04—establishes a replicable methodology for fairness auditing in health care workforce AI.
2. The positive association between SHAP engagement and satisfaction (r=0.456), combined with qualitative evidence of explanations as conflict mediators, is consistent with

XSL•FO
RenderX

explainability functioning as an organizational intervention, though directionality cannot be confirmed from observational data alone.

3. The complete implementation science documentation—governance structure, phased rollout, learning curve data, and sensitivity analyses—provides a practical guide for responsible AI adoption in health care workforce management. WSA across 7 objective-weight configurations confirmed that fairness guarantees remain intact regardless of optimization priority settings (disparate impact ratio P=.12 across all configurations), validating the framework for diverse institutional deployment contexts.

The XAI-NSDSS framework demonstrates that algorithmic decision support can simultaneously achieve efficiency, equity, and transparency. The framework is replicable, scalable, and aligned with emerging standards for responsible AI governance in clinical settings.

Acknowledgments

The authors gratefully acknowledge the participating nurses and nurse managers at Fu Jen Catholic University Hospital for their dedication to this study. We thank the hospital administration for supporting this research and providing access to scheduling data and operational resources. We also acknowledge the contributions of the hospital's Information Technology Department for technical infrastructure support. Generative artificial intelligence (AI) tools were used to assist with drafting and editing selected sections (eg, phrasing and structure). All scientific content, analyses, interpretations, and references were verified by the authors, who take full responsibility for the manuscript.

Funding

This research was supported by the National Science and Technology Council, Taiwan (Grant Number: NSTC 114-2622-8-030-002-TD2). The funding agency had no role in study design, data collection, analysis, interpretation, or manuscript preparation.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to institutional privacy policies protecting employee data but are available from the corresponding author on reasonable request and with appropriate data use agreements. The explainable AI–enabled nurse scheduling decision support system (XAI-NSDSS) source code is available from the corresponding author on reasonable request for noncommercial research purposes, subject to institutional approval.

Authors' Contributions

Conceptualization: BCS, SRW
Data curation: SMP
Formal analysis: BCS, QYZ
Funding acquisition: SRW
Investigation: QYZ, CYL
Methodology: BCS, SRW
Project administration: BCS
Resources: CYL
Software: BCS, SMP
Supervision: SRW
Validation: SMP
Visualization: SMP
Writing – original draft: BCS
Writing – review & editing: QYZ, CYL, SRW
All authors read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflicts of interest. No author has financial relationships with commercial scheduling software companies or other entities that could inappropriately influence this work. The XAI-NSDSS system is not commercially available, and the authors have no plans for commercialization.

Multimedia Appendix 1
Additional figures, tables, extended methods, additional results, and implementation details for the explainable artificial intelligence–enabled nurse scheduling decision support system.
[DOCX File, 1068 KB - nursing_v9i1e94450_app1.docx]

XSL•FO
RenderX

References

1. State of the world's nursing 2020: investing in education, jobs and leadership. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/9789240003279 [accessed 2026-05-19]
2. Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. JAMA 2002;288(16):1987-1993. [doi: 10.1001/jama.288.16.1987] [Medline: 12387650]
3. Dall'Ora C, Ball J, Reinius M, Griffiths P. Burnout in nursing: a theoretical review. Hum Resour Health 2020 Jun 5;18(1):41. [doi: 10.1186/s12960-020-00469-9] [Medline: 32503559]
4. Burke EK, De Causmaecker P, Berghe GV, Van Landeghem H. The state of the art of nurse rostering. J Sched 2004 Nov;7(6):441-499. [doi: 10.1023/B:JOSH.0000046076.75950.0b]
5. Barker HR, Griffiths P, Dall'Ora C. "I don't think there's necessarily a one size fits all" negotiating competing priorities in nurse shift scheduling: a qualitative study. BMC Nurs 2025 Aug 11;24(1):1048. [doi: 10.1186/s12912-025-03705-6] [Medline: 40790199]
6. Turhan AM, Bilgen B. A hybrid fix-and-optimize and simulated annealing approaches for nurse rostering problem. Comput Ind Eng 2020 Jul;145:106531. [doi: 10.1016/j.cie.2020.106531] [Medline: 33868622]
7. Aickelin U, Dowsland KA. An indirect genetic algorithm for a nurse-scheduling problem. Comput Oper Res 2004 Apr;31(5):761-778. [doi: 10.1016/S0305-0548(03)00034-0]
8. Eshghali M, Kannan D, Salmanzadeh-Meydani N, Esmaieeli Sikaroudi AM. Machine learning based integrated scheduling and rescheduling for elective and emergency patients in the operating theatre. Ann Oper Res 2023 Jan 19;332(1-3):1-24. [doi: 10.1007/s10479-023-05168-x] [Medline: 36694896]
9. Knight DRT, Aakre CA, Anstine CV, et al. Artificial intelligence for patient scheduling in the real-world health care setting: a metanarrative review. Health Policy Technol 2023 Dec;12(4):100824. [doi: 10.1016/j.hlpt.2023.100824]
10. Renggli FJ, Gerlach M, Bieri JS, Golz C, Sariyar M. Integrating nurse preferences into AI-based scheduling systems: qualitative study. JMIR Form Res 2025 Jun 4;9:e67747. [doi: 10.2196/67747] [Medline: 40466089]
11. Char DS, Shah NH, Magnus D. Implementing machine learning in health care - addressing ethical challenges. N Engl J Med 2018 Mar 15;378(11):981-983. [doi: 10.1056/NEJMp1714229] [Medline: 29539284]
12. Tun HM, Rahman HA, Naing L, Malik OA. Trust in artificial intelligence-based clinical decision support systems among health care workers: systematic review. J Med Internet Res 2025 Jul 29;27:e69678. [doi: 10.2196/69678] [Medline: 40772775]
13. Tjoa E, Guan C. A survey on explainable artificial intelligence (XAI): toward medical XAI. IEEE Trans Neural Netw Learn Syst 2021 Nov;32(11):4793-4813. [doi: 10.1109/TNNLS.2020.3027314] [Medline: 33079674]
14. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: a review of machine learning interpretability methods. Entropy (Basel) 2020 Dec 25;23(1):18. [doi: 10.3390/e23010018] [Medline: 33375658]
15. Adadi A, Berrada M. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access 2018;6:52138-52160. [doi: 10.1109/ACCESS.2018.2870052]
16. Oh MY, Kim HS, Jung YM, Lee HC, Lee SB, Lee SM. Machine learning-based explainable automated nonlinear computation scoring system for health score and an application for prediction of perioperative stroke: retrospective study. J Med Internet Res 2025 Mar 19;27:e58021. [doi: 10.2196/58021] [Medline: 40106818]
17. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
18. Chang SJ, Lee YW, Chou WJ. Self-scheduling: a win-win for nurses and organizations. Hu Li Za Zhi 2025 Oct;72(5):12-18. [doi: 10.6224/JN.202510_72(5).03] [Medline: 41034164]
19. Gerlach M, Renggli FJ, Bieri JS, Sariyar M, Golz C. Exploring nurse perspectives on AI-based shift scheduling for fairness, transparency, and work-life balance. BMC Nurs 2025 Sep 2;24(1):1161. [doi: 10.1186/s12912-025-03808-0] [Medline: 40898192]
20. Liu Y, Liu C, Zheng J, Xu C, Wang D. Improving explainability and integrability of medical AI to promote health care professional acceptance and use: mixed systematic review. J Med Internet Res 2025 Aug 7;27:e73374. [doi: 10.2196/73374] [Medline: 40773743]
21. Cutillo CM, Sharma KR, Foschini L, et al. Machine intelligence in healthcare-perspectives on trustworthiness, explainability, usability, and transparency. NPJ Digit Med 2020;3:47. [doi: 10.1038/s41746-020-0254-2] [Medline: 32258429]
22. Kang HW, Kim J, Kim KJ, et al. Shift nurses' work quality and job satisfaction after implementing the Inha University hospital nursing AI scheduling system (IH-NASS). BMC Nurs 2025 Jul 1;24(1):792. [doi: 10.1186/s12912-025-03470-6] [Medline: 40597282]
23. Patel V, Deodhar A, Birru D. A multi-objective genetic algorithm for healthcare workforce scheduling. arXiv. Preprint posted online on Aug 28, 2025. [doi: 10.48550/arxiv.2508.20953]
24. Ben Said A, Mouhoub M. Machine learning and constraint programming for efficient healthcare scheduling. Int J Soft Eng Knowl Eng 2026 Jun 30;36(8):1089-1120. [doi: 10.1142/S0218194025501025]
25. Saeed M, Jalil MS, Dahwal FM, et al. The impact of AI on healthcare workforce management: business strategies for talent optimization and IT integration. tajmspr 2025;07(3):136-156. [doi: 10.37547/tajmspr/Volume07Issue03-15]

XSL•FO
RenderX

26. Rieke N, Hancox J, Li W, et al. The future of digital health with federated learning. NPJ Digit Med 2020;3:119. [doi: 10.1038/s41746-020-00323-1] [Medline: 33015372]

Abbreviations

AI: artificial intelligence
ASOC: AI Scheduling Oversight Committee
BDE: binary differential evolution
CV: coefficient of variation
EP: efficiency-priority
EW: equal-weight
FP: fairness-priority
GEE: generalized estimating equations
IP: integer programming
IRR: incidence rate ratio
LMM: linear mixed effects models
MAE: mean absolute error
PP: preference-priority
RMSE: root-mean-square error
SHAP: Shapley Additive Explanations
SQUIRE: Standards for Quality Improvement Reporting Excellence
WHO: World Health Organization
WSA: weight sensitivity analysis
XAI: explainable artificial intelligence
XAI-NSDSS: explainable AI–enabled nurse scheduling decision support system

Edited by E Borycki; submitted 01.Mar.2026; peer-reviewed by BMT Lin, H Koruca; revised version received 06.Apr.2026; accepted 05.May.2026; published 02.Jul.2026.

Please cite as:
Shia BC, Peng SM, Zhang QY, Lo CY, Wang SR
Explainable AI for Equitable Nurse Scheduling: Pragmatic Pre-Post Implementation Study
JMIR Nursing 2026;9:e94450
URL: https://nursing.jmir.org/2026/1/e94450
doi:10.2196/94450

© Ben-Chang Shia, Szu-Ming Peng, Qui-Yang Zhang, Chiung-Yun Lo, Sheng-Ru Wang. Originally published in JMIR Nursing (https://nursing.jmir.org), 2.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effects of Adequate Rest on Nurse Job Satisfaction, Burnout Prevention, and Physical Health in Medical and Emergency Units at a Hospital in Western Jamaica: Qualitative Study

Channon Smith[1*], MPH; Keitumetse-Kabelo Murray[2*], MA, MPhil; Natasha Croome[2*], PhD

[1]Health Services Department, The Heart Foundation of Jamaica, 28 Beechwood Ave, Kingston 5, Kingston, Jamaica

[2]School of Public Health, Faculty of Medicine, Imperial College London, London, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Channon Smith, MPH
Health Services Department, The Heart Foundation of Jamaica, 28 Beechwood Ave, Kingston 5, Kingston, Jamaica

Abstract

Background: The demanding work environment of nurses in medical and emergency units often results in high stress, job dissatisfaction, and burnout. Adequate rest is crucial for maintaining nurses' physical health, mental clarity, and emotional resilience, yet it is often overlooked in these high-pressure settings. This qualitative study explores the perceptions of nurses at a hospital in Western Jamaica regarding the quality and duration of rest they receive and its impact on their professional, mental, physical, and personal well-being. The hospital was selected due to the unique challenges health care workers face in Jamaica, including limited resources, high patient loads, and frequent staff shortages, which may exacerbate rest-related issues.

Objective: This study aimed to explore the perceptions of registered nurses working in the emergency and medical units of the hospital in Western Jamaica regarding their rest experience and its implications for burnout, job satisfaction, and overall health.

Methods: The study used a constructivist epistemological lens and used purposive sampling to select 12 registered nurses. The principal researcher conducted in-depth interviews with each participant via Zoom, using a semistructured guide. Interviews lasted 25 to 45 minutes, were audio-recorded, and attended only by participants and the researcher. Thematic analysis was used to transcribe, code, and analyze the data, culminating in the development of a thematic map of findings.

Results: The findings indicated that nurses face significant challenges in obtaining adequate rest due to staff shortages, heavy workloads, irregular shifts, and limited management support. A total of three primary themes emerged: (1) noncompliance with rest policies, (2) resource limitations, and (3) management issues, each influencing job satisfaction, burnout, and overall health. Within noncompliance, nurses highlighted suboptimal nurse-to-patient ratios, absenteeism, and inadequate break time. For example, ratios as high as "30 to 2" or "60 to 3" were cited, affecting nurses' ability to take breaks. Resource constraints included inadequate staffing, insufficient staff replacement, and the absence of suitable rest areas. Management concerns included weak policy enforcement, inadequate policy awareness, and limited support for rest breaks. These challenges collectively contributed to poor sleep quality, increased stress, and diminished job satisfaction.

Conclusions: The study highlights the need for systemic improvements to address nurse rest and well-being, including increased staffing, structured policies on break enforcement, and enhanced management engagement. While the study is specific to the hospital in Western Jamaica, the findings may have broader implications for health care systems in similarly resource-constrained settings in the Caribbean and other low- and middle-income regions.

(JMIR Nursing 2026;9:e84106) doi:10.2196/84106

KEYWORDS

emergency nursing; Jamaica; job satisfaction; nurse burnout; nurse-patient ratio; nurse well-being; qualitative research; rest; staffing shortages

Introduction

Registered nurses (RNs) are among the most vital resources globally, a reality that was illuminated during the COVID-19 pandemic [1]. The pandemic not only highlighted the essential role that nurses play in maintaining health care systems but also brought global attention to the immense pressures they face [1]. As reported by the British Broadcasting Corporation, nurses were at the frontline, often working extended hours under extreme conditions, which magnified the importance of addressing issues such as adequate rest, burnout, and job satisfaction [2,3]. This growing global realization of the system relevance of the nursing profession continues to shape

XSL•FO
RenderX

discussions on health care reforms and support for the nursing workforce.

While nurses are vital in delivering health care and patient care, there has been less focus on their health promotion. The demanding nature of nursing in medical and emergency units puts professionals at high risk for stress and burnout [3,4]. The World Health Organization defines burnout as emotional exhaustion from chronic workplace stress, leading to fatigue and decreased performance [5,6]. Reports of burnout among RNs are prevalent, particularly in high-pressure environments like the US health care system [7]. However, there is a noticeable gap in the literature regarding the role of rest in preventing burnout among nurses, particularly in Caribbean or low-resource contexts. Most existing research focuses on high-income countries, leaving a lack of context-specific understanding in regions such as Jamaica.

RNs are expected to provide high-quality care while managing intense workloads and irregular hours, underscoring the need for effective strategies to enhance nurse well-being and job satisfaction [7]. Job satisfaction reflects a positive emotional response to one's role and work environment [6,8,9]. Among these strategies, ensuring adequate rest is a crucial yet often overlooked component, essential for physical recovery and maintaining mental clarity and emotional resilience [10,11].

Despite being an upper-middle-income country, Jamaica faces a low nurse-to-patient ratio that negatively affects the rest quality and job satisfaction of RNs [12]. In contrast, socio-economically similar countries like Cuba and the Dominican Republic have much higher ratios [12,13]. Given the significant role of rest in job satisfaction, burnout prevention, and overall health, there is a lack of comprehensive research on its effects in Jamaican health care settings. This study aims to fill this gap by examining the perceptions of nurses at the hospital regarding the quality and duration of their rest and its impact on their well-being. This study focuses specifically on the medical and emergency units due to their particularly high levels of stress and burnout, as well as practical access considerations. These wards also represent high-acuity environments where the effects of inadequate rest are likely to be most pronounced. In fact, burnout rates on these wards are notably higher, ranging from 25% to 55% [14,15].

Through an in-depth analysis of how rest influences nurse satisfaction and burnout, this research seeks to generate evidence that can drive meaningful improvements in health care work environments. By focusing on the specific context of Western Jamaica, an underresearched region facing high patient loads and limited staffing, this study offers context-specific insights into the impact of rest on nurse well-being and performance. Its findings aim to support the development of evidence-based policies and institutional practices that prioritize staff recovery, reduce burnout, and ultimately improve patient outcomes. Beyond contributing to academic understanding, this research aspires to influence workforce planning and retention strategies, offering actionable recommendations to help build a more resilient, efficient, and sustainable health care system in Jamaica.

Methods

Setting

The hospital, established in 1964, is a key facility in western Jamaica. As a type B hospital, it provides 5 basic specialties: general surgery, internal medicine, obstetrics and gynecology, orthopedics, and pediatrics, ensuring comprehensive care for its diverse population. Although its capacity is 190 patients, it currently houses up to 300, leading to significant overcrowding and challenges for the health care system and medical staff [16].

This was a qualitative study that used purposive sampling to identify research participants, grounded in a constructivist epistemological approach. This perspective recognized the coconstruction of knowledge between the researcher and participants, aligning with the study's aim to explore the subjective experiences of nurses [17].

The principal researcher acknowledged their potential influence on the research process, particularly given prior acquaintance with 2 participants. Efforts were made to remain self-aware and neutral during interviews and analysis, to minimize bias and enhance the credibility of the findings.

Participant Recruitment

Permission was obtained from the Director of Nursing Services at the hospital on April 8, 2024, after discussing the research's objectives, methodology, and potential impacts. A formal written request outlining the research aims, methodology, ethical considerations, and data management was submitted on April 11, 2024. Following approval, a signed permission letter was issued. The hospital administration assisted in ethically disseminating recruitment emails to 15 potential participants (see Table 1 for inclusion/exclusion criteria). While all 15 acknowledged receipt, only 12 participated: 1 did not sign the consent form, and 2 failed to respond after signing.

XSL•FO
RenderX

Table . Inclusion and exclusion criteria.

Criteria	Inclusion	Exclusion
Employment	Registered nurses currently employed in the medical or emergency units at the hospital.	Nurses who do not work in the medical or emergency units at the hospital.
Clinical experience	Minimum of 1 year of clinical experience in their respective units.	Nurses with <1 year of clinical experience in their respective units.
Educational qualification	Possession of a Bachelor of Science degree in Nursing or a higher-level nursing qualification.	Nurses who do not possess a Bachelor of Science degree in Nursing or a higher-level nursing qualification.
Willingness to participate	Willingness to participate in qualitative interviews discussing their experiences and perceptions related to rest, job satisfaction, burnout, and physical health.	Nurses who are unwilling to participate in qualitative interviews discussing their experiences and perceptions related to rest, job satisfaction, burnout, and physical health.
Availability	Availability to participate in a 45 - to 60-minute interview session, either in person or virtually.	Nurses who are not available to participate in a 45 - to 60-minute interview session, either in person or virtually.
Language proficiency	Ability to understand and communicate in English effectively, as the interviews will be conducted in English.	Nurses who are unable to understand or communicate effectively in English, as the interviews will be conducted in English.
Male nurses	Male nurses who meet the above criteria are included, ensuring a diverse representation within the study.	Male nurses who do not meet the above criteria are excluded to maintain consistency in the participant pool and ensure a focused analysis.

Data Collection and Analysis

An interview guide was developed with open-ended questions exploring nurses' practices in managing rest, perceptions of rest's impact on satisfaction and burnout, and the role of hospital policies in promoting well-being. It also included recommendations for improving rest in the medical and emergency units at the hospital (see Multimedia Appendix 1).

Prior to the interview, participants were emailed the participant information sheet and subsequently the consent form and were asked to sign them via DocuSign. Interviews, averaging 45 minutes, were all conducted via Zoom using a semistructured guide with only the participants and the principal researcher present. Each session was audio recorded for accuracy, with no photos or videos taken. Data from interviews were transcribed, coded, and analyzed thematically following the 6-step framework of Braun and Clarke [18], which involved familiarization with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and writing up. An inductive approach was used, allowing themes to emerge from the data without being driven by preexisting theories or frameworks [18].

The trustworthiness of the data was reinforced through transparency in the research process and ongoing critical reflection. While one researcher led the initial coding, all authors reviewed the coding framework and contributed to the interpretation of findings. This collaborative process helped ensure consistency, reliability, and analytical depth, supported by clear documentation maintained throughout. Feedback from all authors further shaped the development of the analysis and structure of the paper, enhancing its overall rigor and quality.

Ethical Considerations

The research was conducted in adherence to ethical guidelines outlined by the Declaration of Helsinki and followed institutional protocols to ensure quality, integrity, and ethical responsibility [19]. The methodology was rigorously designed to ensure reliability, validity, and participant protection, following best practices for qualitative research [20].

Ethical approval was granted by the Research Governance and Integrity Team at Imperial College London (ethics application ID 7069331).

Data management was robust, with clear documentation and secure storage. Paper forms were scanned into the primary author's OneDrive space and then securely shredded. The data were securely stored. Any sensitive data archived were encrypted and access was restricted to authorized personnel only. Risks were managed by transparent communication, voluntary participation, and anonymization of data. To anonymize the data, personal identifiers such as names were removed and replaced with pseudonyms. Unique codes were assigned to each participant, and any identifying information was stored separately from the research data to maintain confidentiality.

All participants received a Participant Information Sheet prior to recruitment and provided written informed consent electronically before taking part in the study. Participation was voluntary, and participants were informed of their right to withdraw at any time without consequence. No financial or material compensation was provided for participation.

XSL•FO
RenderX

Results

Participant Characteristics

Twelve RNs participated in the study (see Table 2). The mean age of the participants was 28 (SD 2.2; range 25 - 33) years. All nurses held a Bachelor of Science in Nursing degree, with 1 nurse also possessing a critical care certificate.

Participants had a mean of 3.6 (SD 1.8; range 1.2 - 6) years of experience working at the hospital. The sample was evenly distributed across clinical units, with 50% (6/12) of participants assigned to the Accident and Emergency (A&E) unit and 50% (6/12) to the Medical unit.

Table . Demographic and professional characteristics of nurses at the hospital: female participant 1 - 11, male participant 12; numbers indicate the chronological order of interviews.

Pseudonyms	Age (y)	Educational background	Experience at hospital (y)	Current unit	Sex
Female Participant 1	29	BSN[a]	5	A&E[b]	Female
Female Participant 2	28	BSN	4.5	Medical	Female
Female Participant 3	33	BSN	5.5	Medical	Female
Female Participant 4	28	BSN	5	Medical	Female
Female Participant 5	30	BSN	3	Medical	Female
Female Participant 6	29	BSN	5	A&E	Female
Female Participant 7	27	BSN	1.5	A&E	Female
Female Participant 8	25	BSN	1.2	Medical	Female
Female Participant 9	29	BSN+critical care certificate	6	A&E	Female
Female Participant 10	27	BSN	3	A&E	Female
Female Participant 11	28	BSN	1	Medical	Female
Male Participant 12	32	BSN	2	A&E	Male

[a]BSN: Bachelor of Science in Nursing.

[b]A&E: Accident and Emergency unit.

Themes

Overview

Three key themes emerged from the interviews, on factors influencing rest quality and its impact on job satisfaction, burnout, and physical well-being. These themes were as follows: (1) noncompliance with rest policies, with subthemes of high nurse-patient ratios, high absenteeism, and rest duration; (2) resources, including limited human resources and the absence of rest facilities; and (3) management, focusing on policy improvement and implementation, as well as nonadherence to duties (see Multimedia Appendix 2 for coding table).

Noncompliance With Rest Policies

Overview

Noncompliance with rest policies among nurses in public hospitals is a persistent and multifaceted issue that has serious implications for staff well-being and patient care. Although formal guidelines are in place to ensure that nurses receive adequate breaks during their shifts, various systemic challenges make it difficult to adhere to these policies. Three critical subthemes emerged in relation to this problem: high nurse-patient ratios, high absenteeism, and inadequate rest duration.

High Nurse-Patient Ratios

The issue of noncompliance with rest policies among nurses is exacerbated by unsustainable nurse-patient ratios, making it nearly impossible for nurses to take their designated 1-hour breaks. Participants reported ratios as high as “sometimes 30 to 2, 35 to 2, 60 to 3, it varies” (Male Participant 12: A&E), highlighting the overwhelming workload they face. It is also common for a single RN to manage a unit with only an enrolled assistant nurse, meaning that while the RN receives assistance, they are still solely responsible for the entire unit, including supervising the enrolled assistant nurse. The consensus was clear; without addressing these staffing imbalances, compliance with rest policies will remain a significant challenge.

High Absenteeism

High absenteeism among nurses is a significant consequence of noncompliance with rest policies, as many nurses report having various medical illnesses, feeling overwhelmed, and burnt out, leading them to take frequent sick days. One participant noted, “the call-in rate is very high because when you realize that you are burnt out and tired you’ll find that persons are not coming in” (Female Participant 1: Medical). Another participant echoed this sentiment, stating, “I will wake up in the morning and say, OK, yes, I’m going to make it to work today… I just find that I am tired, not just physically tired, but emotionally tired” (Female Participant 2: Medical). This

XSL•FO
RenderX

chronic fatigue often results in nurses prioritizing their health over work obligations. The pervasive culture of exhaustion and the lack of adequate rest contribute to a cycle of absenteeism that further strains the already limited nursing staff, ultimately impacting patient care and overall hospital operations.

Rest Duration

Rest duration remains a significant issue among nurses in public hospitals in Jamaica. Although the policies mandate a 1-hour break during 8-hour day shifts and 2 hours for night shifts, these rest periods are rarely observed in practice. One participant verbalized, "You're supposed to get one hour in the day shift and two hours in the night, but we don't get that" (Female Participant 6: A&E). Participants highlighted that despite these official guidelines, the reality of high patient complexity and understaffing often makes it impossible to take the full allotted break or even time to eat. Many nurses expressed frustration with the gap between policy and practice, noting that the workload and staffing shortages leave little time for adequate rest. This chronic lack of rest not only exacerbates fatigue, medical illnesses, and burnout but also negatively impacts patient care [21,22].

Resources

Overview

The availability and quality of resources, particularly human resources and physical infrastructure, play a critical role in shaping nurses' ability to rest during their shifts. Inadequate resources contribute significantly to poor rest quality, increased burnout, and decreased job satisfaction. Two key subthemes emerged under this category: limited human resources and the absence of adequate rest facilities.

Limited Resources (Including Human Resources)

The theme of resources, particularly limited human resources, emerged as one of the main factors influencing rest quality and, consequently, job satisfaction, burnout, and physical well-being among nurses. Participants consistently expressed concerns about inadequate staffing levels, which directly impact their ability to take necessary breaks. Among all, 1 nurse articulated the challenge succinctly when asked if management does not actively hire new staff, stating, "If you go there now and say, oh, we need staff, they're going to say based on the quota that they have… but be reminded they have opened a lot of different areas and the population has expanded" (Female Participant 9: A&E). The overwhelming workloads resulting from these staffing shortages leave little room for rest. The lack of adequate resources not only hinders compliance with rest policies but also exacerbates feelings of burnout, as nurses struggle to manage their responsibilities without sufficient support.

Absence of Adequate Rest Facilities

The lack of adequate rest facilities at the hospital severely affects nurses' ability to recuperate during shifts. All participants expressed dissatisfaction with the current designated rest areas, citing issues such as overcrowding, noise from nearby units, and the combination of a bed and lunch area with a microwave in the same space, posing safety and health risks. While some nurses resorted to resting in their cars, others had no choice but to endure the suboptimal conditions. This shows that without proper rest facilities, nurses struggle to fully recover during their shifts, which in turn affects their physical well-being and their ability to provide quality patient care.

Management

Overview

The role of hospital management, particularly nursing leadership, is central to ensuring that rest policies are effectively implemented and that nurses are supported in their demanding roles. However, participants highlighted ongoing management-related challenges that undermine nurse well-being and disrupt the delivery of quality care. Two key subthemes emerged: the need for policy improvement and implementation, and nonadherence to managerial duties.

Policy Improvement and Implementation

Participants expressed a clear need for more effective policies that not only address staffing levels but also prioritize the well-being of nursing staff. Among all, 1 participant noted, "I think we need more policies to actually not just cater for the staffing of the hospital… but also to cater to the nurses" (Female Participant 1: Medical). Despite the existence of policies that outline break times, the implementation of these policies is often lacking. As one nurse stated, "the policy exists… however, there's no implementation of the actual policy" (Female Participant 7: A&E). This sentiment was echoed by another participant who remarked, "I don't think there is a collaborative effort among the hospital, administration, and nurses in promoting nurse well-being through proper rest practices" (Female Participant 3: Medical). The need for management to actively engage in policy enforcement and to create a supportive environment for nurses is paramount, as inadequate rest not only affects nurse satisfaction but also compromises patient care and safety.

Nonadherence to Duties

The issue of nursing managers not adhering to their duties at the hospital has been a significant concern among the interviewees. Participants expressed frustration over the lack of support from nursing managers, particularly during critical times when the unit is short-staffed. One participant noted, "the sisters are supposed to come there and assist and ensure that the unit is running to full capacity… but you find that when you fall into an emergency situation… they either tell you that they are short-staffed or they tell you did you call this ward for this" (Female Participant 4: Medical). This lack of responsiveness leaves nurses feeling overwhelmed and unsupported, as they are often left to manage high patient loads and intense emergency situations without adequate assistance. Another participant highlighted that "most of their tasks, they leave it for the nurses to do while they basically do nothing" (Female Participant 9: A&E), indicating a perceived neglect of managerial responsibilities on the units. The absence of proactive engagement from nursing managers not only exacerbates the challenges faced by nurses but also compromises patient care and safety, as the staff is unable to effectively manage their duties under such conditions.

XSL•FO
RenderX

Differences Based on Gender, Unit Type, and Experience Level

To further contextualize these findings, differences based on gender, unit type, and experience level were also observed among participants, offering deeper insight into how individual and situational factors shape nurses' experiences with rest policy compliance.

Gender Differences

The male nurse often highlighted issues related to workload and understaffing with a strong focus on managerial support and policy enforcement. For example, the male participant in the A&E unit emphasized frustration with the lack of implementation of break policies. Female nurses frequently discussed challenges around balancing work demands with personal responsibilities, such as family care, which impacted their ability to rest adequately during shifts. They also noted more about the emotional toll and burnout symptoms.

Unit Differences (A&E vs Medical)

Nurses working in A&E units reported higher stress levels due to patient complexity and unpredictability of cases. They described fewer opportunities for breaks and greater difficulty taking rest because they were often the only RN on the unit during shifts. Nurses in the Medical units acknowledged the challenges of patient care but reported slightly more opportunities for breaks compared to A&E. However, they also noted the workload increased significantly during night shifts.

Experience and Role

More experienced nurses, working 3 years or more, tended to express frustration with systemic issues such as staffing policies and managerial accountability. Less experienced nurses, working less than 3 years, were more likely to discuss immediate physical fatigue and emotional exhaustion, focusing on day-to-day survival rather than broader systemic changes.

Discussion

Summary of Major Findings

This study examined the role of adequate rest in nurse satisfaction, burnout prevention, and physical well-being in medical and emergency units. It uncovered systemic challenges that hinder nurses from obtaining sufficient rest. High patient-to-nurse ratios, staffing shortages, and lack of managerial support were identified as key contributors to fatigue and burnout. Nurses' personal responsibilities, such as caregiving at home, also affected their ability to prioritize rest. These findings show how both institutional and personal factors compromise nurses' well-being and, in turn, patient care.

Interpretation of Demographic Differences

Demographic differences reflect social, cultural, and professional dynamics shaping nurses' experiences. Female nurses often balance work and caregiving roles, heightening stress and burnout risk [23]. A&E nurses reported higher fatigue levels than those in Medical units, due to the unpredictable nature of emergency care. Experienced nurses highlighted systemic issues, while newer nurses focused on the immediate physical and emotional toll, reflecting their frontline pressures [24].

Interpretation of Findings

The study reaffirms the link between rest and well-being. Nurses with more rest reported better physical and mental health, supporting global research on this topic [25]. Challenges at this hospital, such as high workloads and weak managerial support, mirror trends in other resource-limited settings [26]. For example, a Namibian study also found burnout tied to understaffing and lack of rest infrastructure [27]. Nurses at the hospital in Western Jamaica frequently reported physical exhaustion and long-term health concerns, echoing global findings on rest-related health risks, including musculoskeletal pain and emotional exhaustion, key components of burnout [7,25,28].

Relation to Wider Context and Integration With Existing Literature

Jamaica faces a severe brain drain, with 80% of skilled workers emigrating, including health care professionals [29]. Ranking second globally on the brain drain index [30], Jamaica's workforce shortages worsen burnout. Similar trends are noted in Guyana and Trinidad and Tobago [31]. Cultural and economic pressures often lead nurses to work extended hours without adequate compensation or rest. In contrast, high-income countries enforce stricter work hour regulations. At Spanish Town Hospital in Southeast-Central Jamaica, the nurse-patient ratio is 1:10, far above ratios in wealthier nations, contributing to burnout and absenteeism [32-34]. The Maslach Burnout Inventory indicated "very high" burnout among nurses, especially in emotional exhaustion [22,35]. The International Council of Nurses [36] highlights the negative outcomes that occur from high-income countries attempting to address their nursing shortages through "inequitable international recruitment." Through recruiting via migration, it leaves nursing workforces in low- and middle-income countries without adequate care, masks the underlying issues leading to high turnover, and costs low- and middle-income countries lost training expenses after public investment in education.

Nurses struggle to maintain work-life balance in these settings. Overcrowding and understaffing lead to long hours and little time for self-care [37,22]. The State of the World's Nursing Report 2025 [38] illustrated that only 55% of countries had regulations on working hours and conditions, whereas the remaining 45% had partial or no regulations. Care packages for mental well-being of nurses were implemented in 42% of countries, whereas 64% only implemented partial or no care packages. A study in Iceland showed higher satisfaction among nurses working standard hours versus those on overtime [39]. However, extended shifts remain necessary at this hospital in Western Jamaica, negatively affecting health and morale. These findings call for urgent action to support nurses and improve patient care through systemic reform. This is supported by the International Council of Nurses 2025 report [36], which states that solutions such as "ensuring adequate staffing and a balanced skill mix and workforce capacity aligned with patients demands" need to be implemented.

XSL•FO
RenderX

Strengths and Limitations

This study's strength lies in its qualitative design, which captured rich personal narratives often missed in quantitative research [17]. Familiarity between the researcher and some participants may have encouraged openness. However, it may also have introduced bias. The lead researcher's nursing background may have shaped interpretations. Additionally, excluding non-English-speaking nurses, such as Cuban staff [1], limited the diversity of views. With only 12 participants from 2 units, generalizability is limited. Self-reported data also carry risks of under- or over-reporting.

Future Work

Future studies should explore strategies for staff recruitment and retention to reduce burnout. Research is needed on how managerial practices affect rest and well-being. Peer support systems and cultural change around rest and self-care should be evaluated. Post-COVID recovery efforts should prioritize mental health support and enforceable rest policies [40].

Policy and Intervention Recommendations

The hospital should address high nurse-to-patient ratios by increasing staff, ensuring breaks without compromising care. California's Nurse-to-Patient Ratios Law (1:2 in intensive care units; 1:4 in medical-surgical) reduced burnout and improved outcomes [4]. Structured break policies and cross-unit support systems are vital. UK hospitals implementing scheduled breaks reported reduced stress and higher job satisfaction [41]. Australia's cross-unit model ensures continuity of care and relieves pressure during breaks [39,42,43].

Hospital leadership should prioritize nurse well-being. Programs in Australia train managers to promote self-care and regular breaks, leading to greater satisfaction and lower turnover [44]. At Cleveland Clinic in the United States, managerial training on rest and mental health improved morale and retention [45,46]. Wellness programs like Johns Hopkins' Resilience in Stressful Events provide peer support and counseling, significantly reducing burnout and emotional exhaustion among nurses [47].

Conclusion

This study underscores the role of adequate rest in preventing burnout and enhancing job satisfaction among nurses in medical and emergency units. The findings reveal that systemic barriers, such as high workloads, inadequate staffing, and chaotic work environments, significantly hinder nurses' ability to achieve sufficient rest. Nurses in the emergency unit, in particular, face higher stress levels due to the demanding nature of their work, which worsens fatigue and burnout. The lack of localized studies focusing on the physical and mental well-being of Jamaican nurses, particularly in high-pressure emergency units, creates a significant academic gap.

Addressing these issues can enhance the global understanding of burnout in various contexts while providing region-specific strategies to improve nurse retention, job satisfaction, and overall health care quality in Jamaica. A multifaceted approach is required to tackle these challenges, incorporating policy changes, management training, and the creation of supportive work environments. By prioritizing the well-being of nurses through adequate rest, health care institutions can not only improve nurse satisfaction but also ensure better patient care outcomes.

Acknowledgments

The authors are grateful to the director of nursing services at the hospital where the study was conducted for granting permission to carry out the research. We give a special thank you to all the participants for taking part in the study. We are thankful to Imperial College London for their continued academic support and for covering the article processing fee for the publication of this manuscript. The authors used the generative artificial intelligence tool ChatGPT (OpenAI) for editorial assistance in formatting and aligning the manuscript with JMIR submission guidelines. All content was reviewed and finalized by the authors.

Funding

The researchers did not receive funding from any organization to undertake this research. Imperial College London did cover the costs of open access since the study was undertaken as a master's student project. While affiliated with the study, the institution did not influence the study design, data collection, analysis, interpretation of results, or the drafting of the manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: CS (lead)
Data curation: CS
Formal analysis: CS
Funding acquisition: NC, KKM
Investigation: CS
Methodology: CS
Project administration: CS
Supervision: NC, KKM

XSL•FO
RenderX

Validation: CS
Writing – original draft: CS (lead)
Writing – review & editing: NC (supporting), KKM (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File, 12 KB - nursing_v9i1e84106_app1.docx]

Multimedia Appendix 2
Coding table: from quotes to theme.
[DOCX File, 12 KB - nursing_v9i1e84106_app2.docx]

References

1. Flaubert JL, Le Menestrel S, Williams DR, Wakefield MK, editors. The Future of Nursing 2020-2030: Charting a Path to Achieve Health Equity: National Academies Press; 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK573922/ [accessed 2024-09-10]
2. The nurses changing lives. BBC. 2020. URL: https://www.bbc.com/storyworks/caring-with-courage/the-nurses-changing-lives [accessed 2024-09-10]
3. Jennings BM. Work stress and burnout among nurses: role of the work environment and working conditions. In: Hughes RG, editor. Patient Safety and Quality: An Evidence-Based Handbook for Nurses: National Library of Medicine; 2008. URL: https://www.ncbi.nlm.nih.gov/books/NBK2668/ [accessed 2024-09-10]
4. Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. JAMA 2002;288(16):1987-1993. [doi: 10.1001/jama.288.16.1987] [Medline: 12387650]
5. Burn-out an "occupational phenomenon": International Classification of Diseases. World Health Organization. 2019. URL: https://www.who.int/news/item/28-05-2019-burn-out-an-occupational-phenomenon-international-classification-of-diseases [accessed 2024-09-10]
6. Penn M, Romano JL, Foat D. The relationship between job satisfaction and burnout: a study of human service professionals. Adm Ment Health 1988;15(3):157-165. [doi: 10.1007/BF00869249]
7. Bakhamis L, Paul DP, Smith H, Coustasse A. Still an epidemic: the burnout syndrome in hospital registered nurses. Health Care Manag (Frederick) 2019;38(1):3-10. [doi: 10.1097/HCM.0000000000000243] [Medline: 30640239]
8. Gebregziabher D, Berhanie E, Berihu H, Belstie A, Teklay G. The relationship between job satisfaction and turnover intention among nurses in Axum comprehensive and specialized hospital Tigray, Ethiopia. BMC Nurs 2020;19(1):79. [doi: 10.1186/s12912-020-00468-0] [Medline: 32831645]
9. Locke EA. The nature and causes of job satisfaction. In: Dunnette MD, editor. Handbook of Industrial and Organizational Psychology: Rand McNally; 1976, Vol. 02:1297-1343.
10. Scott AJ, Webb TL, Martyn-St James M, Rowse G, Weich S. Improving sleep quality leads to better mental health: a meta-analysis of randomised controlled trials. Sleep Med Rev 2021 Dec;60:101556. [doi: 10.1016/j.smrv.2021.101556] [Medline: 34607184]
11. Vandekerckhove M, Wang YL. Emotion, emotion regulation and sleep: an intimate relationship. AIMS Neurosci 2018;5(1):1-17. [doi: 10.3934/Neuroscience.2018.1.1] [Medline: 32341948]
12. Jamaica. World Bank Group. URL: https://www.worldbank.org/en/country/jamaica [accessed 2024-09-10]
13. Nurses per capita by country. World Population Review. 2025. URL: https://worldpopulationreview.com/country-rankings/nurses-per-capita-by-country [accessed 2024-09-10]
14. Muir KJ, Keim-Malpass J, LeBaron VT. “You have to ask yourself when you’ve had enough”: an ethnography of multi-level nurse burnout cultural impacts in the emergency department. SSM - Qualitative Research in Health 2022 Dec;2:100111. [doi: 10.1016/j.ssmqr.2022.100111]
15. Burnout: emergency medicine hit hardest. HealthLeaders. 2015. URL: https://www.healthleadersmedia.com/strategy/burnout-emergency-medicine-hit-hardest#:~:text=Rates%20of%20burnout%20vary%20markedly [accessed 2024-09-18]
16. About WRHA. Western Regional Health Authority. URL: https://www.wrha.gov.jm/about-us/ [accessed 2024-09-10]
17. Tenny S, Brannan JM, Brannan GD. Qualitative study. In: StatPearls: StatPearls Publishing; 2022. URL: https://www.ncbi.nlm.nih.gov/books/NBK470395 [accessed 2024-09-10]
18. Braun V, Clarke V. Thematic Analysis: A Practical Guide: Sage; 2022. URL: https://us.sagepub.com/en-us/nam/thematic-analysis/book248481 [accessed 2026-01-03]
19. World Medical Association. World medical association declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]

XSL•FO
RenderX

20. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. Educ Inf 2004;22(2):63-75. [doi: 10.3233/EFI-2004-22201]
21. Melnyk BM. Burnout, depression and suicide in nurses/clinicians and learners: an urgent call for action to enhance professional well‐being and healthcare safety. Worldviews Evid Based Nurs 2020 Feb;17(1):2-5. [doi: 10.1111/wvn.12416] [Medline: 32017437]
22. Khatatbeh H, Pakai A, Al-Dwaikat T, et al. Nurses' burnout and quality of life: a systematic review and critical analysis of measures used. Nurs Open 2022 May;9(3):1564-1574. [doi: 10.1002/nop2.936] [Medline: 33991408]
23. Alenezi L, Gillespie GL, Smith C, Davis KG. Gender differences in burnout among US nurse leaders during COVID-19 pandemic: an online cross-sectional survey study. BMJ Open 2024 Nov 19;14(11):e089885. [doi: 10.1136/bmjopen-2024-089885] [Medline: 39566937]
24. Ten Hoeve Y, Kunnen S, Brouwer J, Roodbol PF. The voice of nurses: Novice nurses' first experiences in a clinical setting. a longitudinal diary study. J Clin Nurs 2018 Apr;27(7-8):e1612-e1626. [doi: 10.1111/jocn.14307] [Medline: 29446496]
25. Kumar A, Sinha A, Varma JR, Prabhakaran AM, Phatak AG, Nimbalkar SM. Burnout and its correlates among nursing staff of intensive care units at a tertiary care center. J Family Med Prim Care 2021;10(1):443-448. [doi: 10.4103/jfmpc.jfmpc_1651_20]
26. Jakovljevic M, Liu Y, Cerda A, et al. The Global South political economy of health financing and spending landscape - history and presence. J Med Econ 2021 Nov;24(sup1):25-33. [doi: 10.1080/13696998.2021.2007691] [Medline: 34866543]
27. Ashipala DO, Nghole TM. Factors contributing to burnout among nurses at a district hospital in Namibia: a qualitative perspective of nurses. J Nurs Manag 2022 Oct;30(7):2982-2991. [doi: 10.1111/jonm.13693] [Medline: 35593041]
28. Miller JF. Burnout and its impact on good work in nursing. J Radiol Nurs 2011;30(4):146-149. [doi: 10.1016/j.jradnu.2011.06.004]
29. Study finds small developing lands hit hardest by 'brain drain'. Migration and Development, Archive.org. 2005. URL: https://web.archive.org/web/20110723210301/http://www.migrationanddevelopment.net/popular-press-news/developing-lands-hit-hardest-by-brain-drain [accessed 2024-09-22]
30. Brain drain countries. World Population Review. 2025. URL: https://worldpopulationreview.com/country-rankings/brain-drain-countries [accessed 2024-09-10]
31. Health workers perception and migration in the caribbean region. PAHO. 2020. URL: https://www.paho.org/en/documents/health-workers-perception-and-migration-caribbean-region [accessed 2024-09-10]
32. Bourne PA. Factors determining "Burnout" among nursing - staffers at a national hospital in Jamaica. JOJPH 2018;3(2):555608 [FREE Full text] [doi: 10.19080/JOJPH.2018.03.555608]
33. Fixing health: the brain drain of our health professionals. Jamaica Gleaner. 2015. URL: http://jamaica-gleaner.com/article/health/20150930/fixing-health-brain-drain-our-health-professionals [accessed 2024-09-10]
34. Maharaj S, Lees T, Lal S. Prevalence and risk factors of depression, anxiety, and stress in a cohort of Australian nurses. Int J Environ Res Public Health 2018 Dec 27;16(1):61. [doi: 10.3390/ijerph16010061] [Medline: 30591627]
35. Mincarone P, Bodini A, Tumolo MR, et al. Association between physical activity and the risk of burnout in health care workers: systematic review. JMIR Public Health Surveill 2024 Mar 18;10(1):e49772. [doi: 10.2196/49772] [Medline: 38498040]
36. International Nurses Day 2025: caring for nurses strengthens economies. : International Council of Nurses (ICN); 2025 URL: https://www.icn.ch/sites/default/files/2025-04/ICN_IND2025_report_EN_A4_FINAL_0.pdf [accessed 2025-10-30]
37. Brigham T, Barden C, Legreid Dopp A, et al. A journey to construct an all-encompassing conceptual model of factors affecting clinician well-being and resilience. NAM Perspectives 2018;8(1). [doi: 10.31478/201801b]
38. State of the world's nursing 2025: investing in education, jobs, leadership and service delivery. : World Health Organisation (WHO) & International Council of Nurses (ICN); 2025 URL: https://iris.who.int/server/api/core/bitstreams/a4173924-a18f-49b6-8bd1-9c2a4a098980/content [accessed 2025-10-30]
39. Bragadóttir H, Kalisch BJ, Flygenring BG, Tryggvadóttir GB. The relationship of nursing teamwork and job satisfaction in hospitals. SAGE Open Nurs 2023;9(9):23779608231175027. [doi: 10.1177/23779608231175027] [Medline: 37214231]
40. Martin B, Kaminski-Ozturk N, O'Hara C, Smiley R. Examining the impact of the COVID-19 pandemic on burnout and stress among U.S. nurses. J Nurs Regul 2023 Apr;14(1):4-12. [doi: 10.1016/S2155-8256(23)00063-7] [Medline: 37035777]
41. Griffiths P, Ball J, Murrells T, Jones S, Rafferty AM. Registered nurse, healthcare support worker, medical staffing levels and mortality in English hospital trusts: a cross-sectional study. BMJ Open 2016 Feb 9;6(2):e008751. [doi: 10.1136/bmjopen-2015-008751] [Medline: 26861934]
42. Twigg D, Duffield C, Bremner A, Rapley P, Finn J. The impact of the nursing hours per patient day (NHPPD) staffing method on patient outcomes: a retrospective analysis of patient and staffing data. Int J Nurs Stud 2011 May;48(5):540-548. [doi: 10.1016/j.ijnurstu.2010.07.013] [Medline: 20696429]
43. West E, Mays N, Rafferty AM, Rowan K, Sanderson C. Nursing resources and patient outcomes in intensive care: a systematic review of the literature. Int J Nurs Stud 2009 Jul;46(7):993-1011. [doi: 10.1016/j.ijnurstu.2007.07.011] [Medline: 17964577]
44. Duffield CM, Roche MA, Blay N, Stasa H. Nursing unit managers, staff retention and the work environment. J Clin Nurs 2011 Jan;20(1-2):23-33. [doi: 10.1111/j.1365-2702.2010.03478.x] [Medline: 21158986]

XSL•FO
RenderX

45. Shirey MR. Authentic leadership, organizational culture, and healthy work environments. Crit Care Nurs Q 2009;32(3):189-198. [doi: 10.1097/CNQ.0b013e3181ab91db] [Medline: 19542970]
46. Williams SG, Fruh S, Barinas JL, Graves RJ. Self-care in nurses. J Radiol Nurs 2022 Mar;41(1):22-27. [doi: 10.1016/j.jradnu.2021.11.001] [Medline: 35431686]
47. Edrees H, Connors C, Paine L, Norvell M, Taylor H, Wu AW. Implementing the RISE second victim support programme at the Johns Hopkins Hospital: a case study. BMJ Open 2016 Sep 30;6(9):e011708. [doi: 10.1136/bmjopen-2016-011708] [Medline: 27694486]

Abbreviations

A&E: Accident and Emergency
RN: registered nurse

Edited by E Borycki; submitted 14.Sep.2025; peer-reviewed by K Kruger, T Stephenson; revised version received 19.Dec.2025; accepted 23.Dec.2025; published 23.Jan.2026.

Please cite as:
Smith C, Murray KK, Croome N
The Effects of Adequate Rest on Nurse Job Satisfaction, Burnout Prevention, and Physical Health in Medical and Emergency Units at a Hospital in Western Jamaica: Qualitative Study
JMIR Nursing 2026;9:e84106
URL: https://nursing.jmir.org/2026/1/e84106
doi:10.2196/84106

© Channon Smith, Keitumetse-Kabelo Murray, Natasha Croome. Originally published in JMIR Nursing (https://nursing.jmir.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Medication-Related Errors Among Nurses by Unit Adaptation Levels: Bayesian Network–Based Exploratory Study

Naotaka Sugimura[1,2*], RN, MSc; Katsuhiko Ogasawara[3,4*], MBA, PhD

[1]Department of Nursing, School of Health Sciences, Tokyo University of Technology, Ota-ku, Tokyo, Japan

[2]Graduate School of Health Sciences, Hokkaido University, Sapporo, Hokkaido, Japan

[3]Faculty of Health Sciences, Hokkaido University, N12-W5, Kita-ku, Sapporo, Hokkaido, Japan

[4]Graduate School of Engineering, Muroran Institute of Technology, Muroran, Hokkaido, Japan

[*]all authors contributed equally

Corresponding Author:
Katsuhiko Ogasawara, MBA, PhD
Faculty of Health Sciences, Hokkaido University, N12-W5, Kita-ku, Sapporo, Hokkaido, Japan

Abstract

Background: Medical errors occur more frequently in health care than in other industries due to challenges in patient safety education for nurses and students. To address this, it is important to identify the factors and structures underlying clinical errors and apply these insights to educational programs. Medication-related errors are highly preventable with appropriate interventions, highlighting the importance of data-driven safety education. Previous research suggests that unit adaptation, rather than clinical experience alone, plays a critical role in error occurrence. Focusing on "adaptive performance," an underexplored concept in nursing, can help identify new educational strategies and interventions.

Objective: This hypothesis-generating study used Bayesian network modeling to examine how unit experience—and, secondarily, total nursing experience—relates to the structure of medication-related errors, with implications for data-driven patient safety education and nursing student instruction.

Methods: This mixed methods study conducted a qualitative root cause analysis of medication error reports to identify causal factors. Bayesian network modeling, an artificial intelligence–based approach, was used to visualize error-generation flows and compare models based on years of experience within the current unit. Data were obtained from 2023 medication-related error reports submitted by nurses to the Japan Council for Quality Health Care.

Results: Among 119 medication-related incidents, the largest proportion occurred in internal medicine departments (n=47 cases, 39.5%), followed by surgical departments (n=27 cases, 22.7%), the intensive care unit (n=10 cases, 8.4%), and the emergency room (n=7 cases, 5.9%). Most incidents occurred in patient rooms or wards (n=97 cases, 81.5%). Root cause analysis identified 10 types of medication-related events and 23 contributing factors. Cases were categorized into low-, moderate-, and high-adaptation groups based on years of experience in the current unit. In the low-adaptation group (n=33), 69.7% of nurses had ≥5 years of total nursing experience, indicating that the group did not consist solely of novice nurses. The average entropy of incident events across the Bayesian network models ranged from 0.37 to 0.78, suggesting moderate to relatively high uncertainty in the inferred error structures. The moderate- and high-adaptation models exhibited fewer complex error networks, with weaker chains of unsafe conditions or actions than the low-adaptation model. The extensive clinical experience did not always prevent errors; rather, it was often linked to lapses in verification behavior.

Conclusions: Unit adaptation appears to play a critical role in shaping error pathways. Nurses with lower unit adaptation demonstrated more complex error structures, indicating higher vulnerability to incident chains. In contrast, total years of nursing experience did not uniformly reduce risk and may even relate to lapses in verification behaviors.

(JMIR Nursing 2026;9:e87436) doi:10.2196/87436

KEYWORDS

adaptive performance; artificial intelligence; Bayesian network modeling; data-driven education; patient safety

Introduction

Data-driven education is an approach that continuously enhances the quality of instruction and learning by analyzing data generated throughout the educational process [1,2]. In recent years, data-driven education has been increasingly applied to evaluate instructional quality by assessing learning outcomes [3]. However, in nursing education, it is particularly important to utilize real-world data not only for analyzing learner performance but also for informing the design and improvement

XSL•FO
RenderX

of educational systems. Patient safety education is a representative example as it remains an underdeveloped area of curriculum design worldwide, with evident challenges in the early stages of program development. Despite the World Health Organization (WHO)'s concern over the high frequency of medical errors in health care compared with other industries [4], patient safety education remains marginal in nursing curricula, often addressed only informally or as part of broader topics, and research on educational interventions is limited [5,6].

At the core of this issue lies the fundamental nature of health care professionals as practitioners whose work directly influences clinical practice, aligned with indispensable clinical competencies. In response, the American Association of Colleges of Nursing has proposed a framework for competency-based education [7]. Grounded in the principle of "beginning with the end in mind" [8,9], competency-based education is essential for defining learning outcomes and designing effective backward curriculum in patient safety education [10-12]. Nurses are positioned at the frontline of patient care and bear critical responsibility as the final safeguard against medical errors and adverse events. However, in the absence of a well-established and systematic educational framework, they often depend heavily on on-the-job training. This imbalance in educational approaches may ultimately undermine efforts to ensure patient safety.

Among medical errors, medication-related errors are the most preventable and represent an area in which educational interventions can yield substantial improvement [13,14]. A previous study employing Bayesian network modeling (BNM) identified multiple knowledge-, behavior-, and environment-related factors that contribute to medication errors among nurses [15]. These findings indicate that many of these factors can be modified through education, suggesting the potential for significant error reduction. In particular, unfamiliarity with unit-specific tasks and insufficient adaptation to work routines have emerged as key contributors, underscoring the importance of unit adaptation.

Adaptive performance—the ability to respond flexibly to unexpected events and environmental changes [16]—plays a critical role in error occurrence. Medication errors in nursing can be understood as outcomes of performance shaped by fatigue, workload, and environmental conditions [17-19]. Two attribute variables have commonly been used to represent a nurse's experience: total years of nursing experience and years of unit experience. The former is well known from the Benner classification of nursing expertise [20], which applied the 5-stage model and has informed criteria in some clinical ladders [21]. By contrast, unit adaptation has received limited attention in nursing and medical research. Consequently, much of the patient safety literature has focused on novice nurses with limited experience [22-25]. However, understanding the developmental stages and individual differences of adaptive performance is essential for designing effective educational strategies for both nurses and nursing students [26]. Focusing on unit adaptation [16] challenges the view that "inexperience" is defined solely by fewer years of clinical practice and enables a more precise identification of targets for educational interventions.

Accordingly, this hypothesis-generating study used BNM to examine how unit experience—and, secondarily, total nursing experience—relates to the structure of medication-related errors, with implications for data-driven patient safety education and nursing student instruction.

Methods

Study Design

This study used a mixed methods approach consisting of 2 main analytical processes: qualitative factor extraction and relational analysis. First, episodes of medication-related errors involving nurses were qualitatively analyzed using root cause analysis (RCA) to identify the contributing factors [15]. Next, the data were categorized into 3 groups based on years of unit experience: low adaptation (>1 y), moderate adaptation (1 to >3 y), and high adaptation (3 years or more). For each group, BNM was conducted to evaluate the relationships between the factors and compare the structural differences in the flow of medication-related errors.

Data Sources and Data Selection

This study used data from the Medical Error Database published by the Japan Council for Quality Health Care (JCQHC) [27]. The JCQHC undertakes various initiatives, including the collection of medical errors and near-miss cases, to maintain public trust in health care services and improve their quality. The database includes reports from a total of 1899 medical institutions across Japan. The JCQHC requires participating medical institutions to report incidents that fall into the following categories (1–3). (1) Cases in which inappropriate medical care or management was clearly identified and resulted in patient death, permanent physical or psychological impairment, or the need for unexpected or more extensive treatment than anticipated. (2) Cases in which it could not be conclusively determined whether inappropriate medical care or management occurred, but which were otherwise comparable to category (1). (3) Cases other than (1) and (2) that provide useful information for preventing and reducing the recurrence of incidents within health care institutions. The reporting form of the JCQHC requires detailed documentation, including the patient's condition and age group, the working status of the staff involved in the error, the circumstances leading to the incident, and measures for future prevention. From this database, 119 cases were selected for analysis based on the following criteria (): (1) a nurse was involved, (2) the error was medication-related, involving adverse events or equivalent medical errors, and (3) the case included sufficient contextual information to allow for RCA.

Competency and Total Clinical Experience

Although there is no universally accepted definition of nursing competency, it is commonly described as "a core ability required to fulfill nursing responsibilities" [28] or as "the knowledge, behaviors, judgment, skills, values, and attitudes necessary to provide safe, effective, and high-quality care" [29]. These definitions are grounded in experiential learning theory [20] and suggest that competency is a transferable ability developed through the integration of knowledge and skills. In many studies,

XSL•FO
RenderX

total years of clinical experience and length of employment are frequently treated as variables that influence competency development [20,30-32]. In this study, informed by the Benner 5-stage model [20] and its application within clinical ladder systems in the world and Japan [21,33], total years of nursing experience were treated as categorical data (<5 y, 5–10 y, and ≥10 y).

Adaptive Performance and Current Unit Experience

In contrast to competency, performance, particularly work performance, is defined as "behaviors or actions that are relevant to the goals of the organization" [34] and as "scalable actions, behaviors, and outcomes" [35]. Work performance is categorized into 4 domains: task, contextual, counterproductive, and adaptive [36]. Fatigue and burnout are closely associated with work performance among nurses [17], which is a critical element influencing patient safety and holds substantial importance in clinical practice.

Adaptive performance refers to "the extent to which an individual adapts to changes in a work system or work roles" [16]. Although this concept has received limited attention in medical and nursing research, it is theoretically associated with experiential learning [21]. In this study, years of current unit experience were treated as categorical data (low adaptation: <1 y; moderate adaptation: 1 to <3 y; high adaptation: ≥3 y) as a pragmatic proxy for the degree of adaptation to a specific clinical environment, informed by the application of clinical ladder systems in Japan and internationally [21,33]. These 3 groups were intended to correspond to levels of clinical practice and functional independence in the workplace and were operationally defined as follows:

- Low adaptation: the stage of learning all tasks with instructions and follow-ups
- Moderate adaptation: the stage of learning some tasks with instructions and follow-ups while performing other tasks independently
- High adaptation: the stage of performing most tasks independently

Analysis

RCA and Factor Extraction

A total of 119 medical error reports were reviewed to classify events and extract contributing factors. Although the primary coding process was conducted by one researcher, the overall analytical framework was developed through discussions between 2 authors. To minimize subjectivity in factor extraction, a fact-based analytical strategy was adopted. Specifically, contributing factors were strictly derived from observable descriptions in the incident reports rather than from interpretative judgments. The following RCA-guiding questions were applied to ensure systematic extraction: "Why did this event occur?," "What actions or conditions involving the nurse or patient contributed to the event?" "If the event could have been prevented, what actions should the nurse have taken?" In addition, predefined factor definitions were established in advance (Table 1), and contributing factors were extracted as objective facts from the free-text descriptions regarding the event sequence and implemented countermeasures based on consistent criteria. Although some reports included subjective statements—such as the reporter's psychological state or fatigue—these were excluded from the analysis because such descriptions varied among reports and could not ensure objective evaluation. The extracted factors were categorized into 4 groups: attributive, system-related, situational, and knowledge/behavioral. Each factor was labeled within its respective categories. Event classification and factor extraction based on the RCA were reported in our previous study [15], and their validity was confirmed through consistency with prior research [37].

XSL•FO
RenderX

Table . Explanations of nodes.

Node	Explanation
Event	
Wrong dose	Administered the wrong dose of medication
Wrong patient	Administered medicine to wrong patient
Improper mixing	Performed improper mixing
Wrong or unnecessary administration	Administered a different medication to the one that should have been administered or administered a medication unnecessarily
Overlooking contraindication	Missed contraindicated medication
Accidental ingestion/ overdose	Patient accidentally ingested or overdosed
Incorrect storage	Stored medicine in an incorrect storage method
Wrong timing	Administered medicine at the wrong time
Default of medication discontinue	Failed to discontinue medicine
Wrong route	Administered medicine by the wrong route
Attributional factors	Factors related to the attributes of nurses who made errors
Year as registered nurse	Divided into 5 years as registered nurse
<5	Including novice, advanced beginner, and competent nurse
5 - 9	Including proficient nurse
≥10	Including expert nurse
Year as current unit	Divided into 3 stages of learning the tasks of a typical Japanese hospital unit
<1	The stage of learning all tasks with instructions and follow-ups
1 - 3	The stage of learning some tasks with instructions and follow-ups, but performing some tasks independently
>3	The stage of performing most tasks independently
System factor	Factors related to the organization's systems, such as electronic medical records and educational systems
Failure to detect by the electronic system	Not equipped with the ability to detect on electronic medical record system
Conditional factors	Factors related to conditions surrounding nurses. For example, the time of day, the type of work, and other conditions when errors had occurred
Weekday/holiday	Occurred during weekday or holiday
Work time	Work time was divided into 3 parts according to the patient's life and the nurse's work behavior
9 AM-4 PM	
4 PM-9 PM	
9 PM-9 AM	
Day shift/night shift	Occurred during day shift or night shift
Patients' age (y)	Age of patients associated with errors
<20	
20 - 59	
≥60	
Time pressure	Time pressure was present due to sudden emergencies or multiple tasks of the patient
Sudden change in patients' schedule	There was a sudden change in schedule for examination or treatment
Failure of doctor's orders	There was a failure of doctor's orders.
Patients who do not follow instructions	Patient did not follow the instructions of health care professionals

XSL•FO
RenderX

Node	Explanation
Knowledge and behavioral factors	Factors related to nurses' knowledge state and behavior.
Failure to confirm the 5 rights	Had not checked the 5 rights
Assumptions and forgetfulness	Assumed or forgot instructions and orders
Invalid double check	Nurses double-checking each other was invalid
Improper use of instruments and equipment	The use of instruments and equipment induced errors
Insufficient knowledge of medications	Did not have sufficient knowledge
Unfamiliarity with operations of medications	Did not know how to properly handle medications due to unfamiliarity
Failure of information communication	Discrepancies in communication when conveying information
Improper handling of medications	Handled medicine in the wrong way
Failure of monitoring	Monitoring was a possibly avoidable error
Task interruption	Task interrupted when error occurs
Non-reconfirming of inappropriate doctor's orders	Did not check again even though nurse felt the doctor's instructions were inappropriate
Lack of explanation to patients	Not enough explanation about medication to the patient
Miscalculation	Miscalculated appropriate medication dosage

Bayesian Network Modeling

BNM is an artificial intelligence (AI)–based probabilistic approach that represents qualitative dependencies among multiple variables through a graphical structure and expresses quantitative relationships using conditional probabilities [15]. In this study, each event and contributing factor was represented as a "node," and causal or temporal relationships were indicated by "edges" connecting these nodes. The analysis was performed using BAYOLINK version 7.0.1 (NTT DATA Mathematical Systems), employing a greedy algorithm with an Akaike information criterion threshold of 0.02 [38]. We used this algorithm and criterion to avoid overlooking even weak associations between incidents and contributing factors, as well as relationships among the factors themselves, because this study was positioned as a hypothesis-generating investigation. While previous patient safety research has primarily relied on qualitative analyses based on individual incident case reviews or simple descriptive surveys, this study aims to explore hypotheses regarding common structural patterns underlying specific incident occurrences and the characteristics of contributing factors derived directly from the data.

Model construction combined algorithmic computation with expert judgment [15]. A major advantage of this method is its ability to infer potential causal relationships from relatively small datasets based on information criteria. During the process of model development, whitelist and blacklist constraints were applied to refine the network structure. No whitelist was specified, and the initial blacklist included attributive and system-related factors, which were conceptually treated as root-level nodes without parent relationships. This classification was informed by human factors–based frameworks [37,39]. These constraints were iteratively updated to remove irrelevant links and maintain logical consistency in the causal and temporal relationships. Final models were constructed for each adaptation group.

Unit experience was classified into 3 adaptation levels based on the general clinical ladder for nurses in Japan: low adaptation (<1 y), moderate adaptation (≥1 and <3 y), and high adaptation (≥3 y) [15]. The low-adaptation group was defined as the stage in which nurses performed all tasks under instruction and supervision. The moderate-adaptation group was defined as the stage in which nurses performed some tasks independently, while still requiring instruction and follow-up for others. The high-adaptation group was defined as the stage in which nurses performed most tasks independently. Consistent with the secondary objective of this study, when the node "Year as registered nurse" (categorical variable: <5 y, 5–9 y, and ≥10 y) was included in a model, the category associated with medical error occurrence was identified. The structural differences between the 3 models were evaluated based on the number of layers and the number and characteristics of factors not directly associated with adverse events or equivalent medical errors.

The average entropy was calculated as an indicator of uncertainty in BNM. When the child node is a binary variable, the average entropy ranges from 0 to 1.0, where lower values indicate lower uncertainty. For example, when the cross-tabulation between a parent node A (binary) and a child node B (binary) is completely uniform, the uncertainty reaches its maximum and the value becomes 1.0. A smaller value of the average entropy indicates that the state of the child node can be more easily determined, given the condition of the parent node. Although there is no universally established threshold for interpreting the magnitude of average entropy, in this study, the average entropy was calculated only for binary nodes, and therefore, the values were restricted to the range of 0 to 1.0. Accordingly, the model was evaluated comparatively within this range.

Ethical Considerations

This study used data from the anonymized and publicly available database of the JCQHC [26]. Therefore, it did not violate any

XSL•FO
RenderX

privacy policies and did not fall under the scope of an ethics review.

Results

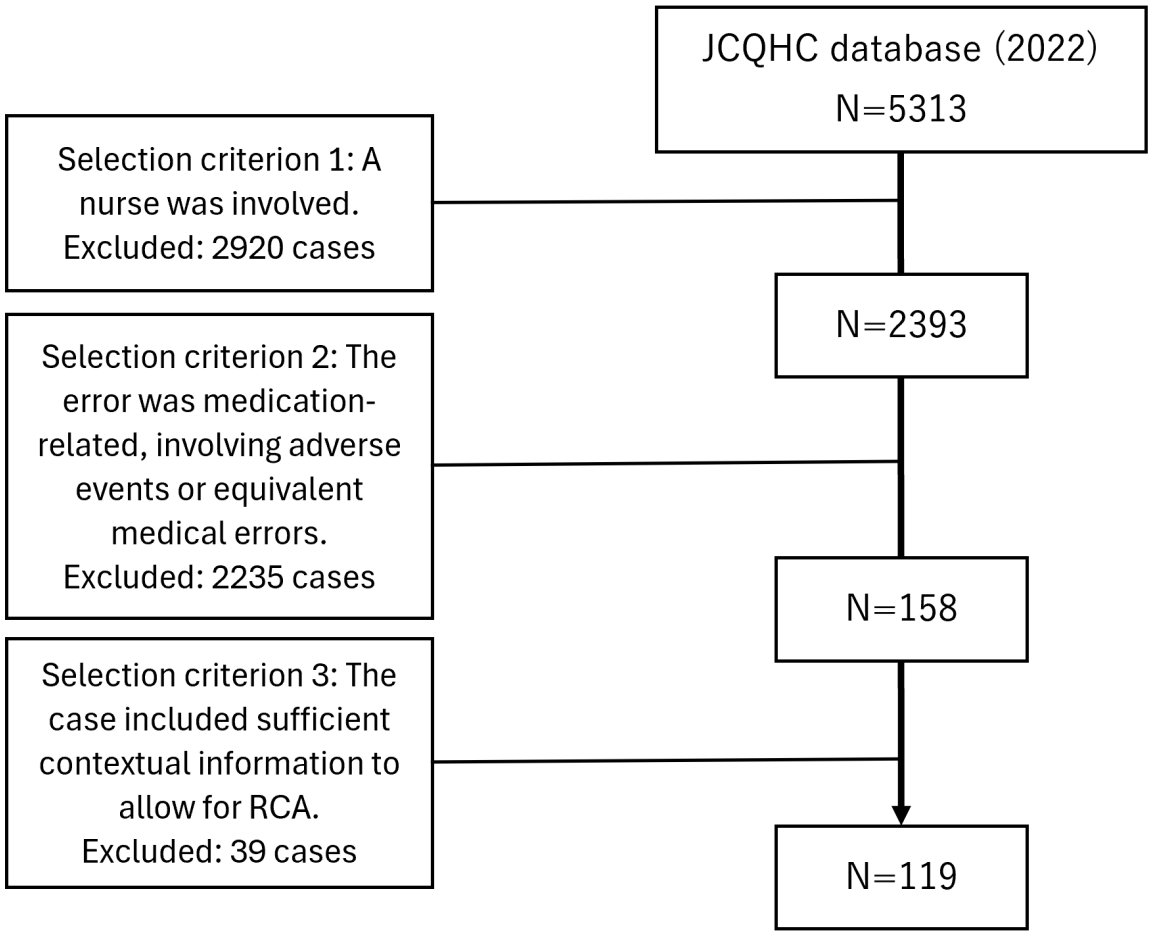
Descriptive Statistics

The distribution of the 119 incidents by clinical department was as follows. A total of 27 (22.7%) incidents occurred in surgical departments: neurosurgery (n=7 cases, 5.9%), orthopedic surgery (n=5 cases, 4.2%), gastrointestinal surgery (n=3 cases, 2.5%), cardiovascular surgery (n=1 case, 0.8%), plastic surgery (n=1 case, 0.8%), and oral and maxillofacial surgery (n=1 case, 0.8%). An additional 11 (9.2%) incidents were categorized as "other surgical departments," referring to cases documented only as "surgery" or from mixed surgical wards (Figure 1).

A total of 47 (39.5%) incidents occurred in internal medicine departments: cardiology (n=7, 5.9%), hematology (n=6, 5.0%), gastroenterology (n=5, 4.2%), pulmonology (n=3, 2.5%), otorhinolaryngology (n=3, 2.5%), metabolism/endocrinology (n=3, 2.5%), neurology (n=2, 1.7%), dermatology (n=1, 0.8%), rheumatology (n=1, 0.8%), oncology (n=1, 0.8%), and general internal medicine (n=1, 0.8%). Another 14 (11.8%) incidents were categorized as "other internal medicine departments," referring to cases recorded only as "internal medicine" or from mixed internal medicine wards.

In addition, 10 (8.4%) incidents occurred in the intensive care unit, 7 (5.9%) in the emergency room, and 4 (3.4%) in the neonatal intensive care unit. Other clinical areas included obstetrics and gynecology (11 incidents), pediatrics (9 incidents), psychiatry (1 incident), and palliative care (1 incident).

Figure 1. Data selection process. JCQHC: Japan Council for Quality Health Care; RCA: root cause analysis.



The distribution of incident locations was as follows. Patient rooms and wards accounted for the majority of cases (n=97 incidents, 81.5%), followed by outpatient clinics (n=12, 10.1%). Other locations were operating rooms (n=5, 4.2%), radiology examination rooms (n=3, 2.5%), dialysis units (n=1, 0.8%), cardiac catheterization laboratories (n=1, 0.8%), and patients' homes (n=1, 0.8%).

Medication-related errors were classified into 10 event types, and 23 contributing factors were extracted (Table 1). Descriptive statistics by the unit experience group are summarized in Table 2. In the adaptation group (n=33), defined as nurses with less than 1 year of current unit experience, 8 (24.2%) had 5 to 9 years' experience as registered nurses and 15 (45.5%) had ≥10 years. These findings indicated that this group was not composed solely of nurses with limited clinical experience.

XSL•FO
RenderX

Table . Summary statistics of nodes categorized by unit adaptation.

Node	All (N=119), n (%)	Low-adaptation model (n=33), n (%)	Moderate-adaptation model (n=36), n (%)	High-adaptation model (n=50), n (%)
Event[a]				
Wrong dose	43 (36.1)	15 (46)	10 (28)	18 (36)
Wrong patient	22 (18.5)	4 (12)	8 (22)	10 (20)
Improper mixing	15 (12.6)	5 (15)	6 (17)	4 (8)
Wrong or unnecessary administration	13 (10.9)	4 (12)	4 (11)	5 (10)
Overlooking contraindication	11 (9.2)	2 (6)	5 (14)	4 (8)
Accidental ingestion/overdose	7 (5.9)	1 (3)	3 (8)	3 (6)
Incorrect storage	6 (5.0)	1 (3)	1 (3)	4 (8)
Wrong timing	4 (3.4)	2 (6)	0	2 (4)
Default of medication discontinue	4 (3.4)	2 (6)	1 (3)	1 (2)
Wrong route	2 (1.7)	0	0	2 (4)
Attributional factors				
Years as registered nurse				
<5	39 (32.8)	10 (30)	18 (50)	11 (22)
5 - 9	28 (23.5)	8 (24)	7 (19)	13 (26)
≥10	52 (43.7)	15 (46)	11 (31)	26 (52)
System factor				
Failure to detect by the electronic system	7 (5.9)	2 (6)	2 (6)	3 (6)
Conditional factors				
Weekday/holiday (during holiday)	24 (20.2)	6 (18)	9 (25)	9 (18)
Work time				
9 AM-4 PM	54 (45.4)	18 (55)	13 (36)	23 (46)
4 PM-9 PM	31 (26.1)	8 (24)	11 (31)	12 (24)
9 PM-9 AM	34 (28.6)	7 (21)	12 (33)	15 (30)
Day shift/night shift (during night shift)	65 (54.6)	15 (46)	23 (64)	27 (54)
Patients' age (y)				
<20	12 (10.1)	4 (12)	2 (6)	6 (12)
20 - 59	34 (28.6)	10 (30)	14 (39)	10 (20)
≥60	73 (61.3)	19 (58)	20 (56)	34 (68)
Time pressure	42 (35.3)	13 (39)	12 (33)	17 (34)
Sudden change in patients' schedule	10 (8.4)	3 (9)	1 (3)	6 (12)
Failure of doctor's orders	10 (8.4)	1 (3)	2 (6)	7 (14)
Patients who do not follow instructions	13 (10.9)	4 (12)	6 (17)	3 (6)
Knowledge and behavioral factors				
Failure to confirm the 5 rights	59 (49.6)	14 (42)	20 (56)	25 (50)

XSL•FO
RenderX

Node	All (N=119), n (%)	Low-adaptation model (n=33), n (%)	Moderate-adaptation model (n=36), n (%)	High-adaptation model (n=50), n (%)
Assumptions and forgetfulness	47 (39.5)	12 (36)	14 (39)	21 (42)
Invalid double check	40 (33.6)	14 (42)	11 (31)	15 (30)
Improper use of instruments and equipment	35 (29.4)	6 (18)	13 (36)	16 (32)
Insufficient knowledge of medications	27 (22.7)	11 (33)	6 (17)	10 (20)
Unfamiliarity with operations of medications	23 (19.3)	10 (30)	7 (19)	6 (12)
Failure of information communication	23 (19.3)	7 (21)	10 (28)	6 (12)
Improper handling of medications	13 (10.9)	5 (15)	4 (11)	4 (8)
Failure of monitoring	13 (10.9)	3 (9)	5 (14)	5 (10)
Task interruption	11 (9.2)	3 (9)	0	8 (16)
Non-reconfirming of inappropriate doctor's orders	10 (8.4)	2 (6)	2 (6)	6 (12)
Lack of explanation to patients	4 (3.4)	2 (6)	2 (6)	0
Miscalculation	4 (3.4)	0	1 (3)	3 (6)

[a]There were cases where multiple error events were segregated from a single incident report.

Comparison of Models by Adaptation Levels

The BNM results for the 3 adaptation groups (low, moderate, and high) are presented in Multimedia Appendices 1-3, and the comparison of their structural characteristics is summarized in Table 3. In each model, the incident event was positioned at the lowest level, while the most distal contributing factors were represented as level 1. The average entropy of the incident events ranged from 0.44 to 0.76 in the low-adaptation model, from 0.55 to 0.78 in the moderate-adaptation model, and from 0.37 to 0.75 in the high-adaptation model. In the low-adaptation model (Multimedia Appendix 1), the structure consisted of up to 5 layers. Four nodes were excluded from the model: 2 event nodes ("Incorrect storage" and "Wrong route") and 2 factor nodes ("Patients who do not follow instructions" and "Miscalculation"). Three factor nodes were included in the model but were not connected to any event nodes: "Lack of explanation to patients," "Failure of information communication," and "Time pressure."

XSL•FO
RenderX

Table . Comparison of model structures.

Comparison criteria	Low-adaptation model (n=33)	Moderate-adaptation model (n=36)	High-adaptation model (n=50)
Maximum number of layers	5	4	4
Number of events not included in the model	2	2	1
Node	Incorrect storage and wrong route	Wrong timing and wrong route	Wrong timing
Number of factors not included in the model	2	2	5
Node	Patients who do not follow instructions, miscalculation	Year as registered nurse, task interruption	Time pressure, assumptions and forgetfulness, failure of monitoring, lack of explanation to patients, miscalculation
Number of factors not related to events	3	7	5
Node	Time pressure, failure of information communication, and lack of explanation to patients	Failure to detect by the electronic system, assumptions and forgetfulness, invalid double check, failure of information communication, non-reconfirming of inappropriate doctor's orders, lack of explanation to patients, miscalculation	Work time, failure of doctor's orders, insufficient knowledge of medications, failure of information communication, non-reconfirming of inappropriate doctor's orders

The moderate-adaptation model (Multimedia Appendix 2) had a maximum of 4 layers, with 4 nodes excluded from the model: 2 event nodes ("Wrong timing" and "Wrong route") and 2 factor nodes ("Years as registered nurse" and "Task interruption"). Seven factor nodes were included in the model but were not connected to any event nodes: "Lack of explanation to patients," "Miscalculation," "Failure of information communication," "Failure to detect by the electronic system," "Invalid double check," "Assumptions and forgetfulness," and "Non-reconfirming of inappropriate doctor's orders."

In the high-adaptation model (Multimedia Appendix 3), the structure consisted of up to 4 layers. Six nodes were excluded from the model: one event node ("Wrong timing") and 5 factor nodes ("Time pressure," "Assumptions and forgetfulness," "Miscalculation," "Lack of explanation to patients," and "Failure of monitoring"). Five factor nodes were included in the model but were not connected to any event nodes: "Insufficient knowledge of medications," "Failure of information communication," "Work time," "Failure of doctor's orders," and "Non-reconfirming of inappropriate doctor's orders."

The node "Years as registered nurse" was included in both the low- and high-adaptation models (Multimedia Appendices 1 and 3). In the low-adaptation model, all categories except nurses with 5 to 9 years were positively associated with "Failure to confirm the five rights," and nurses with ≥10 years were positively associated with "Wrong or unnecessary administration." In the high-adaptation model, nurses with 5 to 9 years were positively associated with "Failure of information communication" and "Invalid double check," whereas all categories except nurses with less than 5 years were positively associated with "Wrong dose."

Discussion

Key Results

This hypothesis-generating study examined whether years of experience of nurses in their current unit, used as an indicator of unit adaptation, might be related to medication-related errors. While prior qualitative studies have highlighted the importance of adaptation [40], our mixed methods approach using an AI-based BNM suggests the potential for a more objective and quantitative assessment.

Incident causation inherently involves uncertainty because incidents arise from complex interactions among contextual and behavioral factors rather than predetermined outcomes. The moderate to slightly high entropy values observed in this study reflect this uncertainty and should be interpreted as relative indicators. Future studies with larger samples and additional variables, such as work environment factors, may help improve model stability and reliability.

BNM integrates expert judgment while reducing subjectivity through probabilistic modeling, differing from approaches that rely solely on RCA. Consistent with the Reason Swiss Cheese Model [41], our analysis visualized a multilayered structure of medication-related errors. Especially in the low-adaptation model, unsafe conditions and behaviors appeared to precede subsequent unsafe actions, suggesting a potential chain of events. The findings provide preliminary insights into how unit adaptation may influence error structures although this study is limited by its sample size and the interpretive constraints inherent to probabilistic modeling.

Objectivity of RCA, Conceptual Validity, Causality, and Robustness of Hierarchical Structure Models

International studies on medication incidents have mainly reported that wrong dose errors account for a large proportion

XSL•FO
RenderX

of events and that adherence to the "five rights" is a critical factor [37]. The descriptive statistics of incident events and contributing factors in the present study were generally consistent with these prior findings. Although the existing literature is limited, no notable bias was observed in the dataset used in this study.

With respect to RCA, factor extraction was deliberately restricted to elements that could be objectively derived from narrative descriptions; subjective elements prone to inconsistent reporting—such as psychological states and perceptions—were excluded. Furthermore, each incident event and its contributing factors were coded using predefined operational definitions to safeguard objectivity in the extraction process.

For the BNM, relationships among factors were specified with reference to human factors theory, drawing on frameworks [37,39]. These specifications informed the blacklist constraints, thereby supporting the conceptual validity of the resulting network structures.

Regarding the hierarchical structure models, the intent was not to establish definitive causal relationships but to conduct a hypothesis-generating exploration. All nurse knowledge and behavioral factors identified through RCA were treated as potential contributors, and the BNM was used to examine shared structural patterns and differences by unit experience. A greedy algorithm guided by the Akaike information criterion was selected to detect weak associations and latent risks without excluding relevant links, allowing modeling even with a relatively small sample size. However, the limited sample size constrains model stability and robustness, and the results should be interpreted as provisional hypotheses rather than confirmed evidence. The models do not determine causal relationships among factors or between factors and incidents. Future studies should increase the sample size and conduct cross-regional analyses to empirically test the hypotheses generated here.

Effect of Adaptation on Patient Safety

As the concept of adaptive performance suggests, nurses' adaptation to their unit is expected to enhance their problem-solving abilities, predictive practice skills, behavioral flexibility, and adjustment to the organizational culture and physical environment in highly uncertain clinical settings [19]. The findings of this study appear to be consistent with these expectations.

As shown in Table 2, incident-related factors did not decline sharply with increasing adaptation. However, the comparison of the 3 model structures (Multimedia Appendices 1-3) suggests that the overall number of layers leading to incident occurrence becomes fewer as adaptation increases. One possible interpretation is that the low-adaptation group shares more homogeneous incident pathways, making certain unsafe conditions more likely to trigger errors. This trend underscores the importance of population-level patient safety education that supports early adaptation among novice and newcomer nurses.

Table 3 further shows that several factors—such as "Work time," "Failure of information communication," and "Non-reconfirming of inappropriate doctor's orders"—were not directly connected to incidents in the high- and moderate-adaptation models. Other factors, including "Time pressure," "Insufficient knowledge of medications," and "Lack of explanation to patients," were not included in the models at all. These patterns suggest two potential effects of unit adaptation: (1) improved adaptation to work processes, such as anticipatory multitasking, increased psychological resilience, acquisition of unit-specific pharmacological knowledge, and the ability to provide appropriate patient explanations [42]; and (2) enhanced adaptation to organizational culture, including smoother communication with colleagues and other health care professionals [43,44]. These interpretations remain tentative, and larger samples are needed to assess the stability and robustness of these structural differences.

Across all models, a consistent structure emerged involving "Wrong patient" and its associated factors—"Failure to confirm the 5 rights" and "Improper use of instruments and equipment." Addressing these issues will likely require approaches beyond those focused solely on promoting unit adaptation.

Considerations Regarding the Impact of Extensive Clinical Experience on Patient Safety

Previous studies on clinical experiences have primarily focused on patient safety measures for novice nurses with limited experience [22-25]. However, the inclusion of the node "Years as registered nurse" in both the low- and high-adaptation models indicated that extensive clinical experiences did not necessarily contribute to reducing errors. Instead, positive associations were observed with behaviors such as "Failure to confirm the five rights" and "Invalid double check," suggesting lapses in verification behaviors.

One possible explanation is the emergence of inappropriate adaptation, in which safety behaviors become ritualized and lose their intended effectiveness. Another possibility is the presence of a cognitive bias whereby experienced nurses are assumed by other staff to be less likely to make errors, leading to the invalidation of double-checking procedures [45]. Although the potential negative effects of extensive clinical experiences have rarely been considered, and the present findings do not allow us to conclude a causal relationship between greater years of experience and errors in verification behaviors, the consistent adherence to safety-verification practices should apply to all nurses, regardless of their level of experience.

Practical Implications

The most significant contribution of this study lies in suggesting that, while previous research has emphasized the lack of clinical experiences among novice nurses as a major issue for patient safety [22-25], our analysis of real-world data indicates that challenges in adapting to clinical units may also underlie this issue. This represents a novel application of AI in nursing education [46]. These findings underscore the potential importance of supporting nurses' adaptation to their assigned units—not only for newly graduated nurses but also for those who have recently transferred to new settings. Educational approaches that emphasize adaptation may also play a meaningful role in the training of other health care professionals. Further hypothesis-testing research will be required.

XSL•FO
RenderX

As possible practical interventions, intensive orientation programs focusing on unit-specific workflows during the initial assignment phase, along with educational initiatives aimed at improving medication knowledge and fostering safe practices, may be effective. In the field of safety behavior training, virtual reality programs have been developed and applied to surgical education [47]. Following this example, scenario-based simulation training can be designed using the results of the present models to reproduce situations in which unsafe nursing behaviors are likely to cascade.

Furthermore, the models presented in this study can be used as feedback tools for safety management and education, reflecting individual differences in clinical experiences and adaptation levels. Although adverse events and medical errors are often influenced by multiple latent factors [37], postevent instructions tend to focus primarily on direct causes due to hindsight bias. The multilayered structures revealed in the current models may help overcome this bias by encouraging more comprehensive and proactive safety-management strategies.

Limitations and Future Work

We conducted RCA based on narrative descriptions of reported medical errors. Therefore, situations or latent factors unrecognized by reporting nurses may not have been captured in the analysis. In particular, nurses with extensive clinical experiences often care for critically ill patients or manage heavier workloads, which may serve as underlying factors contributing to counterproductive behaviors, such as task shortcuts or neglect of verification behaviors [48]. However, this limitation is not unique to this study but rather reflects a structural issue common to existing incident reporting systems worldwide. Indeed, prior studies on nurse fatigue risk management have rarely incorporated variables such as patient load and workload in a systematic manner. In addition, even in routine clinical practice, systems for consistently recording indicators such as the number of patients per nurse or patient acuity for the purpose of workload analysis are often lacking. If quantitative approaches to analyzing incident reports—such as those explored in our study—become more established, they may contribute to future international discussions on revising reporting standards.

Another limitation concerns the balance between global and local optimization in the analysis; that is, the extent to which institutional characteristics are incorporated depends on the research purpose and may affect the generalizability of the results [2]. In this study, organizational factors specific to individual institutions could not be extracted from the report narratives. Future research should compare data across institutions and design unit-specific or institution-specific training programs, followed by an evaluation of their effectiveness and acceptability [49].

The findings from the hierarchical network models in this study should be regarded as exploratory and hypothesis-generating rather than definitive. Future work is needed to expand the sample and to evaluate the stability and robustness of the model structures through random-sampling approaches, such as analyzing pre- and post-COVID-19 datasets [50] and employing bootstrap methods [51].

The variable “Years of nursing experience” spans a broad continuous range, so it was converted into categorical data using 5-year intervals, following classifications commonly used in clinical ladder systems [21,33]. This categorization, however, is not definitive. Greater experiences may also reflect age-related factors, and although declines in work ability with age are well recognized in occupational health [52], they have rarely been examined in medical or nursing contexts. Further research is needed to clarify how nursing experience relates to work performance and to determine more appropriate ways of grouping experience levels.

Conclusions

This hypothesis-generating study applied an AI-based BNM to examine structural patterns underlying medication-related errors, demonstrating the value of a quantitative, probabilistic approach that goes beyond the traditional RCA. The use of BNM represents a novel methodological contribution in patient safety research by combining expert judgment with reduced subjectivity and enabling structural comparison across experience groups. The findings suggest that unit adaptation plays an important role in shaping error pathways: low-adaptation nurses showed more layered and homogeneous structures leading to unsafe actions, indicating a higher vulnerability to incident chains. In contrast, total years of nursing experience did not uniformly reduce risk and may even relate to lapses in verification behaviors. Together, these results highlight the need to support adaptation not only among new graduates but also among nurses transitioning to new units, while reinforcing verification practices across all experience levels. Further empirical studies are needed to validate these preliminary hypotheses.

Acknowledgments

Generative artificial intelligence tools such as DeepL and ChatGPT were used during the manuscript preparation process to standardize citation formatting and to assist with the initial English translation of the Japanese manuscript. Prior to submission, the manuscript was reviewed and professionally edited by an academic editing service to ensure accuracy and clarity.

Funding

The authors declared no financial support was received for this work.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: NS, KO
Data curation: NS
Formal analysis: NS
Methodology: NS
Project administration: KO
Writing – original draft: NS
Writing – review and editing: NS, KO

Conflicts of Interest

None declared.

Multimedia Appendix 1
The low-adaptation model.
[PNG File, 574 KB - nursing_v9i1e87436_app1.png]

Multimedia Appendix 2
The moderate-adaptation model.
[PNG File, 610 KB - nursing_v9i1e87436_app2.png]

Multimedia Appendix 3
The high-adaptation model.
[PNG File, 608 KB - nursing_v9i1e87436_app3.png]

References

1. Hervatis V, Loe A, Barman L, O'Donoghue J, Zary N. A conceptual analytics model for an outcome-driven quality management framework as part of professional healthcare education. JMIR Med Educ 2015 Oct 6;1(2):e11. [doi: 10.2196/mededu.4789] [Medline: 27731840]
2. Mohamed B, Fahy BG. The impact of data-driven learning on graduate medical education. J Clin Anesth 2022 Jun;78:110621. [doi: 10.1016/j.jclinane.2021.110621] [Medline: 34872825]
3. Lin L, Zhou D, Wang J, Wang Y. A systematic review of big data driven education evaluation. Sage Open 2024 Apr;14(2). [doi: 10.1177/21582440241242180]
4. Patient safety. World Health Organization. URL: https://www.who.int/features/factfiles/patient_safety/en/ [accessed 2025-10-13]
5. Tella S, Liukka M, Jamookeeah D, Smith NJ, Partanen P, Turunen H. What do nursing students learn about patient safety? An integrative literature review. J Nurs Educ 2014 Jan 1;53(1):7-13. [doi: 10.3928/01484834-20131209-04] [Medline: 24308538]
6. Lee SE, Morse BL, Kim NW. Patient safety educational interventions: a systematic review with recommendations for nurse educators. Nurs Open 2022 Jul;9(4):1967-1979. [doi: 10.1002/nop2.955] [Medline: 34047058]
7. The essentials: core competencies for professional nursing education. : American Association of Colleges of Nursing; 2021 URL: https://www.aacnnursing.org/Portals/0/PDFs/Publications/Essentials-2026.pdf [accessed 2025-10-13]
8. Wittmann-Price RA. Competency-based curriculum development. In: Wittmann RA, Gittings KK, editors. Fast Facts About Competency-Based Education in Nursing How to Teach Competency Mastery, 1st edition: Springer; 2020:17-35.
9. Grigsby K. Curriculum models for graduate programs. In: Billings DM, Halstead JA, editors. Teaching in Nursing: A Guide for Faculty, 6th edition: Elsevier; 2019:163-180.
10. Currie L, Watterson L. Challenges in delivering safe patient care: a commentary on a quality improvement initiative. J Nurs Manag 2007 Mar;15(2):162-168. [doi: 10.1111/j.1365-2834.2007.00627.x] [Medline: 17352699]
11. Henneman EA, Roche JP, Fisher DL, et al. Error identification and recovery by student nurses using human patient simulation: opportunity to improve patient safety. Appl Nurs Res 2010 Feb;23(1):11-21. [doi: 10.1016/j.apnr.2008.02.004] [Medline: 20122506]
12. Rinaldi K, Messer M, Hanson A, Chan J. Utilization of backward design in health professional education: a rapid review. J Prof Nurs 2025;58:31-38. [doi: 10.1016/j.profnurs.2025.02.004] [Medline: 40368493]
13. Hodkinson A, Tyler N, Ashcroft DM, et al. Preventable medication harm across health care settings: a systematic review and meta-analysis. BMC Med 2020 Nov 6;18(1):313. [doi: 10.1186/s12916-020-01774-9] [Medline: 33153451]
14. Panagioti M, Khan K, Keers RN, et al. Prevalence, severity, and nature of preventable patient harm across medical care settings: systematic review and meta-analysis. BMJ 2019 Jul 17;366:l4185. [doi: 10.1136/bmj.l4185] [Medline: 31315828]

XSL•FO
RenderX

15. Sugimura N, Ogasawara K. Events related to medication errors and related factors involving nurses' behavior to reduce medication errors in Japan: a Bayesian network modeling-based factor analysis and scenario analysis. J Educ Eval Health Prof 2024;21:12. [doi: 10.3352/jeehp.2024.21.12] [Medline: 38858820]
16. Griffin MA, Neal A, Parker SK. A new model of work role performance: positive behavior in uncertain and interdependent contexts. Acad Manage J 2007 Apr;50(2):327-347. [doi: 10.5465/amj.2007.24634438]
17. Sugimura N, Sato M, Sumi N, Yano R. Validity and reliability of the Japanese version of the Nursing Performance Instrument. Jpn J Nurs Sci 2023 Apr;20(2):e12518. [doi: 10.1111/jjns.12518] [Medline: 36344443]
18. Dawson D, Chapman J, Thomas MJW. Fatigue-proofing: a new approach to reducing fatigue-related risk using the principles of error management. Sleep Med Rev 2012 Apr;16(2):167-175. [doi: 10.1016/j.smrv.2011.05.004] [Medline: 21784677]
19. Gander P, Hartley L, Powell D, et al. Fatigue risk management: organizational factors at the regulatory and industry/company level. Accid Anal Prev 2011 Mar;43(2):573-590. [doi: 10.1016/j.aap.2009.11.007] [Medline: 21130218]
20. Benner P, Tanner C, Chesla C. Expertise in Nursing Practice: Caring, Clinical Judgment, and Ethics, 1st edition: Springer; 1996.
21. Slagle A, Wakim N, Gray SE. A global examination of clinical ladder programs—a synthesis of commonalities and opportunities for standardization. Worldviews Evid Based Nurs 2023 Feb;20(1):56-63. [doi: 10.1111/wvn.12622] [Medline: 36604948]
22. Hou M, Wei W, Wang S, Hu J, Chen Y. Challenge-hindrance stressors and novice nurses' safety behaviour: the mediating role of regulatory focus and the moderating role of workplace spirituality. J Adv Nurs 2025 Nov;81(11):7597-7608. [doi: 10.1111/jan.16723] [Medline: 39740076]
23. Labrague LJ, De Los Santos JAA. Transition shock and newly graduated nurses' job outcomes and select patient outcomes: a cross-sectional study. J Nurs Manag 2020 Jul;28(5):1070-1079. [doi: 10.1111/jonm.13033] [Medline: 32315478]
24. Yan J, Li L, Li J, et al. Stepwise interactive situated training program for young nurses' safety behavior and interrupted coping behavior. Healthcare (Basel) 2022 Jun 21;10(7):1157. [doi: 10.3390/healthcare10071157] [Medline: 35885683]
25. Ebright PR, Urden L, Patterson E, Chalko B. Themes surrounding novice nurse near-miss and adverse-event situations. J Nurs Adm 2004 Nov;34(11):531-538. [doi: 10.1097/00005110-200411000-00010] [Medline: 15586075]
26. Battles JB, Shea CE. A system of analyzing medical errors to improve GME curricula and programs. Acad Med 2001 Feb;76(2):125-133. [doi: 10.1097/00001888-200102000-00008] [Medline: 11158830]
27. Outline [Article in Japanese]. Japan Council for Quality Health Care. URL: https://www.med-safe.jp/contents/english/index.html [accessed 2025-10-13]
28. Fukada M. Nursing competency: definition, structure and development. Yonago Acta Med 2018 Mar;61(1):1-7. [doi: 10.33160/yam.2018.03.001] [Medline: 29599616]
29. Mrayyan MT, Abunab HY, Abu Khait A, et al. Competency in nursing practice: a concept analysis. BMJ Open 2023 Jun 1;13(6):e067352. [doi: 10.1136/bmjopen-2022-067352] [Medline: 37263688]
30. Almarwani AM, Alzahrani NS. Factors affecting the development of clinical nurses' competency: a systematic review. Nurse Educ Pract 2023 Nov;73:103826. [doi: 10.1016/j.nepr.2023.103826] [Medline: 37951064]
31. Lin YW, Ni CF, Hsu SF, Tsay SL, Tung HH. Effects of length of employment and head nurse leadership style on the clinical competency of staff nurses in Taiwan. J Nurs Res 2024 Jun 1;32(3):e331. [doi: 10.1097/jnr.0000000000000617] [Medline: 38814996]
32. Saiga M, Yamamoto Y, Okuda R, Fukada M. Relationship between clinical nursing competence and work environment by career stage for nurses with 1-10 years of clinical experience. Yonago Acta Med 2024;67(1):9-21. [doi: 10.33160/yam.2024.02.002] [Medline: 38371280]
33. Career philosophy [Article in Japanese]. Hokkaido University Hospital Nursing Department. URL: https://www.kango-hokudai.jp/career [accessed 2026-03-01]
34. Campbell JP. Modeling the performance prediction problem in industrial and organizational psychology. In: Dunnette MD, Hough LM, editors. Handbook of Industrial and Organizational Psychology: Consulting Psychologists Press; 1990:687-732.
35. Viswesvaran C, Ones DS. Perspectives on models of job performance. Int J Sel Assess 2000 Dec;8(4):216-226. [doi: 10.1111/1468-2389.00151]
36. Koopmans L, Bernaards CM, Hildebrandt VH, Schaufeli WB, de Vet Henrica CW, van der Beek AJ. Conceptual frameworks of individual work performance: a systematic review. J Occup Environ Med 2011 Aug;53(8):856-866. [doi: 10.1097/JOM.0b013e318226a763] [Medline: 21775896]
37. Mitchell RJ, Williamson A, Molesworth B. Use of a human factors classification framework to identify causal factors for medication and medical device-related adverse clinical incidents. Saf Sci 2015 Nov;79:163-174. [doi: 10.1016/j.ssci.2015.06.002]
38. Zhao L, Wang X, Qian Y. Analysis of factors that influence hazardous material transportation accidents based on Bayesian networks: a case study in China. Saf Sci 2012 Apr;50(4):1049-1055. [doi: 10.1016/j.ssci.2011.12.003]
39. Lee JY, Huang CH, Sie YA, Yang PC, Su CC, Chang JT. Applying the human factors analysis and classification system within root cause analysis to prevent medical errors and enhancing patient safety culture: insights from a medical center. Int J Qual Health Care 2025 Feb 27;37(1):mzaf009. [doi: 10.1093/intqhc/mzaf009] [Medline: 39936321]

XSL•FO
RenderX

40. Jung S, Park J. Educational needs for medication safety competence among nurses by clinical ladder stage. PLoS One 2025;20(4):e0319483. [doi: 10.1371/journal.pone.0319483] [Medline: 40238821]
41. Wiegmann DA, Wood LJ, Cohen TN, Shappell SA. Understanding the “Swiss cheese model” and its application to patient safety. J Patient Saf 2022 Mar 1;18(2):119-123. [doi: 10.1097/PTS.0000000000000810] [Medline: 33852542]
42. Alsabri M, Boudi Z, Lauque D, et al. Impact of teamwork and communication training interventions on safety culture and patient safety in emergency departments: a systematic review. J Patient Saf 2022 Jan 1;18(1):e351-e361. [doi: 10.1097/PTS.0000000000000782] [Medline: 33890752]
43. Keshtkar L, Bennett-Weston A, Khan AS, et al. Impacts of communication type and quality on patient safety incidents: a systematic review. Ann Intern Med 2025 May;178(5):687-700. [doi: 10.7326/ANNALS-24-02904] [Medline: 40228297]
44. Kwame A, Petrucka PM. A literature-based study of patient-centered care and communication in nurse-patient interactions: barriers, facilitators, and the way forward. BMC Nurs 2021 Sep 3;20(1):158. [doi: 10.1186/s12912-021-00684-2] [Medline: 34479560]
45. Eldor L, Hodor M, Cappelli P. The limits of psychological safety: nonlinear relationships with performance. Organ Behav Hum Decis Process 2023 Jul;177:104255. [doi: 10.1016/j.obhdp.2023.104255]
46. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on nursing education: scoping review. JMIR Nurs 2021;4(1):e23933. [doi: 10.2196/23933] [Medline: 34345794]
47. Mazur L, Butler L, Mitchell C, et al. Effect of immersive virtual reality teamwork training on safety behaviors during surgical cases: nonrandomized intervention versus controlled pilot study. JMIR Med Educ 2025 May 1;11:e66186. [doi: 10.2196/66186] [Medline: 40311122]
48. Rotundo M, Sackett PR. The relative importance of task, citizenship, and counterproductive performance to global ratings of job performance: a policy-capturing approach. J Appl Psychol 2002;87(1):66-80. [doi: 10.1037/0021-9010.87.1.66] [Medline: 11916217]
49. Lee JWY, Tan JY, Bello F. Technology acceptance model in medical education: systematic review. JMIR Med Educ 2025 Jul 16;11:e67873. [doi: 10.2196/67873] [Medline: 40668663]
50. Clari M, Luciani M, Conti A, et al. The impact of the COVID-19 pandemic on nursing care: a cross-sectional survey-based study. J Pers Med 2021 Sep 23;11(10):945. [doi: 10.3390/jpm11100945] [Medline: 34683086]
51. Broom BM, Do KA, Subramanian D. Model averaging strategies for structure learning in Bayesian networks with limited data. BMC Bioinformatics 2012;13 Suppl 13(Suppl 13):S10. [doi: 10.1186/1471-2105-13-S13-S10] [Medline: 23320818]
52. von Bonsdorff ME, Kokko K, Seitsamo J, et al. Work strain in midlife and 28-year work ability trajectories. Scand J Work Environ Health 2011 Nov;37(6):455-463. [doi: 10.5271/sjweh.3177] [Medline: 21695372]

Abbreviations

AI: artificial intelligence
BNM: Bayesian network modeling
JCQHC: Japan Council for Quality Health Care
RCA: root cause analysis
WHO: World Health Organization

Edited by J Sarvestan; submitted 09.Nov.2025; peer-reviewed by MH Tsai, S Mohanadas; revised version received 17.May.2026; accepted 25.May.2026; published 07.Jul.2026.

Please cite as:
Sugimura N, Ogasawara K
Medication-Related Errors Among Nurses by Unit Adaptation Levels: Bayesian Network–Based Exploratory Study
JMIR Nursing 2026;9:e87436
URL: https://nursing.jmir.org/2026/1/e87436
doi:10.2196/87436

© Naotaka Sugimura, Katsuhiko Ogasawara. Originally published in JMIR Nursing (https://nursing.jmir.org), 7.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Implications for Virtual Nursing Role Development in Acute Nursing Care: 24-Hour Time-and-Motion Study

Jennifer Sumner[1,2], PhD; Emily Hwee Hoon Chew[2], MPH; Sin Yee Peck[3], BSc; Janet Lam[3], BSc; Pauline Chong[3], BSc; Mun Kwong Ching[4], BA; De Sheng Ong[2], MPH; Jacqueline Xiao Xuan Lau[2], BSc; Camille Keck[2], BSc; Abigail Ang[2], BA; Jaminah Mohamed Ali[1], BSc; Hui Wen Lim[1], MSc; Anjali Bundele[1], MPH; Elya Chen[2], MPH; Carol Yap[2], BSc; Evelyn Ho[2], BSc; Amartya Mukhopadhyay[1,5,6], MD; Alexander Wenjun Yip[2], MBBCh

[1]Medical Affairs – Research Innovation & Enterprise, Alexandra Hospital, 378 Alexandra Road, Singapore

[2]Department of Healthcare Redesign, Alexandra Research Centre for Healthcare in a Virtual Environment, Alexandra Hospital, Singapore

[3]Department of Nursing, Alexandra Hospital, Singapore

[4]Department of Data Science and Analytics, Alexandra Hospital, Singapore

[5]Yong Loo Lin School of Medicine, National University of Singapore, Singapore

[6]Division of Respiratory and Critical Care Medicine, National University Hospital, Singapore

Corresponding Author:
Jennifer Sumner, PhD
Medical Affairs – Research Innovation & Enterprise, Alexandra Hospital, 378 Alexandra Road, Singapore

Abstract

Background: Understanding current nursing workflows is essential to informing future workforce redesign strategies, including virtual nursing roles. However, granular insights into current nursing workflows over a 24-hour period and across different staff grades are lacking.

Objective: This study aimed to (1) quantify how registered and enrolled nurses in the general acute medicine wards distribute their time across direct and indirect care tasks over a 24-hour period, (2) identify multitasking burdens and temporal distributions, and (3) identify opportunities for the development of a virtual nursing role.

Methods: Using time-and-motion methodology, we observed registered and enrolled nurses in 3 general medicine wards over a 24-hour period between April 2024 and June 2024. We observed 3 task categories (administrative, communication, and bedside tasks), with multiple individual tasks monitored under each category. Multitasking (ie, the occurrence of 2 or more tasks concurrently) was also tracked. The checklist was piloted and refined before data collection.

Results: We observed a total of 48 nursing shifts. During the daytime, registered nurses spent 70% (587/834 min) of their time on indirect care tasks compared with 54% (412/764 min) of the time for enrolled nurses. At night, the proportion of time spent on indirect care tasks decreased to 58% (410/705 min) for registered nurses and 39% (274/711 min) for enrolled nurses. During a 24-hour period, registered nurses spent 209 (SD 51.8) minutes multitasking in the day and 117 (SD 41.0) minutes at night, whereas enrolled nurses spent 152 (SD 54.7) minutes multitasking in the day and 110 (SD 75.9) minutes at night.

Conclusions: These findings highlight opportunities for virtual nursing roles, which, if thoughtfully designed, may help redistribute indirect care tasks, reduce multitasking burden, and enhance overall efficiency without compromising care quality.

(JMIR Nursing 2026;9:e87930) doi:10.2196/87930

KEYWORDS

time and motion; nursing workflows; multitasking; inpatient; virtual nursing

Introduction

Background

Health systems are facing a myriad of workforce challenges. Populations are aging, increasing the demand for care and the complexity of the care required [1]. Simultaneously, the workforce is shrinking; by 2030, the World Health Organization projects an estimated shortfall of 11 million health workers [2]. Furthermore, high-income nations find it increasingly challenging to recruit and retain home talent, relying on foreign workers to fill staffing gaps [3]. Challenges in recruitment and retention are further exacerbated by the increasing levels of staff dissatisfaction and burnout reported in recent years [4]. Collectively, these challenges raise significant questions about the sustainability of health care if no change occurs.

To address workforce challenges, there is an urgent need for role redesign and the careful application of technology to streamline processes and automate tasks. At Alexandra Hospital

XSL•FO
RenderX

in Singapore, a new approach is being pioneered—the integration of a virtual workforce enabled through technology. Supported by the Ministry of Health, the hospital is investing in the development of remote or "virtual" nurses. In this context, virtual nursing refers to an emerging model of practice in which qualified nurses provide support to clinical team members remotely through digital technologies, thereby complementing bedside nursing. For example, virtual nurses may remotely monitor vital signs and fall risk, act as virtual discharge coordinators, participate in rounding through teleconsultation, or provide remote patient and caregiver education and staff supervision [5-8]. Additionally, the introduction of new virtual roles will provide more flexible work options such as working from home, which can support staff retention efforts [9]. Thus, virtual nursing is a promising model for workforce redesign.

To develop the virtual nursing workforce, understanding the status quo is important. Nurses are responsible for a range of direct (ie, involving direct interaction with patients) and indirect (ie, performed away from the patient) patient care tasks, including administering medication, monitoring patient conditions, coordinating with other health care providers, and managing administrative duties [10]. It is also well known that nurses handle diverse and often conflicting tasks, often leading to overwork and suboptimal practice [11] (eg, the duality of direct and indirect patient care tasks competing for attention). Furthermore, it has been widely reported that nurses frequently multitask (ie, perform multiple tasks simultaneously) throughout the day, which can create inefficiencies, increase the risk of errors, and reduce overall quality of care [12].

At the same time, nursing practice is fundamentally grounded in caring science, where the therapeutic nurse-patient relationship and presence are central to promoting healing and well-being [13]. However, growing administrative burden, competing indirect care demands, and multitasking can displace the relational aspects of care that form the essence of nursing. Therefore, understanding how nurses currently spend their time is essential not only for efficiency but to safeguard opportunities for caring practices. Role redesign—including virtual nursing—must be guided by the principle that technology should preserve and enhance nurses' ability to deliver person-centered, relational care.

Despite this, much of the existing evidence base is qualitative, and there remains a gap in quantitative measurement. Where quantitative data exist, there is limited data capture on nighttime activities, the occurrence and patterns of multitasking over a 24-hour period, and the tasks undertaken by different staff grades (eg, registered nurses, who are qualified to dispense medications and instruct junior nurses, and enrolled nurses, who primarily handle patient care, such as bathing, feeding, and toileting) [10,11,14,15]. This study addresses these gaps by systematically mapping real-world nursing activity over a 24-hour cycle, examining the distribution of direct and indirect care tasks, identifying when multitasking occurs, and comparing activity patterns across staff grades. By doing so, this study provides empirical evidence to inform the design, timing, and scope of virtual nursing roles, identifying activities suitable for remote support without compromising patient-facing care.

Aims

Our aims were as follows:

1. To quantify how registered and enrolled nurses in the general acute medicine wards distribute their time across direct and indirect care tasks over a 24-hour period
2. To identify multitasking burdens and temporal distributions
3. To identify opportunities for the development of a virtual nursing role

Methods

Overview

We used time-and-motion methodology to map and quantify the tasks and time spent on direct and indirect patient care in an inpatient setting. In brief, time and motion involves a period of continuous observation and timing of activities. This study is reported according to the Suggested Time and Motion Procedures checklist [16].

Participants

This study was conducted at Alexandra Hospital in 3 general acute medicine wards. Ward sizes range from 19 to 33 beds, and each has several isolation beds where additional infection control measures are in place (eg, methicillin-resistant *Staphylococcus aureus* precautions). The typical nurse-to-patient ratio is 1 nurse for every 4 or 5 patients, with nurses opting for various full-day (12-hour) shifts or partial day shifts. A typical ward comprises a team of registered nurses (who are qualified to dispense medications and instruct junior nurses), enrolled nurses (who primarily handle patient care, such as bathing, feeding, and toileting), and a nurse manager who oversees ward operations. Each nurse has access to the electronic medical record (EMR) system through desktop computers or mobile phones (ie, Epic Rover).

As the roles of registered nurses and enrolled nurses are distinct, we observed both profiles through convenience sampling during a 3-month period. Inclusion criteria were being a full-time staff member and having worked on the wards for at least 6 months. Exclusion criteria were basic care assistants, part-time workers, or those with less than 6 months of experience on the wards. The month before data collection commenced, briefing sessions were held with the nurse managers to introduce the research project and seek their approval. Observers (health service research staff) had no direct relationship with the nursing staff being observed.

Data Collection

The initial checklist of observation items was developed through consultation with the ward managers and comprised 3 overarching domains (administrative, communication, and bedside tasks), 8 task subcategories (activities of daily care, patient communication, clinical care, administrative tasks, staff communication, staff education, ward observation, and break time), and 27 individual task codes, mapped in Table 1. The checklist was piloted and refined during observations between November 29, 2023, and December 8, 2023, before the actual data collection commenced. A second pilot was conducted from April 9 to 12, 2024, to standardize the observation process and

XSL•FO
RenderX

train the 10 observers. Tasks were observed and tracked concurrently to capture periods of multitasking (ie, when 2 or more tasks were performed at the same time). For example, a communication task could be timed at the same time as a bedside care task. In our study, transition periods (short intervals between the end of one task and the start of another) were recorded as part of the subsequent task whenever the transition was functionally inseparable from task initiation (eg, walking from a workstation to a patient's bedside to begin care).

Data collection consisted of continuous nurse observation over 1 complete work shift (8:30 AM-8 PM or 8 PM-8:30 AM). Observations were conducted between April 15, 2024, and June 24, 2024. A roster was created in Python (version 3.8; Python Software Foundation) randomly allocating observers across different days of the week and weekend slots across the 3 wards. We used the Time Capture Tool TimeCaT on an iPad (Apple Inc) to track activities. TimeCaT is a validated tool frequently used in observations of health care processes [17]. Tasks are entered into the app a priori, and observers select the appropriate task as it occurs, activating a timer. Observers can also add notes to specific data points. During data collection, we also captured bed occupancy, the number of patients discharged, and the number of bed days.

Table . Direct and indirect care tasks, subcategories, and individual task codes.

Task categories and subcategories	Task codes
Direct care tasks	
Supporting activities of daily living	• Ambulation • Bathing • Changing bedding • Diapering or assisting with toileting • Assisting with food and drink
Patient communication	• Nurse to patient • Meal ordering • Attending patients' call bell
Clinical care	• Vital sign taking • Laboratory samples (collection and sending) • Other assessments and procedures • Medication handling • Patient comfort • Patient transfer (diagnostics) • Waste disposal (clinical and nonclinical waste)
Indirect care tasks	
Administrative	• Electronic medical record entry • Record viewing (no data entry) • Nonclinical administrative tasks • Discharge or admission
Staff communication	• In-person staff member to staff member • Remote staff member to staff member • Remote staff member to caregiver • Handover • Team meetings
Staff education	• Staff education
Ward observation	• Observing patients
Break time	• Break

Ethical Considerations

This study was reviewed and approved by the National University of Singapore Ethical Review Board (NUS-IRB-2023-327). A waiver for consent was obtained from the National University of Singapore Institutional Review Board. Data were anonymous and analyzed at the aggregate level. Participants were not individually compensated.

Data Analysis

Before analysis, each observation data file was cleaned. TimeCaT includes a note function that allows observers to comment on specific data points. Each note was reviewed, and revisions were made to the data if required (eg, removal of tasks that were erroneously timed or relabeling of incorrect observations). We then mapped the observed tasks to direct and indirect care tasks, creating subcategories within each domain. Direct care tasks involved direct interaction with the patient

XSL•FO
RenderX

(eg, bathing and procedures), and indirect care tasks were those that were performed away from the patient (eg, documentation and staff communication) but on behalf of the patient (Table 1).

Time-and-motion data were analyzed in multiple steps. To calculate the average time spent on each task, we summed the total time spent on each task across all observations, including instances in which the task was not performed (ie, zero time), and divided it by the total number of observations. This produced the average task load time for a typical day or night shift. We then calculated the total task load time for a shift by summing all average task times and used this value to determine the proportion of time spent on each task as a fraction of the total task load time. To analyze the data by hour, each observed task was assigned to a 1-hour "time bin." For example, a task performed from 9:01 AM to 9:20 AM was categorized under the 9 AM to 10 AM time bin and recorded as 19 minutes. If a task spanned multiple time bins, its duration was divided accordingly. This approach enabled accurate tracking of task distribution across the 24-hour period. We also measured multitasking by calculating the total time during which 2 or more tasks were performed simultaneously (eg, communication and administrative tasks). Periods of no activity were excluded from the analysis. All data processing and analyses were conducted using Python (version 3.8) and Stata (version 19; StataCorp LLC).

Validity and Reliability

The observation checklist was tested and optimized over 6 separate pilot observations (a total of 17.5 hours) between November 29, 2023, and December 8, 2023. Prior to the actual data collection, a training session was conducted (April 8, 2024) to introduce the team of observers to the process and address any initial questions. A set of training observations was then conducted in pairs from April 9 to 12, 2024, to familiarize the observers with the process, standardize coding of observed tasks, and ensure that the checklist was complete. The team met 4 times during this period to discuss coding alignment and resolve any questions. Interrater agreement was assessed during the pilot phase of the study [18]. As tasks across the 3 domains (administrative, communication, and bedside) could occur concurrently, the Cohen κ statistic and percentage of agreement were calculated separately for each domain. For each observer pair, a 30-minute observation period was analyzed. The observational data were divided into 30-second intervals, and agreement between coders was assessed based on the task assigned to each interval. κ values ranged from 0.62 to 0.69, indicating substantial agreement, whereas the percentage of agreement ranged from 78% to 84%, reflecting good consistency between coders. Formal data collection commenced on April 15, 2024.

Results

Overview

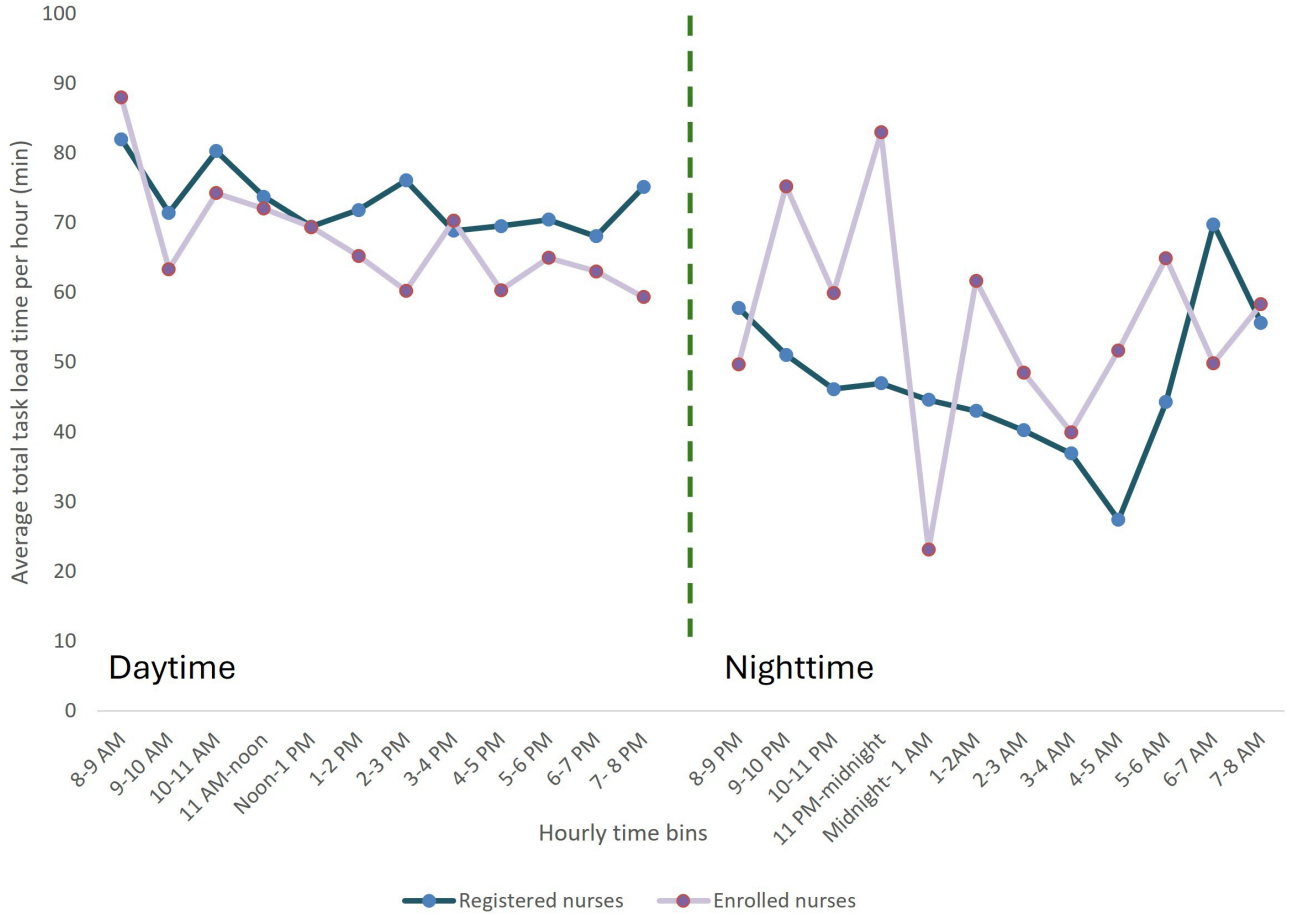
During the study period, we observed 48 shifts, including a total of 30 nurses during the day (8:30 AM-8 PM; n=22, 73.3% registered nurses and n=8, 26.7% enrolled nurses) and 18 nurses at night (8 PM-8:30 AM; n=14, 77.8% registered nurses and n=4, 22.2% enrolled nurses). The total active task observation time for registered nurses was 164 hours and 29 minutes, and for enrolled nurses, it was 47 hours and 22 minutes. During the observation period, bed occupancy ranged from 71% to 90%, and 1044 patients were discharged from the wards, representing 5559 bed days.

Task Load Over a 24-Hour Period

Task load varied over the course of 24 hours, with higher variability at night—particularly for enrolled nurses (Figure 1). During the day, the highest peak in task load for registered and enrolled nurses was between 8 AM and 9 AM, whereas the lowest task load was between 6 PM and 7 PM for registered nurses and between 7 PM and 8 PM for enrolled nurses. In the night shift, the highest task load peak was between 6 AM and 7 AM for registered nurses and between 11 PM and midnight for enrolled nurses. Conversely, the lowest task load was observed from 4 AM to 5 AM for registered nurses and midnight to 1 AM for enrolled nurses.

XSL•FO
RenderX

Figure 1. Total task load (minutes) by hour for a 24-hour period (task load by hour exceeds 60 minutes due to multitasking) for registered nurses (blue) and enrolled nurses (pink).



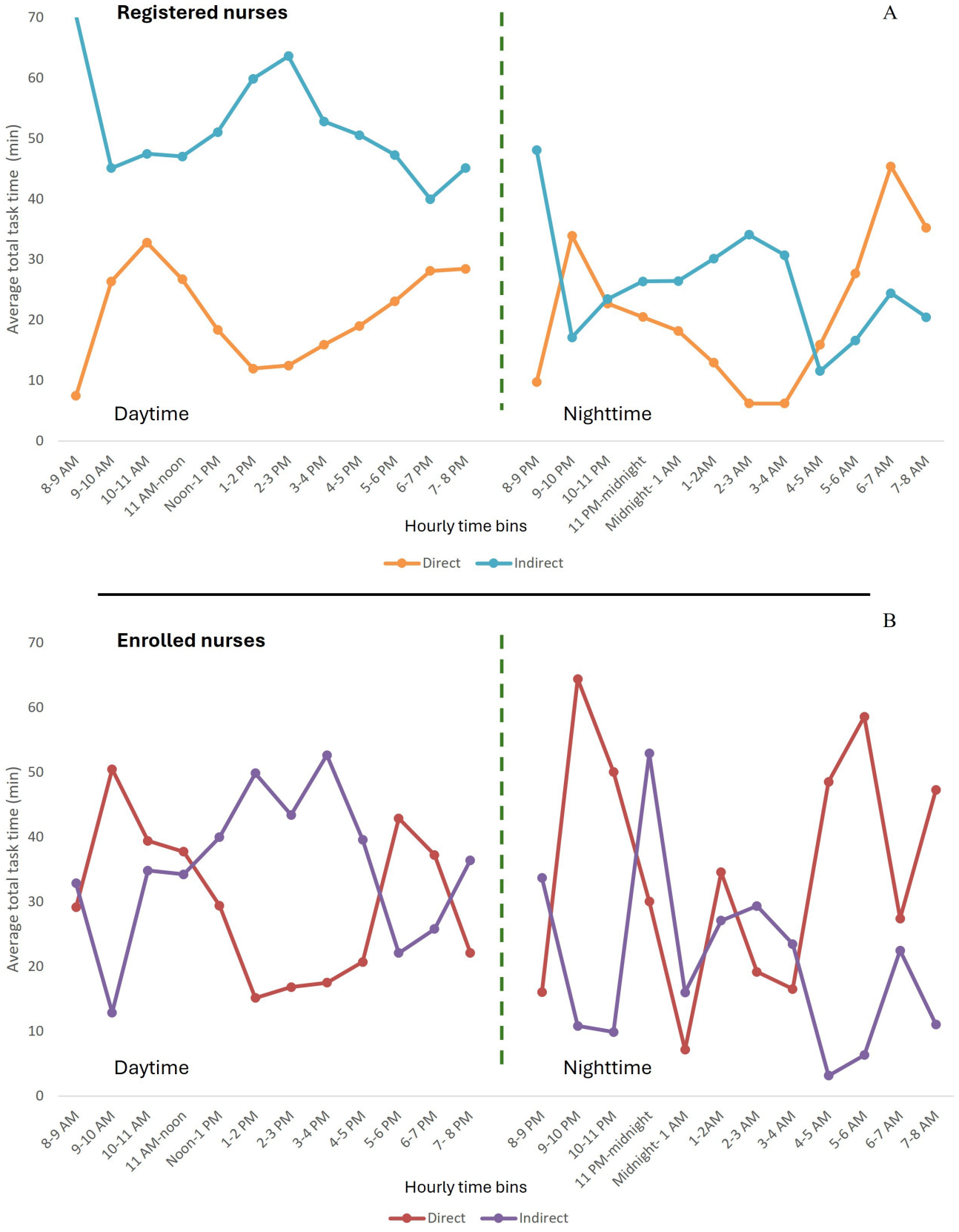
Task Type: Direct and Indirect Nursing Care Tasks

During the day, registered nurses spent an average of 30% (248 min; SD 74.6) of their time on direct clinical care and 70% (587 min; SD 98.9) of their time on indirect care tasks (Figure 2 and Table 2). Enrolled nurses, in contrast, spent an average of 46% (351 min; SD 52.5) of their time on direct care tasks and 54% (411 min; SD 73.2) of their time on indirect care tasks. At night, the proportion of time spent on direct care tasks increased for both groups: registered nurses spent an average of 42% (295 min; SD 78.9) of their time on direct care tasks, whereas enrolled nurses spent 61% (436 min; SD 98.5) of their time on direct care tasks, with the remainder dedicated to indirect care tasks (410 min, 58%; SD 67.2) for registered nurses, and 39% (275 min; SD 31.7) for enrolled nurses.

For registered nurses, indirect tasks dominated, with a consistently higher average total task load during the daytime (Figure 2). Direct task load typically remained at less than 30 minutes per hour, with a peak between 10 AM and noon and between 6 PM and 8 PM. At night, there was a noticeable drop in time spent on indirect tasks, and time spent on direct care tasks gradually decreased until 3 AM. From 3 AM onward, direct care tasks increased, peaking at 6 AM. For enrolled nurses, time spent on direct care tasks tended to dominate over time spent on indirect care tasks for much of the day, especially in the early morning and late afternoon, with a dip in the middle of the day. At night, there was considerable variability in both direct and indirect care tasks. There were notable spikes in direct care tasks (eg, around 11 PM-1 AM and 5 AM - 7 AM) and indirect care tasks (eg, 11 PM-midnight), suggesting time-specific care demands.

XSL•FO
RenderX

Figure 2. Direct and indirect total task load (minutes) by hour for (A) registered nurses and (B) enrolled nurses.



For registered nurses, medication handling (88/834 min, 10%), procedures (46/834 min, 5%), and patient communication (46/834 min, 5%) were the largest contributors to direct care task load during the day (percentage of total task time), with no change at night (Table 2). Staff communication (134/834 min, 16%), data entry into the medical record system (117/834 min, 14%), and record viewing (104/834 min, 12%) were the largest contributors to indirect care task load. At night, data entry and record viewing remained as the top tasks, but time spent on staff communication decreased, whereas handover time increased from 5% (38/834 min) during the day to 8% (59/705 min) at night.

XSL•FO
RenderX

For enrolled nurses, procedures (85/764 min, 11%), assisting with toileting or diaper changes (57/764 min, 8%), and patient communication (57/764 min, 8%) were the largest contributors to direct care task load during the day (percentage of total task time), with no change at night. In addition, the proportion of task time spent on vital sign taking increased from 4% (33/764 min) during the day to 12% (82/711 min) at night. Staff communication (107/764 min, 14%), data entry into the medical record system (90/764 min, 12%), and record viewing (60/764 min, 8%) were the largest contributors to indirect care task load during the day. At nighttime, data entry and record viewing remained high contributors to task time, but staff communication decreased to 2% (11/711 min), whereas observation of patients in the ward became the third largest contributor at 7% (52/711 min).

XSL•FO
RenderX

Table . Average time taken (minutes) on tasks during the day or night for registered nurses (RNs) or enrolled nurses (ENs).

Task category, subcategory, and task codes	Time taken on tasks during a day shift (min), mean (SD)		Time taken on tasks during a night shift (min), mean (SD)	
	RN	EN	RN	EN
Direct care tasks				
Supporting activities of daily living				
Ambulation	2 (2.7)	10 (7.7)	3 (4.7)	7 (2.0
Bathing	13 (17.2)	30 (20.2)	2 (5.7)	7 (6.0)
Changing bedding	4 (5.2)	12 (7.7)	1 (1.8)	2 (1.5)
Diapering or assisting with toileting	12 (13.3)	57 (30.2)	33 (24.7)	103 (65.7)
Assisting with food and drink	9 (11.5)	19 (14.0)	3 (4.2)	18 (7.5)
Patient communication				
Nurse to patient	46 (22.6)	57 (26.7)	62 (41.4)	99 (112.2)
Meal ordering	0.4 (1.1)	4 (6.4)	1 (2.3)	2 (1.4)
Attending the patient call bell	1 (2.6)	2 (2.0)	2 (3.0)	4 (1.3)
Clinical care				
Vital sign taking	5 (6.4)	33 (12.4)	17 (12.9)	82 (23.2)
Laboratory samples	9 (13.1)	1 (2.1)	26 (20.1)	3 (5.8)
Other assessments and procedures	46 (24.7)	85 (25.6)	41 (22.8)	87 (16.5)
Medication handling	88 (44.3)	11 (13.3)	96 (34.7)	4 (3.3)
Patient comfort	4 (3.0)	14 (8.4)	5 (5.2)	4 (3.4)
Patient transfer (diagnostics)	5 (12.0)	8 (17.4)	0.3 (1.2)	5 (10.3)
Waste disposal (clinical and nonclinical waste)	3 (2.1)	9 (2.8)	3 (2.4)	10 (4.2)
Indirect care tasks				
Administrative				
Electronic medical record data entry	117 (37.2)	90 (38.6)	92 (32.3)	59 (23.1)
Record viewing (no data entry)	104 (41.9)	60 (43.8)	92 (39.3)	53 (14.2)
Nonclinical administrative tasks	24 (21.3)	17 (18.4)	9 (14.7)	7 (3.7)
Discharge or admission	19 (16.9)	4 (4.3)	8 (13.6)	3 (2.7)
Staff communication				
In-person staff member to staff member	134 (36.6)	107 (18.4)	46 (41.0)	11 (9.5)
Remote staff member to staff member	22 (11.2)	2 (1.4)	8 (6.6)	0 (0)
Remote staff member to caregiver	1 (1.8)	0 (0)	0.4 (1.0)	0 (0)
Handover	38 (30.5)	16 (15.1)	59 (23.9)	18 (21.4)
Team meetings	22 (14.6)	11 (16.4)	10 (7.1)	10 (14.7)
Staff education				
Staff education	36 (35.5)	4 (6.2)	3 (5.8)	0.4 (0.8)

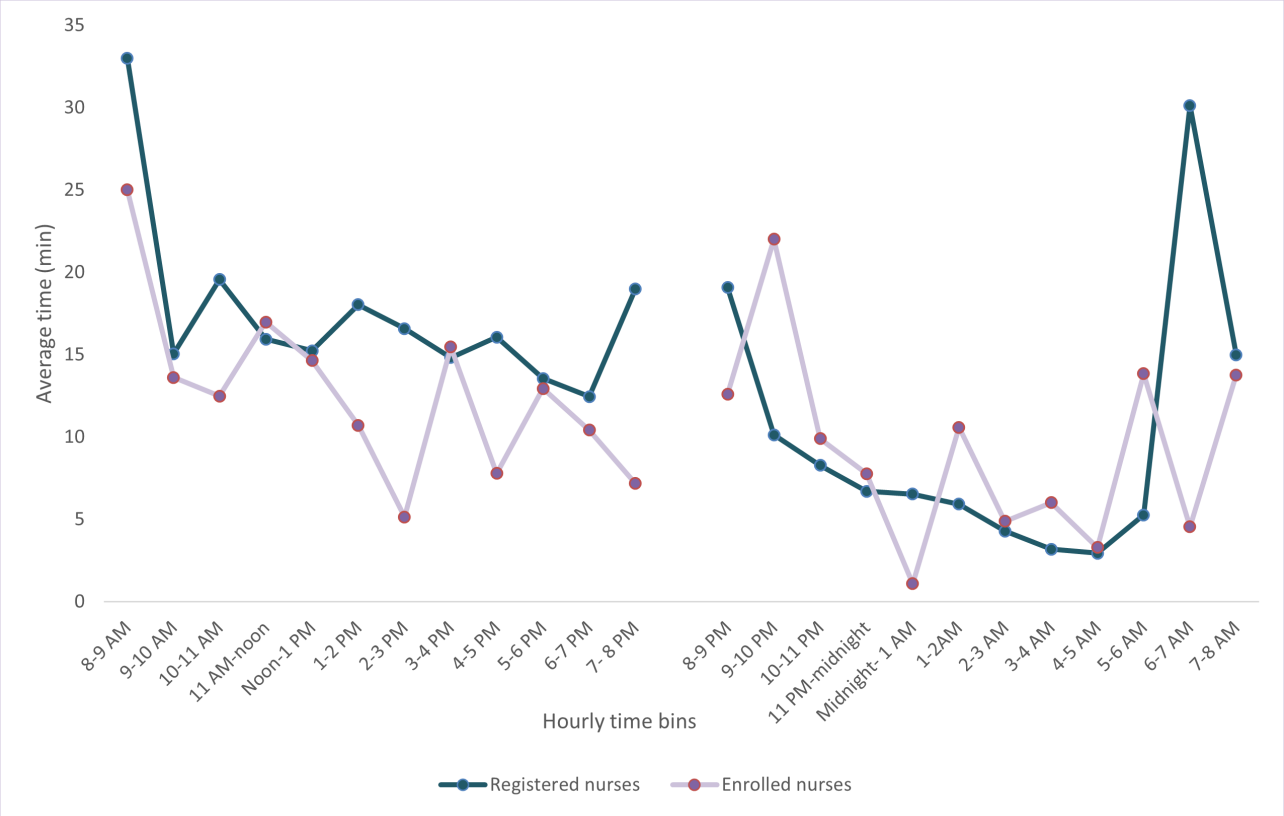
XSL•FO
RenderX

Task category, subcategory, and task codes	Time taken on tasks during a day shift (min), mean (SD)		Time taken on tasks during a night shift (min), mean (SD)	
	RN	EN	RN	EN
Ward observation				
Observing patients	6 (8.3)	8 (5.3)	41 (70.9)	52 (69.7)
Break time				
Break	64 (32.2)	93 (22.4)	40 (27.3)	61 (3.2)

Multitasking Load

During a 24-hour period, registered nurses typically spent 209 (SD 51.8) minutes multitasking in the day and 117 (SD 41.0) minutes at night. For enrolled nurses, 152 (SD 54.7) minutes were spent multitasking in the day and 110 (SD 75.9) minutes at night. The frequency of occurrence varied over the course of the day (Figure 3). During the daytime, the period with the highest multitasking load was between 8 AM and 9 AM for registered nurses (33, SD 14.6 min) and enrolled nurses (25, SD 11.7 min), and the lowest multitasking load was observed between 6 PM and 7 PM for registered nurses and between 2 PM and 3 PM for enrolled nurses. At nighttime, the period with the highest level of multitasking was 6 AM to 7 AM for registered nurses (30, SD 33.1 min) and 9 PM to 10 PM for enrolled nurses (22, SD 16.3 min).

Figure 3. Average time (minutes) spent multitasking in each hour for a 24-hour period by registered nurses (blue) and enrolled nurses (pink).



Discussion

Principal Findings

To address ongoing workforce challenges, there is a pressing need for role redesign and the careful application of technology to support nursing practice. This study provided empirical insights into how nurses in general medicine wards spend their time across a 24-hour period, including the distribution of direct and indirect care activities and the occurrence of multitasking. This study is part of a wider program of work to inform how virtual nursing roles can be designed and integrated into ward workflows by identifying which tasks may be amenable to remote support and during which times support may be most beneficial.

Our findings show that nurses spent only 30% (247/834 min) to 46% (352/764 min) of their time on direct clinical care; the remainder was dedicated to indirect care tasks such as documentation and coordination. Task loads frequently exceeded 60 minutes of activity per hour, indicating significant multitasking and overlapping responsibilities. Such patterns reflect recognized inefficiencies in inpatient workflows, where competing demands require nurses to switch rapidly between tasks, increasing cognitive load and potential for error. We also observed predictable peaks in indirect care duties—particularly around early-morning and evening shift transitions—where

XSL•FO
RenderX

administrative and communication tasks accumulated. In contrast, night shift patterns demonstrated marked variability, with periods of low activity followed by compressed workloads toward the end of the shift, when cumulative fatigue is likely to be highest. These specific patterns represent risk-prone periods within the nursing workday.

From a service design perspective, these findings suggest that virtual nursing support may be particularly beneficial when targeted to predictable high-risk periods, such as shift transitions and late-night hours, rather than implemented as a constant or uniformly distributed resource. Our data show that documentation and coordination activities cluster at these transition points, creating periods where clinical sensemaking and narrative continuity may be compressed or fragmented. While often treated as discrete tasks, documentation and handover are essential to a nurse's understanding of patient care and response over time. Thus, the design of virtual nursing roles should support clinical sensemaking by assisting with preparatory documentation, information synthesis, and record completion while preserving the primary nurse's responsibility for clinical interpretation, evaluation, and relational care. In this way, virtual nursing supports continuity of care rather than displacing the ground nurse's knowledge of the patient.

Our analysis showed that nurses frequently experienced task loads exceeding 1 hour of activity per observed hour, reflecting substantial multitasking and the fragmented nature of nursing work [10,11,14,15]. While some multitasking may improve efficiency, excessive cognitive load increases the likelihood of errors and reduces care quality, underscoring the need to reconsider how workload is organized [19,20]. By mapping when peaks in documentation, coordination, and multitasking occur, we highlight concrete opportunities for virtual nursing integration. For example, indirect care tasks that cluster around shift handover—such as preparing documentation, updating EMR entries, and coordinating care plans—could be supported by virtual nurses operating remotely. Importantly, the high degree of task concurrency observed suggests that virtual nursing should prioritize offloading cognitively competing tasks (eg, documentation and coordination) rather than introducing additional parallel interactions that could further fragment attention.

Night shifts exhibited greater variability, particularly for enrolled nurses, with pronounced peaks and a late-shift surge, when fatigue is likely to impair performance and elevate error risk [21,22]. Prior studies have raised similar concerns, noting that staffing levels during night shifts often fail to meet patient needs, leading to increased stress and task compression or even unresolved tasks [23-25]. Although interventions such as scheduled napping and optimized lighting can mitigate fatigue [26-28], they do not address the underlying workload imbalance. These patterns point to specific opportunities for virtual nursing support, such as documentation, discharge preparation, remote monitoring, and assisting junior staff, to relieve peak pressures and enhance workflow continuity [29-31]. Aligning virtual nursing activity with these high-pressure windows offers a targeted approach to smoothing workflow peaks, reducing multitasking burden, and supporting safer, more sustainable practice.

Analysis of the types of nursing activities revealed that a substantial proportion of time was spent on indirect care tasks, particularly EMR documentation, by both registered and enrolled nurses. This aligns with extensive literature showing that administrative workload, especially documentation, is a major contributor to reduced direct care time, job dissatisfaction, and burnout [10,14,20,32,33]. From the perspective of the theory of human caring by Watson [13], excessive administrative demands diminish nurses' capacity to engage in the relational, patient-centered interactions that underpin caring-healing practices, such as being fully present, developing trust, and supporting patients' emotional and existential needs. Our findings reinforce these concerns by showing that administrative peaks occur at predictable times, particularly during shift changes, thereby reducing opportunities for therapeutic engagement. Rather than treating documentation reduction as a generic goal, our findings indicate that documentation support is most critical during temporally concentrated peaks—particularly at shift transitions, where administrative demands can displace opportunities for patient-facing care. These patterns highlight clear opportunities for workflow redesign, including improved handover protocols, EMR-integrated automation tools, and shifting routine documentation to virtual care teams or dedicated virtual administrative assistants. Early evaluations of virtual nursing models show promise in reducing documentation burden, improving workflow efficiency, and enhancing staff satisfaction [34-37]. Our data help identify where such support would be most impactful, ultimately enabling bedside nurses to devote greater attention to the caring practices central to the nursing profession.

Taken together, these insights offer timely evidence to guide the development of virtual nursing roles within inpatient workflows. Virtual nurses, equipped with real-time access to EMRs, vital sign dashboards, secure messaging, and telepresence tools, could assume a range of indirect care tasks—documentation, discharge coordination, remote education delivery, and virtual rounding [29-31,35-38]. Importantly, we should tailor support by staff role: for registered nurses, reducing documentation and coordination workload allows for greater focus on complex clinical care, whereas for enrolled nurses, virtual teams could assist with supervisory or routine administrative tasks [29-31,35-38]. Over time, this infrastructure also provides a foundation for automating selected processes with artificial intelligence–enabled tools, contributing to more scalable and sustainable workforce models.

At the same time, it is essential that the ethical considerations of virtual nursing are taken into account. Redistributing nursing work to virtual roles raises important ethical considerations related to professional responsibility, accountability, and the nurse-patient relationship. While indirect care activities may be performed remotely, clinical judgment, evaluation of patient responses, and ownership of care outcomes must remain with the primary nurse responsible for the patient. Therefore, our findings highlight the importance of designing virtual nursing models that reduce cognitive and administrative burden without fragmenting accountability or distancing nurses from patient care, ensuring that virtual support enhances—rather than

XSL•FO
RenderX

replaces—the ethical commitments embedded in bedside nursing practice.

Limitations

Our direct observations allowed us to gather unbiased data on nursing processes and report objectively on time use, overcoming the limitations of qualitative interview or survey data, which may be prone to bias. Our observation checklist was comprehensive, allowing for a fine-grained analysis of the task load. However, limitations must be acknowledged. We cannot ignore the Hawthorne effect, where staff may alter their behavior while being observed. We attempted to minimize this by briefing nursing teams and taking multiple sampling points over a 3-month period. There may also have been issues with observer variability. To minimize this risk, we conducted pilot work and training sessions with each observer to standardize the process. Observers were also encouraged to discuss issues and ask for clarifications during regular team meetings. Finally, our sampling took place during a 3-month period, which may not accurately reflect the average task load across the year.

Conclusions

This time-and-motion study documents the substantial and complex workload experienced by inpatient nursing staff over a 24-hour period across different nurse roles. The data reveal consistently high task demands, frequent multitasking, and pronounced peaks and troughs in workload throughout the day and night. These patterns suggest significant cognitive strain, with potential consequences for staff well-being and patient safety. These insights can inform the design of virtual nursing models or other technology-enabled workforce innovations. Future research should test and evaluate these models through staged implementation and outcome assessment, ensuring that virtual and hybrid care teams are safe, effective, and responsive to the realities of inpatient workflows.

Acknowledgments

The authors would like to acknowledge the support of Yang Qiuping, Rose Dimaranan, Norhasidah Binte Wahid, Tim Hwee Mien, Dina Gita Cintanati, Eis Winangsih, Kamilah Qistina, Lim Zhen Tian, Eliona Leandra Norruega, Mihir Bundele, and Nikita Mandyam. The software Grammarly (Superhuman Platform Inc) was used to improve sentence structure and grammar.

Funding

This research is supported by the Singapore Ministry of Health's National Medical Research Council RIE2025 Centre Grant program (grant NMRC/CG3/003/2022-AH/MOH-0010130-00).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JS, EHHC, SYP, AWY
Formal analysis: JS, MKC, DSO
Funding acquisition: AWY, EHHC
Investigation: JS, EHHC, SYP, JXXL, CK, AA, JMA, HWL, AB, EC, CY, EH, AWY
Methodology: JS, EHHC, SYP, JL, PC, AWY
Project administration: AWY, EHHC
Resources: SYT, JL, AM, AWY
Supervision: JS, AWY, EHHC
Validation: JS, EHHC, SYP
Writing—review and editing: JS, EHHC, SYP, JL, PC, MKC, DSO, JXXL, CK, AA, JMA, HWL, AB, EC, CY, EH, AM, AWY

Conflicts of Interest

None declared.

References

1. Ageing and health. World Health Organization. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2024-01-18]
2. Azzopardi-Muscat N, Zapata T, Kluge H. Moving from health workforce crisis to health workforce success: the time to act is now. Lancet Reg Health Eur 2023 Dec;35:100765. [doi: 10.1016/j.lanepe.2023.100765] [Medline: 38115956]
3. Pressley C, Newton D, Garside J, Simkhada P, Simkhada B. Global migration and factors that support acculturation and retention of international nurses: a systematic review. Int J Nurs Stud Adv 2022 Dec;4:100083. [doi: 10.1016/j.ijnsa.2022.100083] [Medline: 38745611]

XSL•FO
RenderX

4. Nigam JA, Barker RM, Cunningham TR, Swanson NG, Chosewood LC. Vital signs: health worker-perceived working conditions and symptoms of poor mental health - quality of worklife survey, United States, 2018-2022. MMWR Morb Mortal Wkly Rep 2023 Nov 3;72(44):1197-1205. [doi: 10.15585/mmwr.mm7244e1] [Medline: 37917563]
5. Brown S, Guillergan J, Beedle E, et al. Best practices in supporting inpatient communication with technology during visitor restrictions: an integrative review. Comput Inform Nurs 2025 Jan 1;43(1):e01200. [doi: 10.1097/CIN.0000000000001200] [Medline: 39436285]
6. Smith JM, Pearson KK, Roberson AE. Interface of clinical nurse specialist practice and healthcare technology. Clin Nurse Spec 2023;37(4):169-176. [doi: 10.1097/NUR.0000000000000755] [Medline: 37410561]
7. Roberson AE, Carlson M, Kohler CM, Harris PA, Volkmann CL. Initiating virtual nursing in general inpatient care. Am J Nurs 2023 Jun 1;123(6):48-54. [doi: 10.1097/01.NAJ.0000938736.42266.5e] [Medline: 37233140]
8. Perpetua Z, Seitz S, Schunk J, et al. Virtual discharge: enhancing and optimizing care efficiency for the bedside nurse. J Nurs Care Qual 2023;38(3):234-242. [doi: 10.1097/NCQ.0000000000000689] [Medline: 36693624]
9. Pressley C, Garside J. Safeguarding the retention of nurses: a systematic review on determinants of nurse's intentions to stay. Nurs Open 2023 May;10(5):2842-2858. [doi: 10.1002/nop2.1588] [Medline: 36646646]
10. Michel O, Garcia Manjon AJ, Pasquier J, Ortoleva Bucher C. How do nurses spend their time? A time and motion analysis of nursing activities in an internal medicine unit. J Adv Nurs 2021 Nov;77(11):4459-4470. [doi: 10.1111/jan.14935] [Medline: 34133039]
11. Kim Y, Lee MJ, Choi M, Cho E, Ryu GW. Exploring nurses' multitasking in clinical settings using a multimethod study. Sci Rep 2023;13(1):5704. [doi: 10.1038/s41598-023-32350-9]
12. Douglas HE, Raban MZ, Walter SR, Westbrook JI. Improving our understanding of multi-tasking in healthcare: drawing together the cognitive psychology and healthcare literature. Appl Ergon 2017 Mar;59(Pt A):45-55. [doi: 10.1016/j.apergo.2016.08.021] [Medline: 27890158]
13. Watson J. Human Caring Science: A Theory of Nursing, 2nd edition: Jones & Bartlett Learning; 2012.
14. Kramer S, Raymond MJ, Hunter P, et al. Understanding the workflow of nurses in acute and subacute medical wards: a time and motion study. J Clin Nurs 2023 Nov;32(21-22):7773-7782. [doi: 10.1111/jocn.16835] [Medline: 37489643]
15. Yen PY, Kelley M, Lopetegui M, et al. Understanding and visualizing multitasking and task switching activities: a time motion study to capture nursing workflow. AMIA Annu Symp Proc 2016;2016:1264-1273. [Medline: 28269924]
16. Zheng K, Guo MH, Hanauer DA. Using the time and motion method to study clinical work processes and workflow: methodological inconsistencies and a call for standardized research. J Am Med Inform Assoc 2011;18(5):704-710. [doi: 10.1136/amiajnl-2011-000083] [Medline: 21527407]
17. Lopetegui M, Yen PY, Lai AM, Embi PJ, Payne PR. Time Capture Tool (TimeCaT): development of a comprehensive application to support data capture for time motion studies. AMIA Annu Symp Proc 2012;2012:596-605. [Medline: 23304332]
18. McHugh ML. Interrater reliability: the Kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
19. Weigl M, Müller A, Holland S, Wedel S, Woloshynowych M. Work conditions, mental workload and patient care quality: a multisource study in the emergency department. BMJ Qual Saf 2016 Jul;25(7):499-508. [doi: 10.1136/bmjqs-2014-003744] [Medline: 26350066]
20. Lavander P, Meriläinen M, Turkki L. Working time use and division of labour among nurses and health-care workers in hospitals - a systematic review. J Nurs Manag 2016 Nov;24(8):1027-1040. [doi: 10.1111/jonm.12423] [Medline: 27581093]
21. Thompson BJ, Stock MS, Banuelas VK. Effects of accumulating work shifts on performance-based fatigue using multiple strength measurements in day and night shift nurses and aides. Hum Factors 2017 May;59(3):346-356. [doi: 10.1177/0018720816677814] [Medline: 27941163]
22. Farag A, Gallagher J, Carr L. Examining the relationship between nurse fatigue, alertness, and medication errors. West J Nurs Res 2024 Apr;46(4):288-295. [doi: 10.1177/01939459241236631] [Medline: 38454783]
23. Griffiths P, Recio-Saucedo A, Dall'Ora C, et al. The association between nurse staffing and omissions in nursing care: a systematic review. J Adv Nurs 2018 Jul;74(7):1474-1487. [doi: 10.1111/jan.13564] [Medline: 29517813]
24. Lasater KB, Aiken LH, Sloane DM, et al. Chronic hospital nurse understaffing meets COVID-19: an observational study. BMJ Qual Saf 2021 Aug;30(8):639-647. [doi: 10.1136/bmjqs-2020-011512] [Medline: 32817399]
25. Ball JE, Murrells T, Rafferty AM, Morrow E, Griffiths P. "Care left undone" during nursing shifts: associations with workload and perceived quality of care. BMJ Qual Saf 2014 Feb;23(2):116-125. [doi: 10.1136/bmjqs-2012-001767] [Medline: 23898215]
26. Cyr M, Artenie DZ, Al Bikaii A, Lee V, Raz A, Olson JA. An evening light intervention reduces fatigue and errors during night shifts: a randomized controlled trial. Sleep Health 2023 Jun;9(3):373-380. [doi: 10.1016/j.sleh.2023.02.004] [Medline: 37080863]
27. Han K, Hwang H, Lim E, et al. Scheduled naps improve drowsiness and quality of nursing care among 12-hour shift nurses. Int J Environ Res Public Health 2021 Jan 20;18(3):891. [doi: 10.3390/ijerph18030891] [Medline: 33498593]

XSL•FO
RenderX

28. Wu CJ, Huang TY, Ou SF, Shiea JT, Lee BO. Effects of lighting interventions to improve sleepiness in night-shift workers: a systematic review and meta-analysis. Healthcare (Basel) 2022 Jul 26;10(8):1390. [doi: 10.3390/healthcare10081390] [Medline: 35893212]
29. Ramirez H, Velastin SA, Meza I, Fabregas E, Makris D, Farias G. Fall detection and activity recognition using human skeleton features. IEEE Access 2021;9:33532-33542. [doi: 10.1109/ACCESS.2021.3061626]
30. Zubrinic M, Vrbanic L, Keshavjee S. Remote telemonitoring is associated with improved patient safety and decreased workload of nurses. JTCVS Open 2023 Dec;16:493-497. [doi: 10.1016/j.xjon.2023.09.014] [Medline: 38204626]
31. Cho I, Cho J, Hong JH, Choe WS, Shin H. Utilizing standardized nursing terminologies in implementing an AI-powered fall-prevention tool to improve patient outcomes: a multihospital study. J Am Med Inform Assoc 2023 Oct 19;30(11):1826-1836. [doi: 10.1093/jamia/ocad145] [Medline: 37507147]
32. Gesner E, Dykes PC, Zhang L, Gazarian P. Documentation burden in nursing and its role in clinician burnout syndrome. Appl Clin Inform 2022 Oct;13(5):983-990. [doi: 10.1055/s-0042-1757157] [Medline: 36261113]
33. Yan Q, Jiang Z, Harbin Z, Tolbert PH, Davies MG. Exploring the relationship between electronic health records and provider burnout: a systematic review. J Am Med Inform Assoc 2021 Apr 23;28(5):1009-1021. [doi: 10.1093/jamia/ocab009] [Medline: 33659988]
34. Roberson AE, Martinez HR, Holligan SA, Carlson M. An evolving virtual nursing model: enhancing care delivery through advanced workflow integration. Nurs Outlook 2025;73(3):102421. [doi: 10.1016/j.outlook.2025.102421] [Medline: 40252518]
35. Tudorache O, Kenemer JB, Pruiett J, et al. Implementing virtual nursing in health care: an evaluation of effectiveness and sustainability. Presented at: 2022 IEEE International Conference on Digital Health (ICDH); Jul 10-16, 2022; Barcelona, Spain p. 129-131. [doi: 10.1109/ICDH55609.2022.00028]
36. Ransford J, Tidwell T, Johnson L, Gitney N, Morgan A, Hauch R. Implementing a virtual discharge nurse pilot: utilizing the pathway to Excellence® framework. J Nurs Adm 2024 Nov 1;54(11):605-611. [doi: 10.1097/NNA.0000000000001498] [Medline: 39475890]
37. Tyransky CE, Paulus K, Langmead E, et al. Integration of virtual technology and artificial intelligence improves satisfaction, patient safety, and nursing workforce efficiency. J Nurs Care Qual 2025;40(3):193-201. [doi: 10.1097/NCQ.0000000000000842]
38. Zhang L, Gallagher R, Du H, et al. Evaluate the effect of virtual nurse-guided discharge education app on disease knowledge and symptom response in patients following coronary events. Int J Med Inform 2025 Apr;196:105818. [doi: 10.1016/j.ijmedinf.2025.105818] [Medline: 39919533]

Abbreviations

EMR: electronic medical record

Edited by M Kleib; submitted 17.Nov.2025; peer-reviewed by B Cross, M Ozkaynak; revised version received 06.Feb.2026; accepted 01.Mar.2026; published 08.May.2026.

Please cite as:
Sumner J, Chew EHH, Peck SY, Lam J, Chong P, Ching MK, Ong DS, Lau JXX, Keck C, Ang A, Mohamed Ali J, Lim HW, Bundele A, Chen E, Yap C, Ho E, Mukhopadhyay A, Yip AW
Implications for Virtual Nursing Role Development in Acute Nursing Care: 24-Hour Time-and-Motion Study
JMIR Nursing 2026;9:e87930
URL: https://nursing.jmir.org/2026/1/e87930
doi:10.2196/87930

© Jennifer Sumner, Emily Hwee Hoon Chew, Sin Yee Peck, Janet Lam, Pauline Chong, Mun Kwong Ching, De Sheng Ong, Jacqueline Xiao Xuan Lau, Camille Keck, Abigail Ang, Jaminah Mohamed Ali, Hui Wen Lim, Anjali Bundele, Elya Chen, Carol Yap, Evelyn Ho, Amartya Mukhopadhyay, Alexander Wenjun Yip. Originally published in JMIR Nursing (https://nursing.jmir.org), 8.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Validation and Feasibility of a Consumer-Grade Wearable Sleep Monitoring Device and an Exploratory Evaluation of Sleep Characteristics and Associated Factors Among Nurses: Observational Cohort Study

Siew Hoon Lim, PhD; Fazila Aloweni, MSc; Phyllis Ong, BSc; Shin Yuh Ang, MBA

Division of Nursing, Singapore General Hospital, Singhealth Tower Level 15, 10 Hospital Boulevard, Singapore, Singapore

Corresponding Author:
Siew Hoon Lim, PhD
Division of Nursing, Singapore General Hospital, Singhealth Tower Level 15, 10 Hospital Boulevard, Singapore, Singapore

Abstract

Background: Shift work disorder and insufficient sleep are prevalent among nurses, leading to fatigue, reduced well-being, and potential safety concerns. The increasing use of wearable sleep-tracking devices presents an opportunity to evaluate nurses' sleep quality objectively.

Objective: The primary objective was to evaluate the feasibility of wearable-based sleep monitoring and to obtain preliminary evidence of agreement with validated actigraphy among nurses. The secondary objective was to describe nurses' sleep characteristics and to examine exploratory associations between sociodemographic characteristics, shift patterns, sleep hygiene, and sleep parameters.

Methods: A 2-phase feasibility observational cohort study was conducted in a tertiary hospital in Singapore. In phase 1, 5 nurses concurrently wore a consumer-grade, wrist-worn Apple Watch Series 10 and a validated actigraph (GENEActiv; ActivInsights Ltd) for 2 weeks. Preliminary agreement between the Apple Watch and GENEActiv was examined using intraclass correlation coefficients. Feasibility was determined through wear-time compliance and data completeness. In phase 2, 50 nurses working rotating or single shifts completed demographic and work-related questionnaires and the Sleep Hygiene Index. Multiple linear regression analyses were performed to examine exploratory associations between selected covariates and sleep parameters, with adjustment for age, sex, BMI, parental status, workplace, total length of service, and sleep hygiene.

Results: The Apple Watch showed preliminary evidence of agreement with GENEActiv for total sleep time, in-bed wake time, and sleep efficiency (intraclass correlation coefficients of 0.95, 0.72, and 0.69, respectively), with high wear compliance and minimal missing data, supporting its feasibility for sleep monitoring. Mean total sleep time was 381 (SD 55) minutes, and mean sleep efficiency was 94.77% (SD 4.11%). Shift nurses reported poorer sleep hygiene than nonshift nurses; however, shift work status was not independently associated with sleep outcomes after adjustment. Higher BMI was associated with shorter total sleep time (unstandardized coefficient B=−3.76 min/kg/m²; P=.01), reduced rapid eye movement sleep (unstandardized coefficient B=−1.18 min; P=.03), shorter core sleep (unstandardized coefficient B=−2.72 min; P=.02), and reduced time in bed (unstandardized coefficient B=−4.05 min; P<.01=.008). Age was negatively associated with deep sleep duration, with older age associated with less deep sleep (unstandardized coefficient B=−1.00 min/y; P=.003).

Conclusions: Apple Watch–based monitoring was feasible and showed preliminary agreement with actigraphy. Nurses slept less than recommended, and BMI and age were associated with sleep outcomes in exploratory analyses. These findings support larger studies and workplace strategies to improve sleep opportunity and healthy sleep behaviors.

(JMIR Nursing 2026;9:e96912) doi:10.2196/96912

KEYWORDS

sleep; nurses; shift work; acute care; wearables

Introduction

Background

Insufficient sleep and sleep disorders are commonly identified health challenges among health care workers who perform shift work [1]. Shift work alters circadian rhythm, causes sleep disturbances, and affects physical and mental health, leading to poor work performance and affecting patient safety [2,3]. This is especially so for nurses who take on night shifts, where the irregularities between their internal circadian rhythms and the work requirements result in alterations in their sleep-wake cycles and light-dark cycles, which lead to a circadian rhythm

XSL•FO
RenderX

sleep-wake disorder known as shift work disorder [3]. Shift work disorder is characterized by chronic insomnia whereby individuals experience difficulties in falling asleep and staying asleep, as well as possible excessive sleepiness [4].

Prolonged irregular working hours increase nurses' vulnerability to the development of sleep disorders, including increased stress and anxiety, depressed mood, and reduced overall mental health [5]. The decrease in sleep quality among nurses has also been reported to increase sick leave days, lower job satisfaction, and burnout [6-8]. Studies have also highlighted possible associations between working night shifts and weight gain, obesity, breast cancer, cardiovascular disease, and metabolic disorders, including type 2 diabetes [9,10]. Regardless of the shift worked, nurses generally report poor sleep quality. Previous studies found that up to 63% of nurses reported poor sleep quality and that 68% highlighted insomnia-related concerns [11].

Sleep Patterns of Nurses on Shift Work

Nurses largely have no fixed work schedule; they have to manage and adjust their schedules according to their shifts. There is no fixed forward-cyclic rostering schedule. The sleep quality of nurses working on different shifts was evaluated in previous studies [12]. Nurses on the night shift generally had lower total sleep time, higher wake after sleep onset, and lower sleep efficiency during their work days when compared to those who worked day and evening shifts [13]. Nurses with rotational shifts in a counterclockwise rotation of more than 50% of shifts (night to evening to day shifts) reported poor sleep quality as well [12]. Whereas nurses who worked evening rotating shifts reported longer total sleep time when compared to those working day or night rotating shifts, as well as fixed day shifts. Furthermore, nurses on backward rotating–shift schedules also reported higher levels of sleep disturbances and less balanced sleep duration as compared to those on forward rotating–shift schedules [14]. Understanding sleep parameters of nurses working on different shifts is crucial to facilitate the arrangement of work shifts in the best possible manner to lower the impact of shift work on their sleep quality.

Sleep Hygiene Behaviors

Sleep hygiene primarily involves individual behaviors that facilitate sleep and minimize behaviors that prevent restful sleep [15]. Good sleep habits include avoiding late afternoon naps; minimizing caffeine, alcohol, and tobacco intake; reducing emotionally demanding activities during bedtime; ensuring a quiet and comfortable bedroom environment; and keeping to the same daily sleep schedule [16]. Inadequate sleep hygiene is defined in the International Classification of Sleep Disorders [17] as a type of sleep disorder due to individual activities that interfere with quality sleep and result in reduced daytime alertness.

The review of the literature found few studies exploring nurses' sleep hygiene behaviors. Recent studies have reported that poor sleep hygiene is associated with poor sleep quality and even with shift work disorder [18,19]. Various common poor sleep hygiene behaviors were identified among nurses, including going to bed at different times, waking at different times, going to bed feeling stressed or angry, as well as planning or worrying when in bed [18]. Poor sleep hygiene practices among nurses were also found to have higher insomnia and anxiety issues. Understanding sleep hygiene practices among nurses working in the acute care setting is crucial to increasing awareness about the importance of sleep hygiene and developing sleep hygiene interventions to improve nurses' sleep hygiene and sleep quality.

Sleep Monitoring Devices

The sensitivity and accuracy of actigraphy have been determined and are commonly used in clinical settings to evaluate sleep quantity and quality [20]. It is a small wristwatch-like device that is convenient for monitoring movements when nurses are at work. Actigraphy-based personal sleep monitoring devices help users establish an objective knowledge about their sleep patterns, including sleep quality, rhythms, and duration. Although actigraphy has been identified as the standard assessment for mobile sleep monitoring, recent rapid development of personal consumer devices has highlighted the possible validity of feasible alternative approaches to actigraphy in the measurement of sleep [21]. One of the advantages of using personal consumer devices is the lower cost, with accompanying additional features, when compared to traditional actigraphy devices. Other favorable features of personal consumer sleep devices include wireless or Bluetooth connections enabling close to real-time data processing and also contain other sensor measurements, including heart rate and respiratory rate, which may provide additional data for sleep algorithms.

Previous studies have evaluated the accuracy of Apple Watch, one of the most widely used consumer-grade sleep tracking devices [22,23]. Roomkham et al [23] compared the sleep parameters recorded on Apple Watch with Philips Actiwatch Spectrum Pro. They found Apple Watch to have overall accuracy (97%) and sensitivity (99%) in detecting actigraphy-defined sleep, and adequate specificity (79%) in detecting actigraphy-defined wakefulness. Apple Watch, on average, only overestimated total sleep time by 6.31 minutes and underestimated wake after sleep onset by 5.74 minutes [23]. Although previous findings have demonstrated that Apple Watch can provide reasonably accurate estimates of total sleep time in general adult populations, evidence remains limited regarding its performance in health care workers, particularly nurses who experience irregular schedules and fragmented sleep. Most validation studies have been conducted under controlled environments, and it is unclear whether consumer wearables perform similarly in occupational settings characterized by variable sleep timing and shift-work demands. Wearable-derived sleep metrics provide objective estimates of sleep timing, duration, continuity, and sleep-stage distribution; however, these data do not capture the behavioral and environmental circumstances that may influence observed sleep patterns. Assessing sleep hygiene alongside objective sleep measures is, therefore, warranted, as it allows consideration of modifiable practices such as irregular sleep schedules, caffeine consumption, presleep cognitive or emotional arousal, and bedroom conditions that may influence wearable-recorded sleep outcomes. The integration of objective sleep monitoring with sleep hygiene assessment may offer a more contextualized

XSL•FO
RenderX

understanding of nurses' sleep health and identify behavioral targets for future sleep-focused interventions.

To address these gaps, this study aimed to evaluate the feasibility of using a commercial sleep monitoring device and to obtain preliminary evidence of agreement with a validated research-grade actigraphy device among nurses working rotating shifts. Agreement was examined by comparing total sleep time, in-bed wake time, and sleep efficiency estimates from the Apple Watch and GENEActiv. Feasibility was examined through device wear compliance and the completeness of nightly data. The secondary aim was to examine how demographic characteristics, shift-work patterns, and subjective sleep hygiene behaviors relate to objectively measured sleep parameters among nurses working rotating shifts.

Methods

Design

A feasibility observational cohort study design was adopted.

Sample and Setting

A convenience sampling approach was used to recruit full-time registered nurses who are working in an acute care tertiary institution in Singapore from May 2025 to December 2025. The exclusion criteria included a history of psychiatric or neurological conditions, part-time employment, or long-term leave from work during the study period.

Sample size was not determined by a formal power analysis because this study was designed as a feasibility study [24]. For the first phase of the study, 5 participants were recruited. The focus of this phase was to conduct a local validation of Apple Watch against actigraphy, rather than to generate statistically powered estimates. Prior findings have demonstrated the relative accuracy of Apple Watch in estimating sleep and wake parameters when compared with polysomnography and actigraphy, showing moderate-to-high agreement for total sleep time and sleep efficiency [25]. In the second phase of the study, another 50 participants were recruited. This sample size was consistent with methodological recommendations for feasibility studies that emphasize obtaining preliminary estimates of outcome variability and adherence rates rather than hypothesis testing [26]. A sample of 55 participants would provide sufficient information on recruitment, attrition, and data completeness to inform the design of a future larger-scale study, while ensuring diversity of participants across demographics and shift patterns within the studied nursing population. The sample was, therefore, considered appropriate for the feasibility objectives of this study, including recruitment, device wear compliance, data completeness, and preliminary estimation of outcome variability, but was not intended to support definitive adjusted comparisons or multivariable regression modeling.

Data Collection

Recruitment and Baseline Data Collection

A study invitation email was sent to all nurses at the institution to solicit their participation. The email included the study details and eligibility criteria. Weekly email reminders were sent to encourage nurses to take part in the study. Written informed consent was obtained from interested participants who approached the study team.

Upon obtaining consent, baseline data, including sociodemographic information (age, sex, race, BMI, marital status, number and ages of children, caregiving for an older family member, educational qualifications, and smoking status), were collected. Work-related information was also obtained, including shift patterns (single-shift duties included day, night, or office hours; rotating-shift duties included 2- or 3-rotation schedules comprising day, evening, and night shifts), clinical areas of work, years of experience in the current work location, job grade, and total length of service. Feasibility was assessed through device wear compliance and the extent of missing sleep data.

Measurement of Sleep Time and Other Vital Signs

For the first phase of the study, 5 nurses wore an Apple Watch Series 10 and a GENEActiv actigraph device (ActivInsights Ltd) during sleep for 2 weeks. All Apple Watch devices, paired iPhones (Apple), and GENEActiv devices used in this study were department-owned equipment designated for research purposes only. The Apple Watch and GENEActiv devices were provided to participants only for the duration of the monitoring period and were returned to the study team after data collection. The Apple Watch Series 10 was selected as the consumer wearable device for this feasibility study because Apple Watch is widely used, is familiar to many users, and is supported by prior validation evidence for estimating sleep duration [23]. Its wrist-worn format, automated sleep tracking, and integration of movement and physiological sensor data also made it a practical option for short-term sleep monitoring among nurses working irregular schedules. Participants were briefed by the study team to manually activate "Sleep Mode" on the Apple Watch before going to sleep and to deactivate it upon waking. For GENEActiv, participants were not required to activate or deactivate the device; sleep start and end times were derived using participants' reported sleep and wake times to define the sleep period for analysis. Participants were reminded during the study briefing to record their sleep and wake times accurately throughout the monitoring period, and daily reminders were provided by the study team to prompt participants to document their sleep and wake times. GENEActiv has been validated against polysomnography and demonstrated reliable performance in detecting sleep and wake patterns (intraclass correlation coefficients [ICCs] of 0.79-0.85) for sleep parameters [27]. Total sleep time, in-bed wake time, and sleep efficiency were compared between GENEActiv and Apple Watch. Apple Watch data were synchronized to the paired iPhone and retrieved from the Apple Health dashboard after each participant completed the monitoring period [28]. GENEActiv raw accelerometer data were extracted in .bin format, with Actiwatch Data files or comma-separated values files generated using ActivInsights software where required [29]. The sleep report was generated using the Actigraphy sleep toolkit [30]. Sleep data were extracted from the Actigraphy sleep report after each participant completed monitoring. All extracted sleep data were reviewed by the principal investigator for completeness and consistency before analysis, while the principal investigator remained blinded to participants' shift roster.

XSL•FO
RenderX

In the second phase of the study, additional sleep metrics from the Apple Watch were collected, including time in bed, which was measured from sleep start to sleep end; in-bed wake time; and amounts of rapid eye movement (REM), light (core), and deep sleep during the time spent in bed. Sleep efficiency was manually calculated as the percentage of time spent asleep while in bed, defined as the total sleep time divided by the time in bed and multiplied by 100.

Sleep Hygiene

The Sleep Hygiene Index was used to evaluate the sleep hygiene behavior of the participants. It was administered at baseline through an online survey before the start of sleep monitoring. The Sleep Hygiene Index is a self-administered 13 items designed to assess environmental and behavioral variables that could disrupt sleep prior to bedtime, including inconsistent bedtimes, feelings of stress, anger, or worry, engaging in stimulating activities in bed such as watching television, consuming caffeine, or exercising right before bedtime, as well as the comfort level of the bedroom environment [31]. Participants were asked to report on the frequency of engagement in the specific behavior (0="never," 1="rarely," 2="sometimes," 3="frequently," and 4="always"). Total scores ranged from 0 to 52, with a higher score representing more maladaptive sleep hygiene practices. The Sleep Hygiene Index scores were presented as percentile rank, with scores above 50 indicating more disruptive sleep behaviors than average, to the extent that they would be more likely to result in poor sleep quality [31]. Scores 35 and above are considered poor sleep hygiene, 27 to 34 are considered normal, and below 26 are considered good sleep hygiene status. The Sleep Hygiene Index reported adequate reliability and validity [31]. The items from the Sleep Hygiene Index were developed based on the diagnostic criteria for insufficient sleep hygiene according to the International Classification of Sleep Disorders by the American Sleep Disorders Association [4]. The Sleep Hygiene Index reported satisfactory Cronbach α (0.66) and good test-retest reliability (r=0.71) [31].

Ethical Considerations

Ethical approval was obtained from the SingHealth Centralised Institutional Review Board, Singapore (ECOS Ref: 2024 - 3932). A participant information sheet detailing the study purpose, procedures, and voluntary nature was provided. Written informed consent was obtained from each participant before participation. Participants' privacy was protected, and the confidentiality of their identities and study data were maintained throughout the study. Participants received reimbursement for the time spent participating in the study.

Data Analysis

Data were entered and analyzed using SPSS version 25.0 (SPSS Inc).

The normality of continuous variables was examined using Shapiro-Wilk tests. Descriptive statistics were computed to summarize participants' demographic, occupational, and behavioral characteristics. Continuous variables were reported as means with SDs, while categorical variables were summarized as frequencies and percentages.

In this feasibility study, preliminary agreement between Apple Watch and GENEActiv was assessed using the ICC for total sleep time, in-bed wake time, and sleep efficiency. Given the small validation sample (n=5), these analyses were intended to provide preliminary evidence of agreement rather than to establish measurement accuracy. ICC values below 0.50 were considered poor, 0.50 to 0.74 moderate, 0.75 to 0.90 good, and above 0.90 excellent [32]. Total sleep time, in-bed wake time, and sleep efficiency were compared directly between devices without epoch alignment [33].

Feasibility was determined by calculating device wear compliance and the completeness of nightly sleep data over the 14-night monitoring period, with higher percentages indicating better acceptability and integration of the device into participants' routines. Wear compliance was expressed as the percentage of nights in which the device was worn. Completeness of nightly data was calculated as the percentage of nights that yielded valid sleep recordings out of the total number of nights. A valid sleep recording was defined as a night in which the device was worn during the intended sleep period and produced an interpretable sleep output with corresponding sleep start and end times. For phase 1 device-comparison analyses, a night was considered valid only when both Apple Watch and GENEActiv had usable data for the same reported sleep episode. Nights with partial recordings, implausible or incomplete sleep or wake entries, nonwear, incorrect wear, synchronization failure, or no recorded sleep output were treated as invalid and excluded from the analysis. Missing data were categorized as nonwear, incomplete or invalid sleep or wake entries, device synchronization failure, or no recorded output despite device wear.

Associations between demographic, occupational, and behavioral variables (age, sex, BMI, shift duties, work unit, parental status, total length of service, and Sleep Hygiene Index) and sleep outcomes were examined using multiple linear regression. Separate multiple linear regression analyses were performed for each sleep parameter: total sleep time, REM sleep, core and deep sleep, time in bed, in-bed wake time, and sleep efficiency. Given the feasibility design and modest sample size, all adjusted analyses were treated as exploratory. Covariates were limited to a small set of demographic, occupational, and behavioral factors considered clinically relevant to nurses' sleep. This variable-selection strategy reduced, but did not eliminate, the risk of overfitting, particularly for categorical predictors that increased the number of model parameters. No formal adjustment was made for multiple comparisons. Accordingly, results were interpreted with reference to effect estimates, CIs, and patterns of consistency across related outcomes rather than statistical significance in isolation. Sleep-stage variables were also considered descriptively, as they represent interdependent components of total sleep time. Statistical significance was set at P<.05.

Results

Participant Characteristics

Fifty-five nurses participated, of whom 36 (65.5%) worked rotating shifts and 19 (34.5%) were on single-shift duties (Table

XSL•FO
RenderX

1). The mean age was 36 (SD 9) years, and most participants were female individuals (92.7%). The mean BMI was 24 (SD 5.3) kg/m², with 43.6% (n=24) overweight or obese (BMI>23 kg/m^2). Around half (n=26, 47.3%) were married, and fewer had children (n=19, 34.5%). Most respondents were graduates of a baccalaureate nursing degree with honors program (60.0%). The mean total length of service was 12.2 (SD 10) years, significantly longer among nonshift nurses compared with shift nurses (15 vs 5.5 y; *P*=.02). The mean Sleep Hygiene Index score was significantly higher among shift nurses (19.8 vs 15.6; *P*=.01), indicating poorer sleep hygiene in this group. Other sociodemographic variables did not differ significantly between groups (Table 1).

XSL•FO
RenderX

Table . Demographic and occupational characteristics of nurses with different shifts.

Characteristics	All nurses (N=55)	Single-shift duty (n=19)	Rotating-shift duties (n=36)	Statistics	*P* value[a]
Age (y), mean (SD)	36 (9.0)	38 (34 - 45)	35 (28 - 38)	1.96 (53)[b]	.06
Sex, n (%)				—[c]	.29[d]
Male	4 (7.3)	0 (0)	6 (100.0)		
Female	51 (92.7)	19 (37.3)	32 (62.7)		
Race, n (%)				3.11 (1)[e]	.38
Chinese	40 (72.7)	16 (40.0)	24 (60.0)		
Malay	6 (10.9)	2 (33.3)	4 (66.7)		
Others (Indian, Filipino, and Eurasian)	9 (16.4)	1 (11.1)	8 (88.9)		
BMI				0.21 (53)[b]	.84
Mean (SD), kg/m^2	24.0 (5.3)	24.2 (5.5)	23.8 (5.2)		
Normal weight (BMI<23 kg/m^2), n (%)	31 (56.4)	9 (29.0)	22 (71.0)		
Overweight (23<BMI<27.4 kg/m^2), n (%)	8 (14.5)	4 (50.0)	4 (50.0)		
Obese (BMI ≥27.5 kg/m^2), n (%)	16 (29.1)	6 (37.5)	10 (62.5)		
Smoking status, n (%)				—[c]	>.99[d]
Yes	1 (1.8)	0 (0.0)	1 (100.0)		
No	54 (98.2)	19 (35.2)	35 (64.8)		
Marital status, n (%)				3.06 (2)[e]	.22
Single	26 (47.3)	6 (23.1)	20 (76.9)		
Married	26 (47.3)	12 (46.2)	14 (53.8)		
Divorced or widowed	3 (5.4)	1 (33.3)	2 (66.7)		
Have children, n (%)				2.11 (1)[e]	.15
Yes	19 (34.5)	9 (47.4)	10 (52.6)		
No	36 (65.5)	10 (27.8)	26 (72.2)		
Caring for older family member, n (%)				2.10 (1)[e]	.16
Yes	9 (16.4)	5 (55.6)	4 (44.4)		
No	46 (83.6)	14 (30.4)	32 (69.6)		
Highest educational level, n (%)				0.86 (2)[e]	.65
Institute of education diploma, polytechnic diploma, or junior college	11 (20.0)	4 (36.4)	7 (63.6)		
Bachelor’s degree	33 (60.0)	10 (30.3)	23 (69.7)		
Master’s degree and above	11 (20.0)	5 (45.5)	6 (54.5)		
Work unit, n (%)				4.87 (3)[e]	.18
Medical ward	10 (18.2)	3 (30.0)	7 (70.0)		

XSL•FO
RenderX

Characteristics	All nurses (N=55)	Single-shift duty (n=19)	Rotating-shift duties (n=36)	Statistics	*P* value[a]
Surgical ward	11 (20.0)	2 (18.2)	9 (81.8)		
Department of Emergency Medicine and Operating Theater	10 (18.2)	2 (20.0)	8 (80.0)		
Others (isolation units, administrative, specialist center, community, radiology, specialist outpatient clinic)	24 (43.6)	12 (50.0)	12 (50.0)		
Job title, n (%)				1.04 (1)[e]	.31
Staff nurse or senior staff nurse	34 (61.8)	10 (29.4)	24 (70.6)		
Assistant nurse clinician or nurse clinician	21 (38.2)	9 (42.9)	12 (57.1)		
Years of working experience in current work unit, n (%)				3.78 (5)[e]	.29
≤2	11 (20.0)	3 (27.3)	8 (72.7)		
>2-5	15 (27.3)	3 (20.0)	12 (80.0)		
>5-10	9 (16.3)	3 (33.3)	6 (66.7)		
≥10	20 (36.4)	10 (50.0)	10 (50.0)		
Total length of service (y), mean (SD)	12.2 (10.0)	16.4 (8.1)	10 (10.3)	2.36 (53)[b]	.02
Sleep Hygiene Index, mean (SD)	18.3 (5.7)	15.6 (5.1)	19.8 (5.6)	−2.68 (53)[b]	.01

[a]Significant at $P<.05$.

[b]Independent 2-sample, 2-tailed *t* test (*df*).

[c]Not available.

[d]Fisher exact test was used since the expected counts were <5 in 50% of cells.

[e]Chi-square test (*df*).

Agreement and Feasibility of the Apple Watch

Five nurses completed up to 14 consecutive nights of concurrent monitoring with the Apple Watch and GENEActiv. Apple Watch showed preliminary agreement with GENEActiv for total sleep time, in-bed wake time, and sleep efficiency (ICC=0.95, 0.72, and 0.69, respectively; Table 2). Figure 1 shows nightly total sleep time for valid paired recordings only. Across participants, the 2 devices showed closely aligned total sleep time patterns.

In phase 1, 5 nurses contributed 70 expected monitoring nights over 14 days. The Apple Watch was worn on 67 of 70 nights, giving a wear compliance of 95.7%, while 3 nights were classified as nonwear. Of the 67 nights when the Apple Watch was worn, 56 nights yielded valid Apple Watch sleep recordings, giving a data completeness rate of 83.6% among worn nights. The remaining 11 worn nights had no Apple Watch sleep output despite the device being charged and initialized before use and were therefore classified as device-related or synchronization-related data loss. For GENEActiv, 67 of 70 expected nights yielded valid recordings, while 3 nights were missing because the device was not worn. Paired device-comparison analyses were restricted to nights with valid recordings from both the Apple Watch and GENEActiv.

Across the 55 nurses, there were 770 expected nights of Apple Watch monitoring. The Apple Watch was worn on 755 nights, corresponding to 98.1% wear compliance, while 15 nights were classified as nonwear. Among the 755 worn nights, 705 yielded valid Apple Watch sleep recordings, corresponding to 93.4% data completeness among worn nights. The remaining 50 worn nights had no recorded Apple Watch sleep output despite device charging and initialization, and these nights were classified as device-related or synchronization-related data loss. Thus, missing Apple Watch data were separated into nonwear nights and nights with device-related missing output despite reported device wear.

XSL•FO
RenderX

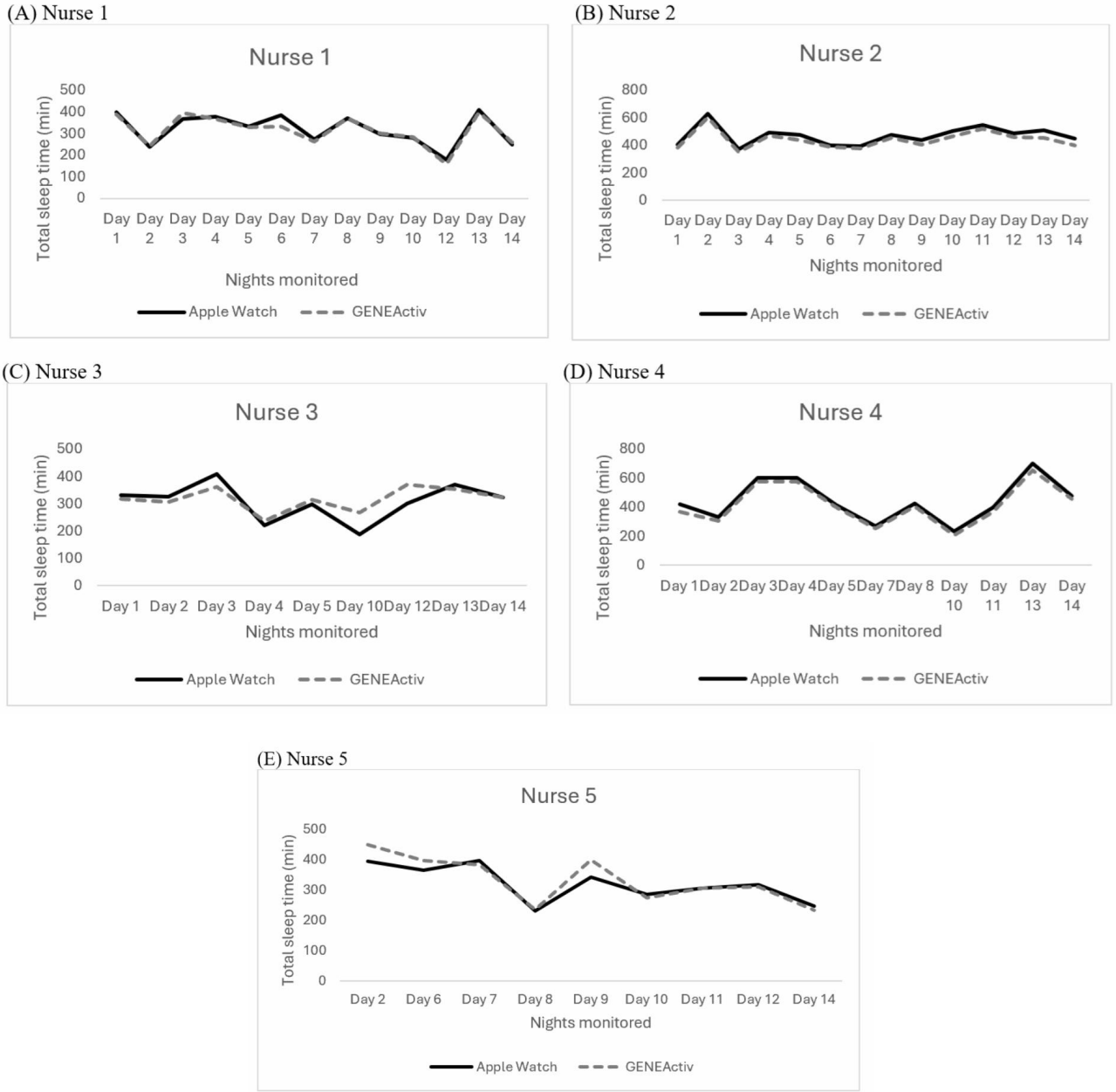
Table . Intraclass correlation coefficients (ICC) examining preliminary agreement between the Apple Watch and GENEActiv for total sleep time, in-bed wake time, and sleep efficiency (N=5)[a].

Sleep parameters	ICC (95% CI)	*P* value[b]
Total sleep time (min)	0.95 (0.06 to 1.00)	<.001
In-bed wake time (min)	0.72 (–0.05 to 0.97)	.05
Sleep efficiency	0.69 (–0.09 to 0.96)	.03

[a]ICCs represent comparisons between the Apple Watch and GENEActiv. Reliability is deemed "poor" if the ICC is <0.4, "fair" if between 0.4 and 0.59, "good" if between 0.60 and 0.74, and "excellent" if above 0.75.

[b]Significant at $P<.05$.

Figure 1. Nightly total sleep time recorded by the Apple Watch and the GENEActiv across 14 nights for (A-E) 5 nurses. Note that the missing data reflect nights when either device was not worn, was worn incorrectly, or had incomplete sleep or wake time entries, and these nights were excluded from the analysis.



Sleep Parameters by Shift Type

The comparison of sleep parameters between nurses working single-shift duty (day, night, or office hours) and those on rotating-shift duties (day, afternoon, and night) is reported in Table 3. Overall, the nurses obtained a mean total sleep time of 380.55 (SD 55.08) minutes. Nurses on rotating shifts recorded a mean total sleep time of 382.03 (SD 55.92) minutes compared

XSL•FO
RenderX

with 377.76 (SD 55.85) minutes among single-shift nurses; this difference was not significant in the unadjusted analysis (mean difference=−4.27, 95% CI −35.87 to 27.33 min; *P*=0.79) and remained nonsignificant after adjustment for age, sex, BMI, presence of children, workplace, total length of service, and sleep hygiene score (adjusted mean difference=−7.38, 95% CI −40.74 to 25.99 min; *P*=.66). Similarly, no significant group differences were found for REM, core, or deep sleep in either unadjusted or adjusted analyses (*P*>.05).

The mean time in bed for all nurses was 401.45 (SD 54.57) minutes, the mean wake time in bed was 19.79 (SD 13.76) minutes, and the average sleep efficiency was 94.77% (SD 4.11%). Mean time in bed was similar between nurses on single shifts (402.46, SD 55.09 min) and rotating shifts (400.92, SD 55.07 min), with no statistically significant differences before or after covariate adjustment. Similarly, in-bed wake time and sleep efficiency did not differ significantly by shift type.

Table . Sleep parameters of nurses with different shifts.

Sleep parameters	All nurses (N=55), mean (SD)	One-shift duty (day, night, or office hours; n=19), mean (SD)	Rotating-shift duties (morning, afternoon, and night; n=36), mean (SD)	Unadjusted mean difference[a] (95% CI)	*t* test (*df*)	*P* value	Adjusted mean difference[b] (95% CI)	*P* value [c]
Total sleep time (min)	380.55 (55.08)	377.76 (55.85)	382.03 (55.92)	−4.27 (−35.87 to 27.33)	–0.27 (53)	.79	−7.38 (−40.74 to 25.99)	.66
Rapid eye movement	80.63 (20.26)	83.89 (18.51)	78.90 (21.17)	4.99 (−6.56 to 16.54)	0.87 (53)	.39	5.58 (−6.42 to 17.58)	.35
Core (light)	240.86 (42.07)	238.20 (45.76)	242.27 (40.60)	−4.07 (−28.20 to 20.06)	–0.34 (53)	.74	−7.81 (−34.08 to 18.46)	.55
Deep	48.29 (10.02)	49.76 (8.77)	47.52 (10.66)	2.24 (−3.48 to 7.96)	0.79 (53)	.44	3.13 (−2.41 to 8.67)	.26
Time in bed (min)	401.45 (54.57)	402.46 (55.09)	400.92 (55.07)	1.54 (−29.79 to 32.86)	0.10 (53)	.92	−0.23 (−33.53 to 33.08)	.99
In-bed wake time (min)	19.79 (13.76)	23.20 (20.08)	17.99 (8.66)	5.21 (−2.56 to 12.98)	1.35 (53)	.18	5.04 (−4.18 to 14.25)	.28
Sleep efficiency (%; total sleep time/time in bed)×100	94.77 (4.11)	93.90 (5.03)	95.22 (3.53)	−1.33 (−3.66 to 1.00)	–1.14 (53)	.26	−1.78 (−4.47 to 0.90)	.19

[a]Independent 2-sample, 2-tailed *t* test.

[b]Adjusted for age, sex, BMI, presence of children, workplace, total length of service, and sleep hygiene score using multiple linear regression.

[c]Significant at *P*<.05.

Associations Between Selected Variables and Sleep Outcomes

Results from multiple linear regression analyses examining exploratory associations between selected variables and sleep parameters derived from Apple Watch devices were presented in Table 4. Given the modest sample size relative to the number of predictors and outcomes, these findings should be regarded as exploratory and hypothesis-generating. After adjustment for age, sex, BMI, shift duties, parental status, work unit, total length of service, and Sleep Hygiene Index, several associations were observed. Higher BMI was associated with shorter total sleep time (unstandardized coefficient *B*=−3.76, 95% CI −6.70 to −0.81 min/kg/m²; *P*=.01), reduced REM sleep duration (unstandardized coefficient *B*=−1.18, 95% CI −2.24 to −0.12 min; *P*=.03), and shorter core (light) sleep duration (unstandardized coefficient *B*=−2.72, 95% CI −5.04 to −0.40 min; *P*=.02). BMI was also associated with reduced time in bed (unstandardized coefficient *B*=−4.05, 95% CI −6.99 to −1.11 min; *P*=.008). Age was negatively associated with deep sleep duration, with older age associated with less deep sleep (unstandardized coefficient *B*=−1.00, 95% CI −1.65 to −0.35 min/y; *P*=.003). Overall, sex, parental status, work unit, total length of service, and Sleep Hygiene Index were not significantly associated with total sleep time, REM sleep, deep sleep, in-bed wake time, or sleep efficiency (*P*>.05).

XSL•FO
RenderX

Table . Multiple linear regression analyses examining sleep parameters among nurses (N=55)[a].

Characteristics	Total sleep time (min)		REM[b] (min)		Core (min)		Deep (min)		Time in bed (min)		In-bed wake time (min)		Sleep efficiency (%)	
	B (95% CI)	*P* value[c]	*B* (95% CI)	*P* value[c]	*B* (95% CI)	*P* value[c]	*B* (95% CI)	*P* value[c]	*B* (95% CI)	*P* value[c]	*B* (95% CI)	*P* value[c]	*B* (95% CI)	*P* value[c]
Age (y)	−0.57 (−4.45 to 3.32)	.77	−0.69 (−2.09 to 0.71)	.33	0.87 (−2.20 to 3.93)	.57	−1.00 (−1.65 to 0.35)	.003	0.16 (−3.72 to 4.04)	.93	0.29 (−0.79 to 1.36)	.59	−0.15 (−0.47 to 0.16)	.33
Sex (female)	35.55 (−24.61 to 95.71)	.24	7.22 (−14.42 to 38.86)	.51	28.57 (−18.81 to 75.94)	.23	6.02 (−3.97 to 16.00)	.23	25.95 (−34.12 to 86.01)	.39	−6.43 (−23.04 to 10.19)	.44	2.95 (−0.47 to 0.16)	.33
BMI (kg/m^2)	−3.76 (−6.70 to −0.81)	.01	−1.18 (−2.24 to −0.12)	.03	−2.72 (−5.04 to −0.40)	.02	−0.32 (−0.81 to 6.72)	.19	−4.05 (−6.99 to −1.11)	.008	−0.06 (−0.87 to 0.75)	.88	0.03 (−0.21 to 0.26)	.83
Have children (yes)	11.26 (−22.95 to 45.48)	.51	7.77 (−4.54 to 20.08)	.21	7.91 (−19.04 to 35.85)	.56	1.04 (−4.64 to 6.72)	.71	10.23 (−23.94 to 44.39)	.55	−1.47 (−10.92 to 7.98)	.76	0.38 (−2.37 to 3.13)	.78
Work unit	1.17 (−11.51 to 13.84)	.85	0.77 (−3.79 to 5.33)	.74	−2.20 (−12.18 to 7.78)	.66	0.32 (−1.78 to 2.43)	.76	1.58 (−11.07 to 14.23)	.80	0.61 (−2.99 to 4.11)	.73	−0.01 (−1.03 to 1.01)	.98
Total length of service (y)	−0.80 (−4.39 to 2.79)	.66	−0.18 (−1.47 to 1.11)	.78	−1.24 (−4.07 to 1.59)	.38	0.50 (−0.10 to 1.09)	.10	−1.20 (−4.78 to 2.39)	.51	−0.04 (−1.03 to 0.95)	.94	0.06 (−0.23 to 0.35)	.70
Sleep Hygiene Index	−1.31 (−4.13 to 1.51)	.36	−0.38 (−1.40 to 0.63)	.45	−1.15 (−3.37 to 1.07)	.30	0.03 (−0.43 to 0.50)	.88	−1.02 (−3.84 to 1.80)	.47	0.07 (−0.71 to 0.85)	.86	−0.11 (−0.33 to 0.12)	.35

[a]Values are unstandardized coefficients (*B*) with 95% CIs and *P* values. Each column represents a separate regression model adjusted for all listed predictors; reference categories: male (sex) and no children.

[b]REM: rapid eye movement.

[c]Significant at $P<.05$.

Discussion

Principal Findings

This study evaluated the feasibility of consumer wearable sleep monitoring among nurses and obtained preliminary evidence of agreement with validated actigraphy in a small validation sample for total sleep time, in-bed wake time, and sleep efficiency. These findings are consistent with previous studies showing that consumer sleep technologies can provide reasonably accurate estimates of total sleep duration compared with research-grade actigraphy devices [33,34]. Much of the existing literature on nurses' sleep has relied on subjective measures, which capture perceived sleep quality but may not accurately quantify sleep duration, sleep-wake patterns, or awakenings. Previous studies have also shown limited agreement between subjective sleep reports and objective measures among nurses [35]. Accordingly, actigraphy and wearable devices are increasingly used to support continuous sleep assessment during clinical schedules [36], although further validation and longer monitoring periods are needed to establish their accuracy and utility in occupational settings [37].

Wear compliance and data completeness supported the feasibility of Apple Watch monitoring in this setting, although some data loss resulted from nonwear or technical issues, as expected in real-world wearable studies [38,39]. Beyond measurement, wearable feedback may complement sleep education by increasing awareness of sleep routines, consistent with behavioral self-efficacy and behavior change models [40,41].

Nurses slept an average of 381 minutes, or approximately 6.4 hours per night, which is below the recommended 7 to 9 hours for adults [42]. Although sleep efficiency was high (94.77%), this should not be interpreted as unequivocal evidence of robust sleep quality. In the context of short total sleep duration and limited time in bed, elevated sleep efficiency may instead reflect constrained sleep opportunity, whereby nurses slept for most of the limited time available rather than achieving sufficient or restorative sleep. This finding is consistent with studies reporting

XSL•FO
RenderX

a high prevalence of poor sleep quality and disrupted sleep among nurses in Asian and shift-working populations [43-46]. Insufficient sleep has been associated with fatigue, impaired cognitive and psychomotor performance, burnout, emotional dysregulation, and occupational fatigue, underscoring the clinical relevance of the short sleep duration observed in this cohort [47-51].

Contrary to expectations, shift duty status was not significantly associated with any sleep parameter after covariate adjustment. Instead, BMI showed the most consistent associations, with higher BMI associated with shorter total sleep time, reduced REM and core sleep, and less time in bed. Although shift nurses reported poorer sleep hygiene than nonshift nurses, Sleep Hygiene Index scores were not independently associated with sleep outcomes in the adjusted models. These findings suggest that individual lifestyle and behavioral factors may be associated with sleep outcomes beyond work schedule alone. Similar observations have been reported in studies showing associations between daily routines, diet, sleep timing around shifts, and sleep quality among nurses [52]. Multiple correlated sleep outcomes were examined across several predictors without adjustment for multiple comparisons; therefore, the regression findings should be interpreted cautiously, with greater emphasis on effect sizes, CIs, and consistency across related outcomes rather than isolated *P* values.

Age also emerged as a significant determinant of sleep outcomes, with older nurses showing reduced deep sleep duration. This pattern is consistent with well-established evidence showing that deep sleep declines across adulthood as a result of age-related changes in cortical activity and sleep homeostasis [53]. Recent studies using both consumer wearables similarly reported a reduction in deep sleep with age among working adults, including health care staff [38,54]. Given that deep sleep contributes to physiological recovery, cognitive functioning, and metabolic regulation, reduced deep sleep among older nurses may have important consequences for fatigue, resilience, and day-to-day clinical effectiveness. Supporting older nurses through workplace policies and tailored sleep-health interventions may be an important component of broader efforts to support and maintain the long-term sustainability of the nursing workforce.

Implications for Nursing Practice

These findings have practice-relevant implications for the nursing workforce. Organizational strategies should prioritize sleep opportunity and fatigue risk management through rostering review, protected rest periods, and minimization of excessive consecutive shifts or overtime. Older nurses may also benefit from tailored fatigue-management support.

Wearable sleep tracking may complement staff well-being initiatives by providing objective feedback alongside sleep education, particularly for nurses who underestimate sleep disruption based on self-report alone. Accordingly, any recommendation for broader Apple Watch use should be balanced against considerations of cost, scalability, and local implementation needs, including whether lower-cost wearables, validated sleep questionnaires, or targeted actigraphy may provide a more proportionate approach for workplace sleep-health initiatives.

Third, the association between higher BMI and poorer sleep parameters underscores the importance of adopting a holistic approach to health promotion within the nursing workforce. Initiatives that support physical activity, healthy nutrition, weight management, and stress regulation may contribute not only to improved sleep outcomes but also to overall well-being.

Finally, poorer sleep hygiene among shift nurses indicates the value of incorporating practical sleep-hygiene guidance into staff well-being programs across shift types.

Limitations

Several limitations should be considered. First, the validation component involved only 5 nurses; therefore, the device-comparison findings should be interpreted as preliminary feasibility evidence rather than established measurement accuracy. Although small samples are acceptable in feasibility studies, recruiting the cohort from a single institution may limit generalizability. An a priori power analysis was not conducted because the study primarily evaluated feasibility. The adjusted analyses included several covariates and multiple interdependent sleep outcomes relative to the sample size and may therefore have been underpowered, overfitted, or imprecise. As no adjustment for multiple comparisons was applied, statistically significant findings should be regarded as exploratory and interpreted cautiously until confirmed in adequately powered studies with prespecified analytic plans.

Self-selection bias may also have occurred, as nurses who volunteered may have had a greater interest in sleep monitoring or have perceived sleep difficulties. Self-reported questionnaires, including the Sleep Hygiene Index, may be affected by recall or social desirability bias. The nonsignificant associations between Sleep Hygiene Index scores and objective sleep outcomes should be interpreted cautiously, as they may reflect measurement limitations, restricted variability in this relatively homogeneous cohort, and limited power to detect small associations. In addition, manual inspection of sleep data was performed only by the principal investigator; although the principal investigator was blinded to participants' shift rosters, the absence of independent interrater verification may have introduced subjective judgment in determining data completeness and consistency.

Chronotype was not assessed in this study, although individual circadian preference may influence tolerance for shift work and could partly explain why shift status was not independently associated with sleep outcomes. Future studies should examine whether chronotype moderates the relationship between shift schedules and sleep outcomes to guide chronobiologically informed scheduling. Finally, although total sleep time, in-bed wake time, and sleep efficiency were compared between the Apple Watch and the GENEActiv, device-specific sleep-window definitions, algorithms, scoring rules, and firmware updates may affect wake-related estimates and reproducibility over time.

Despite these limitations, this study provides objective sleep data among nurses in an Asian context, as well as preliminary evidence supporting the feasibility of Apple Watch–based

XSL•FO
RenderX

monitoring and its agreement with validated actigraphy for selected sleep parameters. The observed associations remain exploratory given the observational design and modest sample size.

Conclusions

This study showed that nurses obtained less sleep than recommended, underscoring concerns about sleep opportunity and recovery in this workforce. The Apple Watch provided preliminary evidence of feasibility and agreement with GENEActiv for total sleep time, in-bed wake time, and sleep efficiency in a small validation sample. BMI and age were associated with sleep outcomes in exploratory analyses. These findings support workplace strategies that combine sleep education, sleep hygiene training, and broader fatigue-management efforts to promote healthier sleep among nurses.

Acknowledgments

Apple Inc provided Apple Watch devices for the research. Apple was not involved in the design of the research, nor was it involved in the collection, analysis, or interpretation of the research data or in the content of this or any related publication. Grammarly was used solely for language editing, including grammar, spelling, and clarity. It was not used to generate scientific content, analyze or interpret data, or create references or citations. The authors reviewed and verified the final manuscript and take full responsibility for its accuracy, originality, and integrity.

Funding

This research is supported by the Singapore Ministry of Health's National Medical Research Council under its Clinician-Scientist Individual Research Grant New Investigator Grant (CS-IRG NIG; NMRC Project No. MOH-CNIG24jan-0002) and the SingHealth Academic Medicine Innovation Institute Nursing Innovation Grant 2024 (Grant No. 2024AMII-May03).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

References

1. Weaver MD, Landrigan CP, Sullivan JP, et al. National improvements in resident physician-reported patient safety after limiting first-year resident physicians' extended duration work shifts: a pooled analysis of prospective cohort studies. BMJ Qual Saf 2023 Feb;32(2):81-89. [doi: 10.1136/bmjqs-2021-014375] [Medline: 35537821]
2. Chang WP, Peng YX. Meta-analysis of differences in sleep quality based on actigraphs between day and night shift workers and the moderating effect of age. J Occup Health 2021 Jan;63(1):e12262. [doi: 10.1002/1348-9585.12262] [Medline: 34392580]
3. James SM, Honn KA, Gaddameedhi S, Van Dongen HPA. Shift work: disrupted circadian rhythms and sleep—implications for health and well-being. Curr Sleep Med Rep 2017 Jun;3(2):104-112. [doi: 10.1007/s40675-017-0071-6] [Medline: 29057204]
4. Sateia MJ. International classification of sleep disorders-third edition: highlights and modifications. Chest 2014 Nov;146(5):1387-1394. [doi: 10.1378/chest.14-0970] [Medline: 25367475]
5. Al Maqbali M, Al Sinani M, Al-Lenjawi B. Prevalence of stress, depression, anxiety and sleep disturbance among nurses during the COVID-19 pandemic: a systematic review and meta-analysis. J Psychosom Res 2021 Feb;141:110343. [doi: 10.1016/j.jpsychores.2020.110343] [Medline: 33360329]
6. Booker LA, Sletten TL, Alvaro PK, et al. Exploring the associations between shift work disorder, depression, anxiety and sick leave taken amongst nurses. J Sleep Res 2020 Jun;29(3):e12872. [doi: 10.1111/jsr.12872] [Medline: 31144389]
7. Dogan HD. Examining sleep quality and job satisfaction of critical care nurses. Clin Exp Health Sci 2020;10(3):247-256. [doi: 10.33808/clinexphealthsci.639710]
8. Chang WP, Chang YP. Relationship between job satisfaction and sleep quality of female shift-working nurses: using shift type as moderator variable. Ind Health 2019 Nov 29;57(6):732-740. [doi: 10.2486/indhealth.2018-0258] [Medline: 30930373]
9. Shan Z, Li Y, Zong G, et al. Rotating night shift work and adherence to unhealthy lifestyle in predicting risk of type 2 diabetes: results from two large US cohorts of female nurses. BMJ 2018 Nov 21;363:k4641. [doi: 10.1136/bmj.k4641] [Medline: 30464025]
10. Fagundo-Rivera J, Gómez-Salgado J, García-Iglesias JJ, Gómez-Salgado C, Camacho-Martín S, Ruiz-Frutos C. Relationship between night shifts and risk of breast cancer among nurses: a systematic review. Medicina (Kaunas) 2020 Dec 10;56(12):680. [doi: 10.3390/medicina56120680] [Medline: 33321692]

XSL•FO
RenderX

11. Lee S, Gonzalez BD, Small BJ. My job impacts my sleep: signs and symptoms of insomnia among healthcare workers. Ind Health 2021;59(2):86-98. [doi: 10.2486/indhealth.2020-0191] [Medline: 33762517]
12. Wangsan K, Chaiear N, Sawanyawisuth K, Klainin-Yobas P, Simajareuk K, Boonsawat W. Which shiftwork pattern is the strongest predictor for poor sleep quality in nurses? Int J Environ Res Public Health 2022 Oct 27;19(21):13986. [doi: 10.3390/ijerph192113986] [Medline: 36360864]
13. Niu SF, Miao NF, Liao YM, Chi MJ, Chung MH, Chou KR. Sleep quality associated with different work schedules: a longitudinal study of nursing staff. Biol Res Nurs 2017 Jul;19(4):375-381. [doi: 10.1177/1099800417695483] [Medline: 28627307]
14. Shiffer D, Minonzio M, Dipaola F, et al. Effects of clockwise and counterclockwise job shift work rotation on sleep and work-life balance on hospital nurses. Int J Environ Res Public Health 2018 Sep 18;15(9):2038. [doi: 10.3390/ijerph15092038] [Medline: 30231514]
15. Lichstein KL, Morin CM, editors. Treatment of Late-Life Insomnia: SAGE Publications; 2000. [doi: 10.4135/9781452225555]
16. Prasad R, Pandey AK. Sleep hygiene and good sleep habits: a review. Indian J Sleep Med 2012;7(4):134-138. [doi: 10.5958/j.0974-0155.7.4.020]
17. Thorpy M. International Classification of Sleep Disorders. In: Sleep Disorders Medicine: Basic Science, Technical Considerations and Clinical Aspects: Springer; 2017:475-484. [doi: 10.1007/978-1-4939-6578-6_27]
18. Booker LA, Barnes M, Alvaro P, et al. The role of sleep hygiene in the risk of Shift Work Disorder in nurses. Sleep 2020 Feb 13;43(2):zsz228. [doi: 10.1093/sleep/zsz228] [Medline: 31637435]
19. Jeong JY, Gu MO. The influence of stress response, physical activity, and sleep hygiene on sleep quality of shift work nurses. J Korea Acad Ind Coop Soc 2016 Jun 30;17(6):546-559. [doi: 10.5762/KAIS.2016.17.6.546]
20. Yuan H, Hill EA, Kyle SD, Doherty A. A systematic review of the performance of actigraphy in measuring sleep stages. J Sleep Res 2024 Oct;33(5):e14143. [doi: 10.1111/jsr.14143] [Medline: 38384163]
21. Baron KG, Duffecy J, Berendsen MA, Cheung Mason I, Lattie EG, Manalo NC. Feeling validated yet? A scoping review of the use of consumer-targeted wearable and mobile technology to measure and improve sleep. Sleep Med Rev 2018 Aug;40:151-159. [doi: 10.1016/j.smrv.2017.12.002] [Medline: 29395985]
22. Miller DJ, Sargent C, Roach GD. A validation of six wearable devices for estimating sleep, heart rate and heart rate variability in healthy adults. Sensors (Basel) 2022 Aug 22;22(16):6317. [doi: 10.3390/s22166317] [Medline: 36016077]
23. Roomkham S, Hittle M, Cheung J, Lovell D, Mignot E, Perrin D. Sleep monitoring with the Apple Watch: comparison to a clinically validated actigraph. F1000Res 2019;8:754. [doi: 10.12688/f1000research.19020.1]
24. Tickle-Degnen L. Nuts and bolts of conducting feasibility studies. Am J Occup Ther 2013;67(2):171-176. [doi: 10.5014/ajot.2013.006270] [Medline: 23433271]
25. Chinoy ED, Cuellar JA, Huwa KE, et al. Performance of seven consumer sleep-tracking devices compared with polysomnography. Sleep 2021 May 14;44(5):zsaa291. [doi: 10.1093/sleep/zsaa291] [Medline: 33378539]
26. Montgomery R. Sample size justification in feasibility studies: moving beyond published guidance. Pilot Feasibility Stud 2025 Jun 23;11(1):88. [doi: 10.1186/s40814-025-01675-9] [Medline: 40551266]
27. Rahimi MM, Phillips CL, Marshall NS, et al. Moderately strong intraclass correlations between actigraphic and polysomnographic total sleep time and sleep efficiency in older adults with sleep disturbance. Sleep Breath 2025 Apr 15;29(2):161. [doi: 10.1007/s11325-025-03326-y] [Medline: 40232641]
28. Track your sleep with Apple Watch. Apple Support. 2026. URL: https://support.apple.com/en-sg/guide/watch/apd830528336/watchos [accessed 2026-07-25]
29. Public health research, raw data accelerometer. ActivInsights. 2026. URL: https://activinsights.com/remote-monitoring-wearables/geneactiv-raw-data-accelerometer-public-health-research [accessed 2026-07-25]
30. Actigraphy Sleep Toolkit. 2025. URL: https://sleeptoolkit.activinsights.net [accessed 2026-07-25]
31. Mastin DF, Bryson J, Corwyn R. Assessment of sleep hygiene using the Sleep Hygiene Index. J Behav Med 2006 Jun;29(3):223-227. [doi: 10.1007/s10865-006-9047-6] [Medline: 16557353]
32. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6(4):284-290. [doi: 10.1037/1040-3590.6.4.284]
33. Kainec KA, Caccavaro J, Barnes M, Hoff C, Berlin A, Spencer RMC. Evaluating accuracy in five commercial sleep-tracking devices compared to research-grade actigraphy and polysomnography. Sensors (Basel) 2024 Jan 19;24(2):635. [doi: 10.3390/s24020635] [Medline: 38276327]
34. Jaworski DJ, Park EJ. Apple Watch sleep and physiological tracking compared to clinically validated actigraphy, ballistocardiography and polysomnography. Presented at: 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); Jul 24-27, 2023. [doi: 10.1109/EMBC40787.2023.10340725]
35. Dujmić Ž, Mikšić Š, Barać I, et al. Sleep quality in shift-working nurses: subjective and objective evaluation. Int J Environ Res Public Health 2025 Dec 31;23(1):64. [doi: 10.3390/ijerph23010064] [Medline: 41595858]
36. Chang WP, Li HB. Influence of shift work on rest-activity rhythms, sleep quality, and fatigue of female nurses. Chronobiol Int 2022 Apr;39(4):557-568. [doi: 10.1080/07420528.2021.2005082] [Medline: 34906006]

XSL•FO
RenderX

37. Ishinuki T, Goda E, Tatsumi H, et al. Utility of a wearable tracker to assess sleep quality in nurses and their spouses: a prospective cohort study. SAGE Open Nurs 2025;11:23779608241267079. [doi: 10.1177/23779608241267079] [Medline: 39872373]
38. de Zambotti M, Cellini N, Goldstone A, Colrain IM, Baker FC. Wearable sleep technology in clinical and research settings. Med Sci Sports Exerc 2019 Jul;51(7):1538-1557. [doi: 10.1249/MSS.0000000000001947] [Medline: 30789439]
39. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016 Oct 24;355:i5239. [doi: 10.1136/bmj.i5239] [Medline: 27777223]
40. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
41. Hittle BM, Hils J, Fendinger SL, Wong IS. A scoping review of sleep education and training for nurses. Int J Nurs Stud 2023 Jun;142:104468. [doi: 10.1016/j.ijnurstu.2023.104468] [Medline: 37080122]
42. Chaput JP, Dutil C, Featherstone R, et al. Sleep duration and health in adults: an overview of systematic reviews. Appl Physiol Nutr Metab 2020 Oct;45(10 (Suppl. 2)):S218-S231. [doi: 10.1139/apnm-2020-0034] [Medline: 33054337]
43. Jeon BM, Kim SH, Shin SH. Effectiveness of sleep interventions for rotating night shift workers: a systematic review and meta-analysis. Front Public Health 2023;11:1187382. [doi: 10.3389/fpubh.2023.1187382] [Medline: 37427284]
44. Supian NS, Ibrahim MI. Factors influencing sleep quality among nurses in tertiary hospitals in Kelantan, Malaysia: a cross-sectional study. BMC Nurs 2024 Dec 3;23(1):882. [doi: 10.1186/s12912-024-02557-w] [Medline: 39627828]
45. Zeng LN, Yang Y, Wang C, et al. Prevalence of poor sleep quality in nursing staff: a meta-analysis of observational studies. Behav Sleep Med 2020;18(6):746-759. [doi: 10.1080/15402002.2019.1677233] [Medline: 31672062]
46. Dong H, Zhang Q, Zhu C, Lv Q. Sleep quality of nurses in the emergency department of public hospitals in China and its influencing factors: a cross-sectional study. Health Qual Life Outcomes 2020 Apr 29;18(1):116. [doi: 10.1186/s12955-020-01374-4] [Medline: 32349759]
47. Alameri RA, Almulla HA, Al Swyan AH, Hammad SS. Sleep quality and fatigue among nurses working in high-acuity clinical settings in Saudi Arabia: a cross-sectional study. BMC Nurs 2024 Jan 18;23(1):51. [doi: 10.1186/s12912-023-01693-z] [Medline: 38238714]
48. Di Muzio M, Diella G, Di Simone E, et al. Nurses and Night Shifts: Poor Sleep Quality Exacerbates Psychomotor Performance. Front Neurosci 2020;14:579938. [doi: 10.3389/fnins.2020.579938] [Medline: 33154716]
49. Haji Mohamud RY, Mohamed NA, Abdi AA, et al. Prevalence and associated factors of poor sleep quality among nurses in a tertiary care hospital: a cross-sectional study. Risk Manag Healthc Policy 2025;18:975-986. [doi: 10.2147/RMHP.S511543] [Medline: 40161898]
50. Membrive-Jiménez MJ, Gómez-Urquiza JL, Suleiman-Martos N, et al. Relation between Burnout and Sleep Problems in Nurses: A Systematic Review with Meta-Analysis. Healthcare (Basel) 2022 May 21;10(5):954. [doi: 10.3390/healthcare10050954] [Medline: 35628091]
51. Zhao Y, Wang X, Liu T. Correlations between empathy fatigue, occupational identity, and sleep quality in nursing staff: a cross-sectional study. Front Public Health 2024;12:1401044. [doi: 10.3389/fpubh.2024.1401044] [Medline: 38932768]
52. Xiao Q, Huang X, Yang T, et al. Determinants of sleep quality and their impact on health outcomes: a cross-sectional study on night-shift nurses. Front Psychiatry 2024;15:1506061. [doi: 10.3389/fpsyt.2024.1506061] [Medline: 39777195]
53. Li J, Vitiello MV, Gooneratne NS. Sleep in normal aging. Sleep Med Clin 2022 Jun;17(2):161-171. [doi: 10.1016/j.jsmc.2022.02.007] [Medline: 35659071]
54. Zheng NS, Annis J, Master H, et al. Sleep patterns and risk of chronic disease as measured by long-term monitoring with commercial wearable devices in the All of Us Research Program. Nat Med 2024 Sep;30(9):2648-2656. [doi: 10.1038/s41591-024-03155-8] [Medline: 39030265]

Abbreviations

ICC: intraclass correlation coefficient
REM: rapid eye movement

Edited by A Sowan, E Borycki; submitted 02.Apr.2026; peer-reviewed by JM Santos-Gago, J Kohyama; revised version received 07.Jul.2026; accepted 16.Jul.2026; published 19.Aug.2026.

Please cite as:
Lim SH, Aloweni F, Ong P, Ang SY
Validation and Feasibility of a Consumer-Grade Wearable Sleep Monitoring Device and an Exploratory Evaluation of Sleep Characteristics and Associated Factors Among Nurses: Observational Cohort Study
JMIR Nursing 2026;9:e96912
URL: https://nursing.jmir.org/2026/1/e96912
doi:10.2196/96912

XSL•FO
RenderX

© Siew Hoon Lim, Fazila Aloweni, Phyllis Ong, Shin Yuh Ang. Originally published in JMIR Nursing (https://nursing.jmir.org), 19.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Data Strategy for the Secondary Use of Nursing Real-World Data: Feasibility Study of the NuDaK Framework

Renate Nantschev, RN, MSc; Elske Ammenwerth, PhD

Institute of Medical Informatics, UMIT TIROL – Private University for Health Sciences and Health Technology, Eduard Wallnöfer Zentrum 1, Hall in Tirol, Tirol, Austria

Corresponding Author:
Renate Nantschev, RN, MSc
Institute of Medical Informatics, UMIT TIROL – Private University for Health Sciences and Health Technology, Eduard Wallnöfer Zentrum 1, Hall in Tirol, Tirol, Austria

Abstract

Background: Routinely collected nursing real-world data (RWD) offer opportunities for nursing research to generate new knowledge and to improve quality of care. Furthermore, advanced analytics with nursing routine data opens the development of predictive models that support risk detection and decision support. However, using RWD in nursing remains challenging due to insufficient data standardization, variable data quality, and difficulty for organizations implementing a targeted data strategy for the secondary use of nursing data. To address this gap, we developed the NuDaK nursing data strategy framework as a practical guide to enable and optimize the secondary use of routinely collected nursing documentation data.

Objective: This study aimed to evaluate the completeness and applicability of the NuDaK framework for the secondary use of nursing routine data in a clinical nursing setting.

Methods: We conducted a formative feasibility study in collaboration with a large acute care hospital in Austria. The NuDaK framework was piloted on a surgical ward and evaluated across 4 implementation phases. Data were collected via phase-specific workshops and project meetings, continuous self-evaluation protocols by the project lead, and a final online focus group interview. Data were analyzed using deductive qualitative content analysis regarding the completeness and applicability of NuDaK. Evaluation results were aggregated and used to update NuDaK.

Results: The 4 core NuDaK aspects (“Purpose & Benefits,” “Dataset & Nursing Documentation,” “Data Modeling & Software Requirements,” and “Data Integration & Data Quality”) were confirmed as conceptually complete. Based on the evaluation results, NuDaK was updated by (1) adopting a circular structure with 4 temporal phases, (2) defining interdisciplinary stakeholder involvement across phases, (3) specifying professional roles and responsibilities across phases, and (4) integrating an iterative nursing information model spanning dataset development, specification, and refinement. Further NuDaK additions comprise a data quality concept with operationalized dimensions and validation procedures.

Conclusions: The NuDaK nursing strategy framework was refined through real-world piloting. It now provides organizations and researchers with a nursing-specific, end-to-end nursing data strategy to enable systematic RWD use. While its feasibility was demonstrated in a real-world nursing setting, further implementation studies should assess transferability across heterogeneous institutions with different documentation systems, governance structures, and levels of digital maturity.

(JMIR Nursing 2026;9:e95546) doi:10.2196/95546

KEYWORDS

nursing informatics; nursing real-world data; secondary use of nursing data; data strategy framework; nursing outcome analytics

Introduction

Background

Digitization in nursing care has advanced considerably. In Germany, digital nursing and treatment documentation averaged 68.4% in 2023 [1]. Similar levels are reported in Austria and Switzerland, where 76% of nursing documentation in hospitals is electronic and 88% of facilities use IT-supported nursing documentation in at least half of their units [2]. Likewise, in the United States, over 96% of hospitals used an electronic health record (EHR) system in 2022 for clinical notes, most with integrated digital nursing documentation [3].

This digital transformation not only changes documentation processes but also involves the routine collection of vast amounts of real-world data (RWD).

RWD are defined as data routinely collected during care provision relating to patient health status or the delivery of health care from various sources, for example, EHRs and documentation systems, disease-specific medical research registries, and data from wearable or mobile devices [4].

XSL•FO
RenderX

RWD can support both nursing research and quality improvement. They offer several advantages because RWD are routinely collected from various health-related sources; they can provide comprehensive and timely representations of patient care and care delivery [5]. The use of RWD in nursing research can help generate evidence-based knowledge for practice, improve the quality of patient care, and further develop nursing practice by, for example, assessing the effectiveness of nursing interventions in achieving specific nursing outcomes based on routine data. In addition, RWD-based analyses can support personalized, patient-centered care while optimizing resource use [6,7]. Advanced analytics with nursing routine data also opens the development of predictive models that support applications such as risk detection and decision support [8-10].

Despite these potentials, the secondary use of nursing RWD remains limited. One major challenge is that nonstandardized data entry limits data reuse. In many settings, standardized nursing terminologies are insufficient or inconsistently applied [2]. This lack of standardization makes it difficult to harmonize and integrate data for meaningful analysis [4,11]. Additionally, the widespread use of free-text documentation, while valued by clinicians for flexibility and completeness, makes data analysis difficult [2,12].

Another challenge is low data quality. Incomplete or incorrect documentation can lead to data gaps and analytical errors and reduce the reliability and validity of results [4].

Finally, many institutions lack a comprehensive data strategy that addresses all aspects relevant to nursing research and nursing outcome evaluation [2]. Without a coordinated approach, data are often collected in isolated application systems, using inconsistent formats and varying documentation standards and terminologies. This fragmentation limits data interoperability, complicates data linkage across settings, and restricts the ability to aggregate information for large-scale analyses [13].

To address these challenges, the literature describes several foundational components that should be addressed to enable the secondary use of nursing RWD. A key component is the standardization and structured documentation of nursing documentation. The consistent use of standardized nursing terminologies and structured data formats enhances the ability to characterize nursing care and analyze its effects on nursing outcomes. Structured and standardized documentation practices are essential for ensuring high data quality, interoperability, and meaningful integration with other clinical data sources [14,15]. Equally important is the precise definition of secondary use purposes and the identification of outcome indicators. Together, these elements provide the foundation for outcome-oriented and evidence-based nursing research [16,17]. Nursing-sensitive outcome indicators serve as key measures in this context. They capture outcomes directly influenced by nursing care, such as patient falls, pressure injuries, pain management, or infection rates, and provide a quantitative basis for evaluating care quality and effectiveness [18-26].

Another essential component for the secondary use of nursing data is the establishment of a defined dataset. A clearly specified dataset determines which data elements are relevant for describing nursing practice and nursing outcomes, ensuring that the collected information is complete, consistent, and comparable across institutions and care settings [27,28].

To ensure that these data are reliable and interpretable, data information models specify the structure and relationships among these datasets and define their meaning in a standardized form. These models bridge clinical and technical perspectives by translating nursing knowledge into formal data specifications, reducing ambiguity and enabling accurate data capture in electronic documentation systems [27,29].

A further critical component is data quality, which determines the reliability and validity of any secondary analysis. High-quality nursing data to be used for RWD analysis must be complete, correct, current, consistent, and plausible, which is supported by systematic data validation and transparent data reporting [30-33].

Finally, data integration is required to consolidate data from multiple sources into an analysis-ready structure (eg, via a data warehouse or data mart), which typically requires close interdisciplinary collaboration among IT professionals, data scientists, and nursing researchers to ensure clinical relevance and technical feasibility [16].

While these components represent core building blocks identified in the literature for enabling the secondary use of nursing RWD, existing frameworks for this task often address only selected aspects rather than providing an integrated, nursing-specific strategy.

The SPIRIT (Systematic Planning of Intelligent Reuse of Integrated Clinical Routine Data) framework offers a secondary use strategy for the entire organization, comprehensive analyses to define system and user requirements, and the implementation of clinical data integration, management, and analysis solutions. However, it does not address specific nursing outcomes and nursing routine data [16].

The framework for mining and analyzing standardized nursing care plan data facilitates the extraction of valuable knowledge from nursing care plan data, demonstrating its applicability in improving nursing outcomes, supporting nursing staff, and enhancing administrative functions [9]. However, the handling of nonstandardized care plan data remains insufficiently addressed, particularly in terms of data harmonization and quality analysis, which are essential to ensuring the reliability and interoperability of the extracted information. The Health Data Quality Framework provides an automated approach for assessing and visualizing the quality of EHR data to support analytics such as predictive modeling and quality measures [33]. This framework primarily targets data quality and thus does not constitute an end-to-end strategy for the reuse of nursing routine data.

Finally, the Governance Framework developed for secondary use of routine health care data in nursing homes outlines governance, legal, and ethical principles for responsible reuse [34]. This framework is not intended to be a comprehensive nursing data strategy across the full life cycle.

Overall, most available frameworks focus on individual dimensions (eg, governance, data quality, or analytic reuse)

XSL•FO
RenderX

without systematically combining all key components into a coherent, nursing-specific data strategy.

In summary, no existing framework comprehensively supports the secondary use of routinely collected nursing documentation data in a nursing-specific way. Current approaches do not adequately address specific nursing outcomes and the particular characteristics of nursing routine data. As a result, an important gap remains in how nursing RWD can be systematically leveraged for nursing-related purposes.

To address this gap, we developed the NuDaK nursing data strategy framework to enable and optimize the secondary use of routinely collected nursing documentation data.

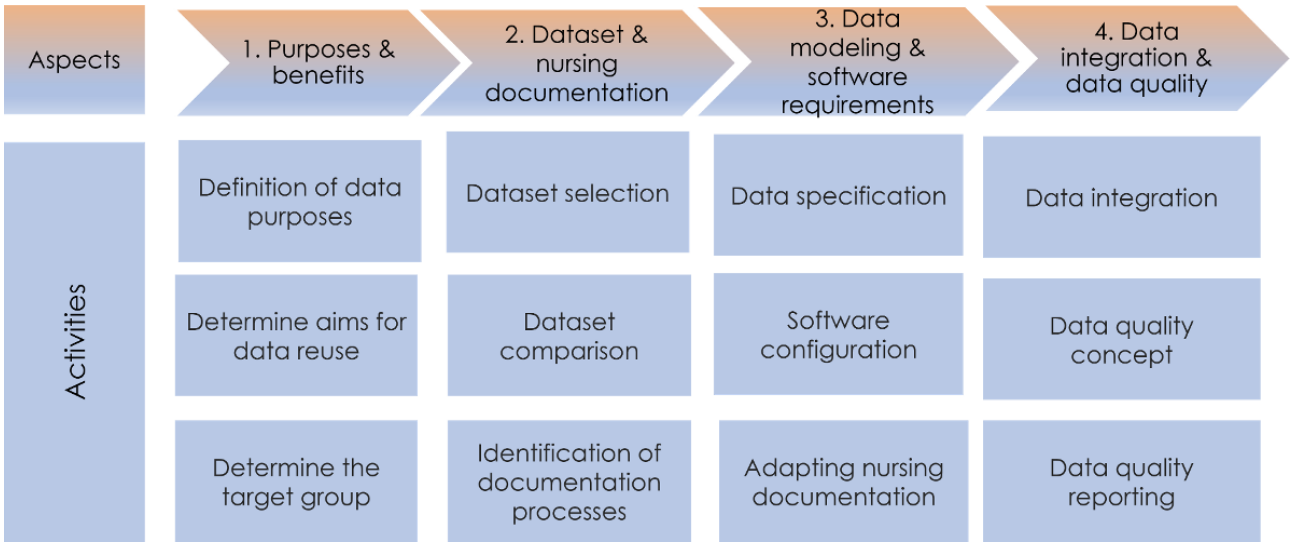
NuDaK does not aim to be universally exhaustive but focuses on addressing nursing data, nursing-related outcomes, and the specific characteristics of routine nursing data. The NuDaK framework serves as a practical guide for both organizations and research teams by structuring how clinical routine data can be prepared, governed, and analyzed for nursing purposes.

In organizational settings, within single institutions or across organizations, NuDaK supports assessing the effectiveness of nursing interventions on nursing outcomes, strengthening nursing quality management, enabling personalized patient-centered care while optimizing resource use, and informing workflow steering.

In research settings**,** NuDaK underpins robust evidence generation in nursing by improving data quality, standardization, and governance, thereby enabling advanced analytics such as data mining, AI-based predictive modeling, risk detection, and decision support.

The NuDaK framework encompasses 4 interrelated aspects that reflect the key components of a comprehensive data strategy (see Figure 1 for an overview) [35,36].

Figure 1. NuDaK nursing data strategy framework (2022) for the secondary use of routine nursing data.



1. Defining the purposes and benefits of using routinely collected data
2. Establishing a standardized nursing dataset and adapting nursing documentation to nursing workflows
3. Data modeling and specification of software requirements
4. Implementing processes for data integration and quality assurance [35,36]

Some information on NuDaK has been published in [35,36]. While NuDaK was built on the results of a narrative-integrative literature review with conceptual synthesis, a comprehensive description and evaluation have not yet been conducted. This is therefore the objective of this paper.

Objective

This study aimed to evaluate the completeness and applicability of the NuDaK data strategy framework in a real clinical nursing setting and, based on these findings, to refine NuDaK as a practical guide for organizations and researchers to enable the secondary use of routinely collected nursing documentation data.

Methods

Study Design

This study is reported in accordance with the STARE-HI (Statement on Reporting of Evaluation Studies in Health Informatics) guideline to enhance the quality and transparency of reporting for quantitative and qualitative health informatics evaluation studies. In line with STARE-HI, the *Methods* section is structured to provide sufficient detail on the study design and methods used in this study [37].

We conducted a 1-year formative feasibility study (February 2023 to June 2024) to pilot and evaluate the NuDaK framework in a real clinical nursing setting. The evaluation focused on 2 primary dimensions: (1) applicability (how well it fits the health care environment) and (2) completeness (whether it covers all necessary aspects and activities comprehensively).

The study used 3 complementary evaluation methods: a formative qualitative external evaluation, a qualitative self-evaluation, and an online focus group interview at the end of the piloting phase.

XSL•FO
RenderX

Theoretical Background

Formative evaluations are particularly employed when a program, product, concept, or intervention is still in development or testing. They primarily serve an optimization function: the goal is to continuously improve the evaluation object during its development or application. Continuous feedback of evaluation results into practice is beneficial, enabling improvements to be implemented and tested through multiple feedback loops [38].

Qualitative methods play an important role in the context of formative evaluations. They enable differentiated capture of experiences, assessments, pros and cons, and concrete change requests from evaluation participants and users [38]. For evaluating the NuDaK framework, this approach is appropriate because the framework is still in the testing phase and requires continuous improvements through iterative feedback loops.

Setting

The study was conducted in collaboration with a large acute care hospital in Austria and the software company that provides the electronic nursing documentation. Together with the stakeholders, we selected a surgical ward to pilot the NuDaK framework.

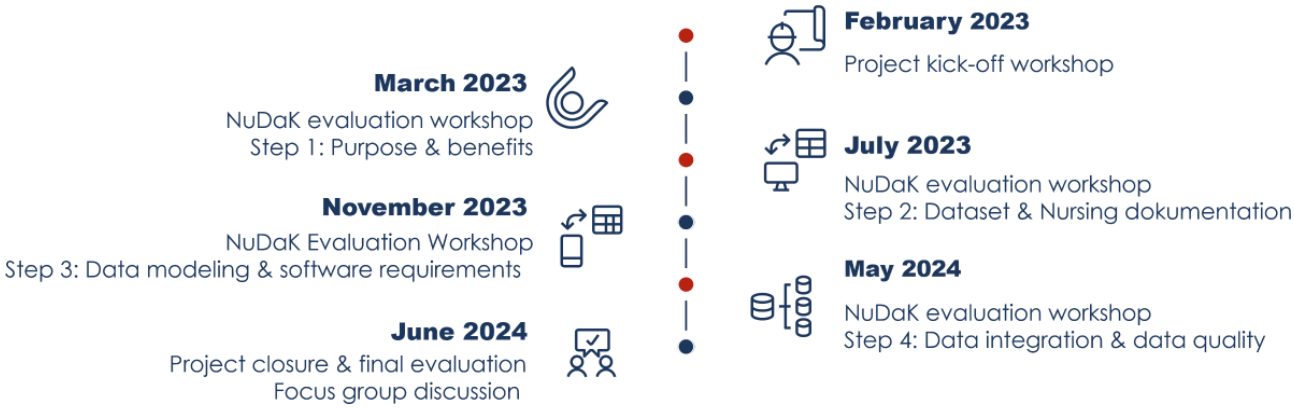
The stakeholders involved in the project were systematically selected through a stakeholder analysis.

In total, 15 stakeholders were identified. Ten active stakeholders participated directly in project activities (director of nursing, experts from the unit of nursing development, ward manager, a practical nurse at the piloting station, IT expert at the hospital, staff member of the software company, and expert in data security and quality management). Five passive stakeholders (executive management, IT business management, and IT requirements and implementation management) supported strategic decision-making and were regularly informed about progress but did not participate in the planned project activities.

Study Flow

The pilot implementation followed a detailed project plan aligned with the 4 aspects of the NuDaK framework. Overall project management was led by the principal researcher, who also served as the project leader. On the clinical side, a designated nurse representative served as the primary point of contact and coordinated all on-site activities. In addition, the software vendor assigned a dedicated project lead to support implementation and technical coordination. The study was conducted from February 2023 to June 2024 and comprised 6 phases: a kick-off phase (February 2023) for initial training and framework introduction; 4 piloting steps corresponding to NuDaK implementation steps 1 to 4, with final evaluation workshops; and project closure (June 2024) for final evaluation. Figure 2 outlines the evaluation timeline within the piloting steps.

Figure 2. NuDaK framework evaluation activities within the piloting steps.



Outcome Measure

This study evaluated the NuDaK framework using 2 primary outcome measures: completeness and applicability.

Completeness addresses whether the framework covers all necessary aspects and activities comprehensively, ensuring there are no gaps in the intended scope. Applicability examines how well the framework fits the real health care environment, including its clarity, usability, and suitability for end users in practice.

The evaluation addressed these predefined criteria using the following guiding question, as presented in Table 1.

XSL•FO
RenderX

Table . Evaluation questions on completeness and applicability of the NuDaK framework.

Main evaluation question	Subquestions
Completeness evaluation	
Q1. Is the set of aspects covered in the NuDaK framework complete and sufficient?	• Are additional aspects or themes required to improve comprehensiveness?
Q2. Are the activities assigned to each NuDaK aspect logically connected?	• Do the activities follow a coherent and meaningful sequence?
Q3. Is each activity within the NuDaK framework feasible to perform in practice?	• Are temporal dependencies between activities clearly defined and logically structured?
Q4. Is each NuDaK framework activity feasible in practice, given the required resources and time?	• Does the framework link resource needs to specific aspects, activities, or methods? • Is the time required for each step realistic and proportionate to available resources?
Applicability evaluation	
Q5. To what extent is the NuDaK framework applicable in practice?	• Are the described aspects and activities understandable and logically structured?
Q6. How clear and comprehensible are the NuDaK framework's aspects and activities for end users?	• Can users easily interpret the purpose of each activity? • Are the expected outcomes of each activity clear? • Are ambiguities or overlaps present between framework components?
Q7. Are the methods and tools recommended within the NuDaK framework suitable for achieving the intended objectives?	• Do the recommended methods adequately support data reuse processes in the nursing context?
Q8. Are the designated expert roles appropriately aligned with the respective NuDaK framework activities?	• Do responsibilities and competencies correspond to the requirements of each activity?

Data Collection

Overview

We combined 3 complementary data collection methods (workshops, internal self-evaluation, and focus group interviews) to address all evaluation questions regarding completeness and applicability of the NuDaK framework. This multimethod approach enabled triangulation of findings and captured both external perspectives (workshops and focus group) and internal methodological reflections (self-evaluation).

Workshops (External Formative Evaluation)

Four workshops were conducted with actively involved project participants to evaluate the applicability of the methods applied in each step and to identify practical barriers, ambiguities, or missing components in the NuDaK framework.

Each workshop included 6 to 10 participants, comprising the director of nursing, ward manager, nursing staff of the piloting station, nursing representatives, IT experts, a staff member of the software company, and an expert in data security and quality management.

Participants were selected based on their active involvement in the piloting steps and their direct experience with NuDaK framework activities.

The workshop questions were adapted to align with the overarching evaluation questions from Table 1, with context-specific modifications for each piloting step.

For example:

- Is the BPMN (Business Process Model and Notation) modeling method suitable for visualizing real documentation processes on the ward?
- Are nursing staff able to use this visualization to assess processes on the ward for completeness and accuracy?
- Are the activities assigned to each NuDaK aspect logically connected and do they follow a coherent sequence?
- Is each activity feasible to perform in practice, given the available resources and time?
- Are the described aspects and activities understandable to end users?

All workshops were audio-recorded and supplemented with written workshop documentation (participant notes, process diagrams, and feedback sheets). Audio files were stored securely and transcribed for qualitative analysis.

Internal Self-Evaluation

Throughout the project, the researcher conducted a structured self-evaluation to assess whether the methods used achieved the goals of each specific aspect and activity of the NuDaK framework. This involved a continuous review process using the project implementation guidelines as a reference framework.

The project guidelines helped identify whether the planned NuDaK activities could be implemented effectively and whether they were applicable and suitable for achieving the goals of each specific aspect and activity. They also helped determine whether any activities were missing, whether they were applied

XSL•FO
RenderX

in the correct sequence, and whether the overall timeline was appropriate. Furthermore, the self-evaluation considered whether the active stakeholders involved in the project had sufficient competencies for their assigned tasks. The researcher maintained continuous self-evaluation notes throughout the study. These notes were stored and served as supplementary data for qualitative analysis.

Focus Group Interview (External Formative Evaluation)

The final online focus group interview (n=5) followed a semistructured guideline and was developed to (1) explore the conceptual understanding and perspectives on the secondary use of routine nursing data, (2) assess the applicability, comprehensibility, and practical relevance of the NuDaK framework, (3) explore challenges, barriers, and facilitating factors in implementing a data strategy for the secondary use of routine data, (4) critically reflect on the composition, roles, and contributions of the stakeholder group, and (5) explore the perceived institutional and organizational added value resulting from the project.

The focus group included 5 project participants (the director of nursing, an IT expert, a member of the software company, and 2 designated nursing representatives). The participants were selected based on their comprehensive involvement across all 4 piloting steps and their ability to provide synthesized perspectives on the framework's overall implementation.

The focus group interview was audio-recorded via the online platform and automatically transcribed.

Data Analysis

Data obtained from workshops and focus group interviews were audio-recorded and transcribed in summary form after each session.

We applied a deductive qualitative content analysis approach based on the 8 predefined evaluation questions (Q1-Q8) from Table 1. The coding process included initial coding based on the evaluation questions, and the analysis was supported by MAXQDA Analytics Pro 2022 (VERBI Software GmbH).

Findings from the internal self-evaluation were triangulated with external qualitative data from the workshops and focus group discussions to identify convergent and divergent assessments, identify refinement needs, and generate recommendations to optimize the NuDaK framework.

Ethical Considerations

This study was conducted in accordance with the standards laid out by the Research Committee for Scientific Ethical Questions. Research Committee for Scientific Ethical Questions is an independent body of the Tyrolean Private University UMIT that examines planned research projects that, due to legal provisions, do not fall under an ethics committee's jurisdiction.

Results

Overview

This section presents the results of the evaluation of the NuDaK framework (2022). Six deductive categories were developed to assess both its applicability and completeness. The categories include definitions and the corresponding evaluation questions, as described in the *Methods* section (see Table 2).

Table . Evaluation categories for completeness and applicability of the NuDaK framework.

Completeness	Applicability
Completeness of framework components (Q1)	Traceability of activities and comprehensibility for users (Q5 and Q6)
Structure, distinctness, and temporal logic of the framework components (Q2)	Suitability of methods and tools (Q7)
Feasibility of NuDaK framework activities in practice (Q3 and Q4)	Roles and responsibilities (Q8)

Completeness of the Framework Components (Q1)

The pilot evaluation indicates that the NuDaK framework's core aspects "Purpose & Benefits," "Dataset & Nursing Documentation," "Data Modelling & Software Requirements," and "Data Integration & Data Quality," cover the essential domains required to plan and implement the secondary use of nursing data. Accordingly, the aspect set was generally perceived as sufficient at the conceptual level.

However, the pilot also showed that several framework-level components require more explicit specification to make the framework complete for practical use, such as (1) explicitly embedding a temporal dimension (phases 1 - 4), (2) defining interdisciplinary stakeholder involvement throughout the phases, and (3) specifying professional roles with role profiles and responsibilities.

In addition, the pilot revealed targeted content gaps within individual aspects. Within "Data Modelling & Software Requirements," an explicit documentation and adaptation concept emerged as necessary to guide systematic changes in the documentation system. Within "Data Integration & Data Quality," additional activities were required, including a strategy workshop to clarify key contextual conditions (eg, technical infrastructure, legal constraints, and staffing) and the development of a data quality concept.

Finally, as a cross-cutting addition, the pilot highlighted the importance of a standalone, iterative information model that is explicitly integrated into the framework. Such a model supports both top-down and bottom-up refinement across dataset development, specification, and revision (see Multimedia Appendix 1). Early and continuous involvement of an interdisciplinary group was considered essential for defining shared objectives and balancing diverging interests. A

XSL•FO
RenderX

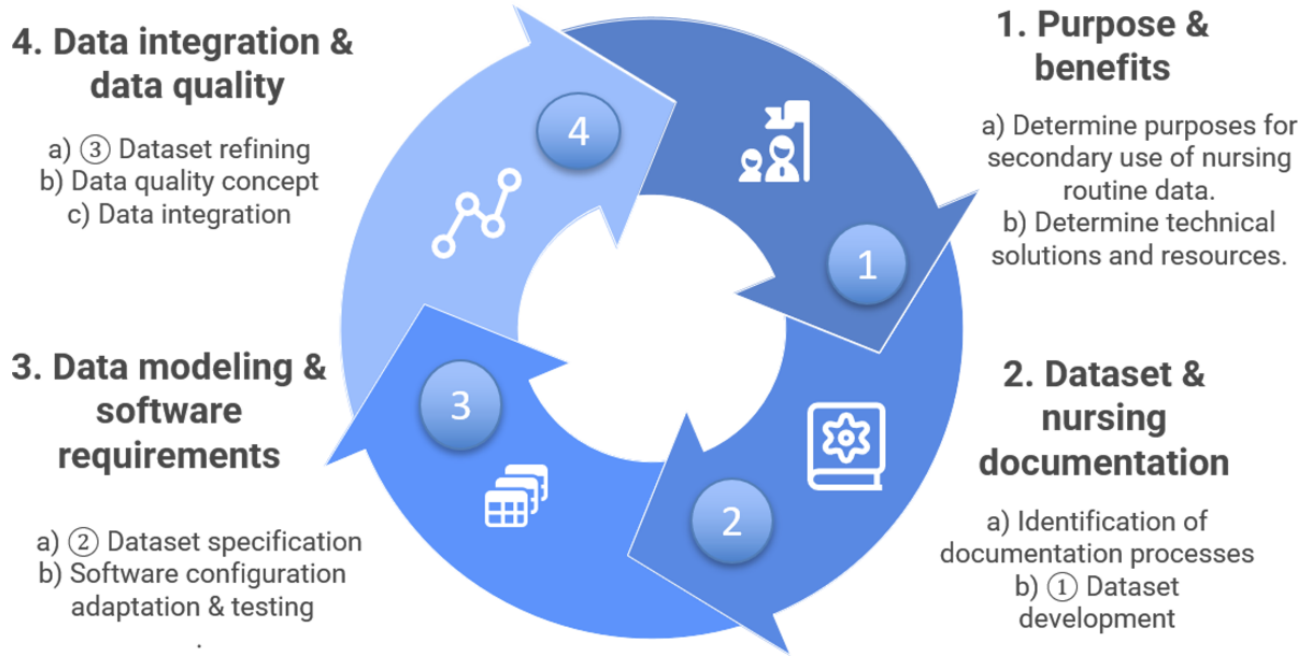
combination of direct nursing practice, nursing management, nursing informatics/nursing development, and local IT was considered effective.

Structure, Distinctness, and Temporal Logic of the Framework Components (Q2)

The evaluation showed that the original linear and sequential representation of activities did not align well with practical requirements. Instead, it revealed a circular logic with recurring activities across multiple phases. Consequently, activities should not only be arranged in a single sequence but must also be distinguished by their temporal nature as either chronological key activities or iterative, cross-aspect activities.

In response, the revised NuDaK framework adopts a circular design comprising 4 phases, each assigned to an overarching aspect (see Figure 3). This structure makes iterative loops explicit and distinguishes between phase-specific key activities and recurring, cross-aspect activities.

Figure 3. NuDaK nursing data strategy framework 2025, representing aspects and activities.



Moreover, the pilot revealed that some activities must be scheduled more precisely to avoid late rework and loss of focus. A strategy workshop to clarify technical infrastructure, legal constraints, and human resources should be anchored early and assigned to “Purpose & Benefits” (phase 1). Likewise, a strategy for data querying and downstream processing should ideally be available by the end of phase 2 or, at the latest, at the beginning of phase 3. Finally, participants indicated that, in phase 4, the purpose and objectives of secondary use often require renewed clarification to maintain the intended focus of the data analysis.

Feasibility in Practice in Terms of Resources and Time (Q3 and Q4)

Overall, the evaluation suggests that the NuDaK framework activities are feasible to perform in practice. Nevertheless, feasibility is conditional on (1) an adequate time budget, (2) clearly assigned responsibilities, (3) access to relevant expertise, and (4) sufficient technical infrastructure.

At the activity level, participants described workshops, process modeling, and iterative concept development (eg, defining data purposes, modeling documentation processes, and specifying data quality criteria) as realistic and implementable when key stakeholders can be convened. In contrast, activities linked to implementation in the documentation system were perceived as particularly resource-intensive, especially testing adaptations of new applications, coordinating change management, and training staff after adaptations (including practical materials such as handouts and explicit clarification of “who does what”).

Participants repeatedly reported that more time would have been beneficial, indicating that the planned time requirements of individual steps were perceived as tight in the pilot setting. In addition, the pilot underscored that sustainable continuation of a data strategy requires building organizational know-how over time, for example, through collaborations and structured training pathways (eg, university partnerships).

Taken together, the findings support making resource requirements and time demands more explicit within the framework by indicating resource-critical activities (eg, documentation system adaptations, data harmonization, and training) and by providing more realistic time estimates or planning guidance for key steps.

Traceability of Activities and Comprehensibility for Users (Q5 and Q6)

The evaluation showed that the NuDaK framework is generally applicable in practice. Participants reported that they could independently develop a data strategy based on the framework, provided that the data purpose and supporting roles are defined early within an interdisciplinary setting.

Comprehensibility and traceability increased throughout the pilot project. However, at the outset, the framework was

XSL•FO
RenderX

perceived as insufficiently concrete. Participants specifically noted that additional start-up input would have been helpful, including concrete examples and use cases.

A conceptual hurdle regarding the meaning of secondary use and the logic of routine data became apparent. Participants described the question "How can we benefit from the data and analyses?" as difficult to answer, and clarification of what secondary use entails emerged only during discussion. In addition, nursing managers and nursing developers often framed the topic in documentation-oriented terms and immediately sought documentation changes to enable analyses, which can conflict with the logic of routine data. These observations indicate that early alignment and clarification, supported by examples, are required from the outset.

Suitability of Methods and Tools (Q7)

Overall, the results suggest that the methods and tools recommended in the NuDaK framework are suitable for achieving the intended objectives and enabling the secondary use of nursing data. Workshops were described as central platforms for strategic decision-making, and audio recording was highlighted as a helpful documentation instrument to secure outcomes.

At the same time, the pilot indicated that methods often need to be adapted in application and, in some cases, applied repeatedly to achieve the desired results. Within "Dataset & Nursing Documentation," developing scenarios for documentation processes, BPMN modeling, and evaluation with nursing practitioners were perceived as effective and goal-oriented. BPMN modeling was considered particularly useful because it is understandable to nursing staff and provides a transparent basis for identifying software-related adaptation needs and relevant data sources.

For "Data Integration & Data Quality," the 2-step approach to developing the data quality concept was considered appropriate: selecting suitable quality dimensions from the literature (eg, completeness, concordance, plausibility, and timeliness) and subsequently operationalizing and validating them in a workshop with IT and nursing experts using practical examples and testing criteria. Moreover, data preparation requires continuous feedback from nursing staff, and harmonizing concepts (data objects) into a uniform data format is indispensable for larger datasets. In summary, the recommended methods most effectively support the intended objectives when applied iteratively and complemented by concrete examples/use cases and appropriate documentation instruments (eg, audio recordings).

Roles and Responsibilities (Q8)

The pilot implementation clearly demonstrated that professional roles, competencies, and responsibilities must be explicitly embedded within the framework, as they are essential for carrying out the activities.

Nursing informatics expertise emerged as central. Implementing a data strategy requires strong competencies in nursing informatics, and nursing informatics specialists should assume the lead and coordinating role. In addition, competencies in data science/data mining and data warehousing, as well as knowledge of programming languages and databases, must be combined with nursing/medical domain knowledge to interpret data correctly and understand nursing processes.

The designation of an independent data administrator role was identified as advantageous to ensure that data are available in the required format and granularity, particularly in phase 3 of implementation. Furthermore, topic-specific nursing expertise in the early phases (eg, a "pain nurse") was considered essential to align activities with nursing requirements.

To support practical implementation, responsibilities should be conceptualized not only as individual roles but also as clearly defined interdisciplinary expert groups for each phase to enable joint strategic alignment, holistic requirements consideration, and the balancing of diverging interests.

NuDaK Framework 2025: A Data Strategy to Reuse Nursing RWD

The evaluation of the applicability and completeness of the NuDaK Framework 2022 provided essential insights for its further development and refinement. Building on these findings, the framework was systematically adapted. In the following section, we present a short version of the updated NuDaK Framework 2025. A comprehensive description of the NuDaK Framework 2025, including detailed activities, methods, and role descriptions, is provided in Multimedia Appendix 1.

Major changes to the updated NuDaK framework comprise a clear temporal structure (phases 1 - 4). Furthermore, it is now conceptualized as a combination of chronological key activities and iterative, cross-aspect activities. In addition, the framework was strengthened by defining interdisciplinary stakeholder involvement, specifying professional roles, responsibilities, and competencies, and highlighting key prerequisites for feasibility.

Alongside this, a standalone, iterative information model was integrated to continuously improve the dataset (see Multimedia Appendix 1).

The NuDaK framework 2025 is a practice-oriented data strategy that structures the secondary use of routinely collected nursing data into 4 temporal phases and 4 corresponding aspects: (1) Purpose & Benefits, (2) Dataset & Nursing Documentation, (3) Data Modeling & Software Requirements, and (4) Data Integration & Data Quality (Figure 3). In addition to these aspects, the framework specifies activities and associated methods. Across all phases and aspects, implementation is enabled by defined professional roles and responsibilities (eg, nursing informatics specialist, IT professionals, data administrator, ward management, and domain experts) that support the implementation of NuDaK.

In *Phase 1 (Purpose & Benefits*), organizations define the aims, expected benefits, and scope of secondary data use. Objectives are established through interprofessional consensus workshops and aligned with feasibility considerations (eg, data processing location, technical infrastructure, data science competencies, and regulatory constraints). Key deliverables are a purpose statement and an agreed analytical scope.

XSL•FO
RenderX

Phase 2 (Dataset & Nursing Documentation) focuses on understanding documentation practices and deriving a robust dataset. Documentation workflows are captured using scenario-based analysis and process modeling (eg, BPMN), followed by expert review. In parallel, the dataset is developed as the first stage ① of the *information model* using a top-down approach (a literature-informed theoretical dataset), refined against real documentation through use cases, and validated by domain experts.

In *Phase 3 (Data Modeling & Software Requirements)*, the dataset is formally specified (information model stage ② [dataset specification]) by standardizing elements, defining attributes and permissible values, and adding missing fields as required. These specifications are translated into software requirements, followed by system configuration, user-centered adaptation, training support, and a structured testing period to ensure applicability and completeness of data capture.

Phase 4 (Data Integration & Data Quality) operationalizes information model stage ③ (dataset refinement) and establishes prerequisites for valid analyses. Iterative test queries are used to verify availability, formats, value ranges, and data types across sources. A data quality concept is developed by selecting core quality dimensions (eg, completeness, concordance, plausibility, and currency) and defining practical evaluation examples and validation procedures with clinical and technical experts. Finally, data integration consolidates validated data into a unified, analysis-ready structure while ensuring traceability, governance, and alignment with refined research questions. Further details are provided in Multimedia Appendix 1.

Discussion

Principal Results

This feasibility study evaluated the completeness and applicability of the NuDaK framework (2022) in a real-world nursing setting and used the findings to refine the framework into NuDaK 2025. Overall, the core set of aspects was confirmed as sufficient to cover key domains required for the secondary use of nursing routine data. However, NuDaK implementation showed that a linear representation does not reflect the iterative character of data strategy work in clinical environments.

Consequently, NuDaK 2025 now adopts a circular structure with 4 temporal phases and makes feedback loops explicit. In addition, the pilot indicated the need to (1) embed an explicit temporal dimension (phases 1 - 4), (2) formalize professional roles and responsibilities, and (3) introduce a standalone, iterative information model spanning dataset development, specification, and refinement across phases 2 to 4. Taken together, these refinements indicate that feasibility in this context depends not only on the content of the framework, but also on its ability to reflect the actual workflow logic of clinical implementation.

Results in Relation to Other Frameworks

NuDaK 2025 offers a nursing-specific, end-to-end strategy that connects purpose definition, dataset and documentation work, technical specification and software adaptation, and, finally, data integration and quality assurance. This distinguishes it from frameworks that address only selected parts of the reuse pathway.

The SPIRIT framework offers valuable guidance for organizational planning and system-level preparation, but it is not tailored to nursing purposes and does not operationalize the pathway from documentation to a reusable nursing dataset in the same level of detail [16].

Likewise, the Health Data Quality Framework is for assessing and visualizing EHR quality but does not provide an integrated life cycle strategy for the secondary use of nursing data that includes documentation processes and software implementation [33]. Nursing data mining approaches, including predictive modeling, focus more strongly on analytic reuse and less comprehensively on the upstream prerequisites, such as governance, integration, and quality management, that are needed before advanced analysis can take place [34].

Against this background, NuDaK extends existing approaches by combining strategic, organizational, technical, and methodological elements within one nursing-specific framework.

It covers the entire process from nursing documentation to a usable dataset for nursing analytics in a practical, step-by-step manner.

Methodological Implications

This exploratory study also shows methodological implications for the development and evaluation of complex frameworks in nursing informatics. First, it shows that formative feasibility work can reveal structural requirements that may not be fully visible at the conceptual design stage, especially when the framework is tested in real clinical practice. Second, the findings suggest that framework completeness should be interpreted not only as coverage of content domains but also as clarity of sequence, timing, roles, and feedback mechanisms. Third, the study illustrates the value of combining external qualitative evaluation with internal self-evaluation because these perspectives complement one another and help identify both practical barriers and areas for refinement.

Meaning and Generalizability

NuDaK 2025 provides a structured pathway for health care organizations and informatics specialists to establish the secondary use of routine nursing data. In particular, it can be used by nursing managers, hospital directors, or health informatics experts to plan, build, and implement the use of routine nursing data to improve nursing practice and support nursing research.

Its practical relevance lies in addressing common barriers that limit the reuse of nursing data. NuDaK particularly addresses insufficient standardization, variable data quality, and missing governance across the documentation-to-dataset pathway.

By aligning technical specification, software adaptation, and data quality assurance with nursing workflows, NuDaK can support robust analyses of large nursing datasets and thereby strengthen nursing quality management and evidence generation.

XSL•FO
RenderX

In the longer term, it may also contribute to establishing the data prerequisites for advanced analytics and AI-related applications. However, the implementation of such downstream applications requires additional technical, organizational, and methodological work beyond the scope of the present feasibility study. Given that AI applications in nursing rely on standardized, complete, and high-quality datasets, NuDaK may provide a relevant methodological foundation [10].

Overall, NuDaK 2025 should be understood as a strategy framework that structures the secondary use of routinely collected nursing data for quality improvement, research, and future analytic development.

Strengths and Limitations

A key strength of this feasibility study is that NuDaK was evaluated in a real clinical nursing setting and across multiple implementation phases, which allowed practical barriers and refinement needs to emerge from actual use rather than from theoretical reflection alone. The combination of workshops, focus group data, and internal self-evaluation also provided complementary perspectives on the same framework.

This study has several limitations. It was conducted in a single hospital ward, so transferability to other organizations with different workflows, documentation systems, or nursing terminologies is limited. In addition, the findings are based on qualitative perceptions and researcher self-evaluation rather than objective outcome measures, so they should be interpreted as exploratory. Future studies should therefore examine NuDaK in additional settings and complement qualitative findings with more objective implementation and outcome data.

Conclusions

Building on the need to better leverage routinely collected nursing documentation data, we developed the NuDaK framework as a practical guide for health organizations and research teams. Refined through a feasibility study in a real-world nursing setting, NuDaK now offers a nursing-specific, end-to-end data strategy across 4 phases, with explicit feedback loops, defined professional roles, and an iterative information model. The present feasibility study supports the conceptual refinement of the framework.

However, several issues remain only partially resolved. Further research is needed to evaluate the performance and transferability of the framework across heterogeneous institutions with different documentation systems, governance arrangements, vendors, and levels of digital maturity. Multisite studies are also needed to examine data harmonization and interoperability requirements across settings and software solutions. Future studies should also validate NuDaK through concrete nursing use cases and exemplar analyses while assessing its longitudinal sustainability and impact in routine practice.

Acknowledgments

The authors gratefully acknowledge all experts and practice partners for their valuable feedback and contributions to piloting and evaluating the NuDaK framework. We also thank the participating health care organization and the software company for their support in the feasibility study and technical implementation.

Generative AI Use

During the preparation and revision of this manuscript, the authors used ChatGPT (GPT-6.1 Sol; OpenAI) to support language editing, improve clarity and readability, and refine the structure and wording of selected passages. Napkin AI was used to support the graphical visualization of concepts. Perplexity AI and Elicit were used to assist with literature searches and the identification of relevant scientific publications. All AI-assisted outputs were critically reviewed and verified by the authors, who take full responsibility for the accuracy, originality, and integrity of the manuscript, including all references and citations.

Funding

The authors declared that no financial support was received for this work.

Authors' Contributions

Conceptualization: RN (lead)
Data curation: RN
Formal analysis: RN (lead), EA (supporting)
Investigation: RN (lead), EA (supporting)
Methodology: RN (lead), EA (supporting)
Project administration: RN
Visualization: RN
Writing – original draft: RN
Writing – review & editing: RN (lead), EA (supporting).

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
NuDaK nursing data strategy framework 2025.
[DOCX File, 298 KB - nursing_v9i1e95546_app1.docx]

References

1. Hübner U, Liebe JD, Bicker L, et al. IT-Report Gesundheitswesen: Wie veränderte sich die Gesundheits-IT aus User-Perspektive? [Report in German]. : Hochschule Osnabrück; 2024 URL: https://www.hs-osnabrueck.de/fileadmin/HSOS/Homepages/IT-Report_Gesundheitswesen/IT_Report_2024.pdf [accessed 2026-11-09]
2. Nantschev R, Ammenwerth E. Challenges using electronic nursing routine data for outcome analyses: a mixed methods study. Int J Nurs Sci 2022 Jan;9(1):92-99. [doi: 10.1016/j.ijnss.2021.11.003] [Medline: 35079610]
3. National trends in hospital and physician adoption of electronic health records. Health IT Research & Analysis. 2026. URL: https://healthit.gov/data/quickstats/national-trends-hospital-and-physician-adoption-electronic-health-records/ [accessed 2026-02-22]
4. Colman BD, Zhu Z, Qi Z, van der Walt A. From real world data to real world evidence to improve outcomes in neuro-ophthalmology. Eye (Lond) 2024 Aug;38(12):2448-2456. [doi: 10.1038/s41433-024-03160-8] [Medline: 38844583]
5. Nahm ES, Zhu S, Seidl K, Chen L, Day J, Seong H. Real-world data for interdisciplinary health care research: a case example. ANS Adv Nurs Sci 2023;46(4):349-362. [doi: 10.1097/ANS.0000000000000496] [Medline: 37102714]
6. McGrow K. Artificial intelligence in nursing: a journey from data to wisdom. Nursing 2025 Apr 1;55(4):16-24. [doi: 10.1097/NSG.0000000000000165] [Medline: 40122866]
7. Westra BL, Christie B, Johnson SG, et al. Modeling flowsheet data to support secondary use. Comput Inform Nurs 2017 Sep;35(9):452-458. [doi: 10.1097/CIN.0000000000000350] [Medline: 28346243]
8. Raju D, Su X, Patrician PA, Loan LA, McCarthy MS. Exploring factors associated with pressure ulcers: a data mining approach. Int J Nurs Stud 2015 Jan;52(1):102-111. [doi: 10.1016/j.ijnurstu.2014.08.002] [Medline: 25192963]
9. Khokhar A, Lodhi MK, Yao Y, Ansari R, Keenan G, Wilkie DJ. Framework for mining and analysis of standardized nursing care plan data. West J Nurs Res 2017 Jan;39(1):20-41. [doi: 10.1177/0193945916672828] [Medline: 27756852]
10. Wei Q, Pan S, Liu X, Hong M, Nong C, Zhang W. The integration of AI in nursing: addressing current applications, challenges, and future directions. Front Med (Lausanne) 2025;12:1545420. [doi: 10.3389/fmed.2025.1545420] [Medline: 40007584]
11. Safran C. Update on data reuse in health care. Yearb Med Inform 2017 Aug;26(1):24-27. [doi: 10.15265/IY-2017-013] [Medline: 29063535]
12. Vuokko R, Mäkelä-Bengs P, Hyppönen H, Lindqvist M, Doupi P. Impacts of structuring the electronic health record: results of a systematic literature review from the perspective of secondary use of patient data. Int J Med Inform 2017 Jan;97:293-303. [doi: 10.1016/j.ijmedinf.2016.10.004] [Medline: 27919387]
13. Topaz M, Pruinelli L. Big data and nursing: implications for the future. Stud Health Technol Inform 2017;232:165-171. [Medline: 28106594]
14. Macieira TGR, Chianca TCM, Smith MB, et al. Secondary use of standardized nursing care data for advancing nursing science and practice: a systematic review. J Am Med Inform Assoc 2019 Nov 1;26(11):1401-1411. [doi: 10.1093/jamia/ocz086] [Medline: 31188439]
15. Liljamo P, Kinnunen UM, Saranto K. Assessing the relation of the coded nursing care and nursing intensity data: towards the exploitation of clinical data for administrative use and the design of nursing workload. Health Informatics J 2020 Mar;26(1):114-128. [doi: 10.1177/1460458218813613] [Medline: 30516092]
16. Hackl WO, Ammenwerth E. SPIRIT: Systematic Planning of Intelligent Reuse of Integrated Clinical Routine Data: a conceptual best-practice framework and procedure model. Methods Inf Med 2016;55(2):114-124. [doi: 10.3414/ME15-01-0045] [Medline: 26769124]
17. Leiner F, Gaus W, Haux R, Knaup-Gregori P. Medical Data Management: A Practical Guide: Springer; 2003.
18. Afaneh T, Abu-Moghli F, Ahmad M. Nursing-sensitive indicators: a concept analysis. Nurs Manag (Harrow) 2021 Jun 1;28(3):28-33. [doi: 10.7748/nm.2021.e1982] [Medline: 33829713]
19. Blume KS, Dietermann K, Kirchner-Heklau U, et al. Staffing levels and nursing-sensitive patient outcomes: umbrella review and qualitative study. Health Serv Res 2021 Oct;56(5):885-907. [doi: 10.1111/1475-6773.13647] [Medline: 33723857]
20. Heslop L, Lu S, Xu X. Nursing-sensitive indicators: a concept analysis. J Adv Nurs 2014 Nov;70(11):2469-2482. [doi: 10.1111/jan.12503] [Medline: 25113388]
21. Koch D, Kutz A, Volken T, et al. Derivation and validation of a prediction model to establish nursing-sensitive quality benchmarks in medical inpatients: a secondary data analysis of a prospective cohort study. Swiss Med Wkly 2022;152:w30152. [doi: 10.4414/smw.2022.w30152] [Medline: 35315266]
22. Liu Y, Avant KC, Aungsuroch Y, Zhang XY, Jiang P. Patient outcomes in the field of nursing: a concept analysis. Int J Nurs Sci 2014 Mar;1(1):69-74. [doi: 10.1016/j.ijnss.2014.02.006]
23. Oner B, Zengul FD, Oner N, Ivankova NV, Karadag A, Patrician PA. Nursing-sensitive indicators for nursing care: a systematic review (1997-2017). Nurs Open 2021 May;8(3):1005-1022. [doi: 10.1002/nop2.654] [Medline: 34482649]

XSL•FO
RenderX

24. Stalpers D, van der Linden D, Kaljouw MJ, Schuurmans MJ. Using publicly reported nursing-sensitive screening indicators to measure hospital performance: the Netherlands experience in 2011. Nurs Res 2016;65(5):362-370. [doi: 10.1097/NNR.0000000000000170] [Medline: 27579504]
25. Tuinman A, De Greef MHG, Finnema EJ, Roodbol PF. A systematic review of the association between nursing staff and nursing-sensitive outcomes in long-term institutional care. J Adv Nurs 2021 Aug;77(8):3303-3316. [doi: 10.1111/jan.14840] [Medline: 33764569]
26. Twigg DE, Kutzer Y, Jacob E, Seaman K. A quantitative systematic review of the association between nurse skill mix and nursing-sensitive patient outcomes in the acute care setting. J Adv Nurs 2019 Dec;75(12):3404-3423. [doi: 10.1111/jan.14194] [Medline: 31483509]
27. Guidance on a data quality framework for health and social care. : Health Information and Quality Authority (HIQA); 2018 URL: https://www.hiqa.ie/sites/default/files/2018-10/Guidance-for-a-data-quality-framework.pdf [accessed 2026-11-03]
28. Westra BL, Johnson SG, Ali S, et al. Validation and refinement of a pain information model from EHR flowsheet data. Appl Clin Inform 2018 Jan;9(1):185-198. [doi: 10.1055/s-0038-1636508] [Medline: 29539649]
29. Berner S, Möller J, Müller S. Whitepaper information modelling: a fundamental tool for leaders in the digital change. : foryouandyourcustomers; 2019 URL: https://foryouandyourcustomers.com/insights/whitepaper/information-modelling/ [accessed 2025-10-13]
30. Weiskopf NG, Bakken S, Hripcsak G, Weng C. A data quality assessment guideline for electronic health record data reuse. EGEMS (Wash DC) 2017 Sep 4;5(1):14. [doi: 10.5334/egems.218] [Medline: 29881734]
31. Kahn MG, Callahan TJ, Barnard J, et al. A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. EGEMS (Wash DC) 2016;4(1):1244. [doi: 10.13063/2327-9214.1244] [Medline: 27713905]
32. Bian J, Lyu T, Loiacono A, et al. Assessing the practice of data quality evaluation in a national clinical data research network through a systematic scoping review in the era of real-world data. J Am Med Inform Assoc 2020 Dec 9;27(12):1999-2010. [doi: 10.1093/jamia/ocaa245] [Medline: 33166397]
33. Johnson SG, Pruinelli L, Hoff A, et al. A framework for visualizing data quality for predictive models and clinical quality measures. AMIA Jt Summits Transl Sci Proc 2019;2019:630-638. [Medline: 31259018]
34. Wieland-Jorna Y, Verheij RA, Francke AL, et al. Setting up a governance framework for secondary use of routine health data in nursing homes: development study using qualitative interviews. J Med Internet Res 2023 Jan 25;25:e38929. [doi: 10.2196/38929] [Medline: 36696162]
35. Nantschev R. NuDaQ- framework: a strategy to reuse electronic routine nursing process data for outcome analyses. Presented at: Nursing-generated data: predictive analytics and eHealth strategies; Mar 16-18, 2023 URL: https://acendio.net/wp-content/uploads/2023/03/Proceedings-ACENDIO_14.03.2023.pdf [accessed 2026-11-03]
36. Nantschev R. Vom datenfriedhof zur datennutzung: sekundärnutzung von pflegerischen routinedaten – chancen der digitalisierung [Article in German]. Österreichische Pflegezeitschrift 2021 [FREE Full text]
37. Brender J, Talmon J, de Keizer N, Nykänen P, Rigby M, Ammenwerth E. STARE-HI - Statement on Reporting of Evaluation Studies in Health Informatics: explanation and elaboration. Appl Clin Inform 2013;4(3):331-358. [doi: 10.4338/ACI-2013-04-RA-0024] [Medline: 24155788]
38. Döring N, Bortz J. Forschungsmethoden Und Evaluation in Den Sozial- Und Humanwissenschaften [Book in German], 5th edition: Springer; 2016. [doi: 10.1007/978-3-642-41089-5]

Abbreviations

BPMN: Business Process Model and Notation
EHR: electronic health record
RWD: real-world data
SPIRIT: Systematic Planning of Intelligent Reuse of Integrated Clinical Routine Data
STARE-HI: Statement on Reporting of Evaluation Studies in Health Informatics

Edited by E Borycki; submitted 17.Mar.2026; peer-reviewed by G Gioiello, KD Groot; revised version received 14.Sep.2026; accepted 21.Oct.2026; published 16.Nov.2026.

Please cite as:
Nantschev R, Ammenwerth E
A Data Strategy for the Secondary Use of Nursing Real-World Data: Feasibility Study of the NuDaK Framework
JMIR Nursing 2026;9:e95546
URL: https://nursing.jmir.org/2026/1/e95546
doi:10.2196/95546

XSL•FO
RenderX

© Renate Nantschev, Elske Ammenwerth. Originally published in JMIR Nursing (https://nursing.jmir.org), 16.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating Human-Robot Interaction With a Socially Assistive Robot in Pediatric Health Care: Mixed Methods Pilot Study

Janika Leoste[1,2], Prof Dr; Slavko Rakić[1,3], Prof Dr; Kristel Marmor[1], MSci; Kadi Krinal[4], PhD; Liina-Harjus Soostar[5], PhD, MEd; Katariina Maribel Vink[1]; Andres Käver[1], MSc; Katrin Aasma[1], BA; Katie Beatty[6]; Leo Benjamin Leoste[1]

[1]IT College, Tallinn University of Technology, Ehitajate Tee 5, Tallinn, Harjumaa, Estonia

[2]School of Educational Sciences, Tallinn University, Tallinn, Harjumaa, Estonia

[3]Faculty of Technical Sciences, University of Novi Sad, Novi Sad, Serbia

[4]Centre for Digital Health, Department of Health Technologies, Tallinn University of Technology, Tallinn, Harjumaa, Estonia

[5]Tallinn Children's Hospital, Tallinn, Harjumaa, Estonia

[6]The Ohio State University, Ohio, OH, United States

Corresponding Author:
Janika Leoste, Prof Dr
IT College, Tallinn University of Technology, Ehitajate Tee 5, Tallinn, Harjumaa, Estonia

Abstract

Background: Socially assistive robots are increasingly explored in health care, but evidence from routine pediatric outpatient settings remains limited.

Objective: This study examined how contextual support and human facilitation shaped visitor interaction with the Temi V3 robot and which sociotechnical conditions were associated with practical acceptance according to staff.

Methods: A mixed methods pilot combined structured behavioral observation across baseline, poster support, and staff facilitation conditions with 4 staff focus group interviews. The observation unit was a visitor passage through the lobby zone; the qualitative sample comprised 12 staff members with direct exposure to the robot. Quantitative indicators were analyzed descriptively, and staff interviews underwent deductive-inductive qualitative content analysis.

Results: In total, 675 visitor passages (children: n=398, 59% and adults: n=277, 41%) were recorded (baseline: n=257; poster support: n=160; and staff facilitation: n=258). Looking at the robot increased from 8 (3.1%) passages at baseline to 30 (18.8%) with poster support and 80 (31%) with staff facilitation. Screen or app use increased from 2 (0.8%) passages at baseline to 4 (2.5%) with poster support and 14 (5.4%) with staff facilitation. Observer notes indicated stronger visible responsiveness among children, although age category was not linked to individual behavior codes. Staff emphasized role clarity, child and family value, human mediation, functional reliability, multilingual content, and organizational ownership.

Conclusions: Observed visitor interaction was higher with poster support and staff facilitation than at baseline. These descriptive findings and staff perceptions informed a provisional 4-domain framework comprising role clarity, interaction affordance, social mediation, and organizational embedding. The framework identifies potential implementation priorities; visitors' own experiences and acceptance require direct assessment.

(JMIR Nursing 2026;9:e95702) doi:10.2196/95702

KEYWORDS

socially assistive robots; human-robot interaction; hospital navigation; mixed methods; technology acceptance

Introduction

Background

Socially assistive robots (SARs) are designed to support people through social interaction without providing direct physical assistance [1,2]. In health care, they have been explored as guides, companions, educators, telepresence platforms, and communication mediators [3-5]. Hospitals are especially relevant settings because they combine logistical complexity with emotional strain, repeated information needs, and heavy staff workload [3,6,7].

Pediatric outpatient environments add a further layer of complexity. Children often face uncertainty, waiting, unfamiliar spaces, and procedure-related stress, while parents must simultaneously manage navigation, appointments, and the child's emotional state. Prior reviews suggest that social robots can support children's well-being during medical treatment through distraction, engagement, and emotionally lighter interaction, but the evidence remains heterogeneous and strongly context dependent [8,9]. Hospital human-robot interaction (HRI)

XSL•FO
RenderX

reviews likewise show that pediatric settings are promising, while many studies remain short-term, small-scale, or conducted under controlled conditions instead of routine clinical conditions [3,7].

Implementation is an additional challenge. Health care robot adoption is rarely determined by technical capability alone. Acceptance depends on whether users understand what the robot is for; whether it fits work processes; whether it is easy to use; and whether organizations provide the training, maintenance, and responsibility structures needed for sustained use [10-12]. Recent European survey evidence also suggests that public support for SARs in health care is conditional. People appear more comfortable with bounded assistive roles, such as information provision and guidance, than with higher-stakes gatekeeping or quasi-clinical decision roles, and privacy remains a recurring concern [13].

Study Context

The present study examined a Temi V3 robot in the second-floor outpatient area of Tallinn Children's Hospital in Tallinn, Estonia. The Temi V3 is a mobile platform with screen-based interaction, speech output, and navigation capabilities [14]. The robot was introduced as a public-facing assistive technology for wayfinding, simple informational support, and child-friendly engagement during waiting. Related earlier work by the research team examined scenarios for telepresence robots in health care [15]; the present manuscript reports a separate 2025 field pilot focused on a SAR in a pediatric outpatient setting.

Theoretical Background

SARs in Hospitals and Pediatric Care

The health care robotics literature includes several overlapping categories, including service robots, social robots, SARs, companion robots, and telepresence robots [2,4,5]. The term SAR is the most relevant term for the robot used in the study (Temi V3), as it was not intended to perform invasive or physically assistive clinical work. Its intended value lay in communication, guidance, and emotionally supportive interaction, which aligns with assistance through social engagement [1].

Systematic reviews show that hospital SARs have been implemented in reception, wayfinding, education, rehabilitation, companionship, telepresence, and administrative support [3,4,7]. Broader surveys similarly suggest that near-term adoption is most plausible in bounded domains such as communication, navigation, logistics, telepresence, and patient-facing service functions, where robots can add value without assuming clinical agency [6]. Many hospital robots therefore occupy an intermediate position: they are neither entertainment devices nor autonomous care professionals, but hybrid systems whose usefulness depends on how well they extend existing human service processes.

In pediatric care, published work has frequently focused on play, distraction, procedure support, emotional comfort, and engagement during waiting or treatment. Moerman et al [8] concluded that social robots may support children's well-being, especially by reducing distress and stimulating positive engagement, while also noting that much of the evidence comes from small or highly contextual studies. Kabacińska et al [9] found that children, parents, and health care providers may all see value in pediatric hospital robots, while emphasizing workflow fit, co-design, and responsible implementation.

The field nevertheless remains uneven. Reviews report positive interaction outcomes alongside recurring friction points, including technical unreliability, overestimated autonomy, narrow interaction repertoires, poor personalization, and weak integration into the social context of care [5,7,10]. A robot's observed use in a hospital cannot therefore be understood as a property of the machine alone. It also reflects the information environment, user expectations, staff mediation, physical setting, and local routines.

Acceptance and Implementation of Health Care Robots

Technology acceptance research offers a useful starting point. The technology acceptance model (TAM) identifies perceived usefulness and perceived ease of use as central drivers of adoption [16]. The UTAUT (unified theory of acceptance and use of technology) adds social influence and facilitating conditions, which are particularly relevant when technologies are introduced through organizational channels [17]. Staff and visitors assess technical capability together with usefulness, understandability, approachability, and visible social endorsement.

For social robots, the Almere model extends acceptance thinking by incorporating trust, enjoyment, social presence, and anxiety [18]. Although developed in work with older adults, its broader insight is relevant here: acceptance of a social agent reflects both functional evaluation and the felt quality of interaction. In pediatric contexts, playful engagement, attention redirection, and perceived warmth may matter alongside instrumental usefulness [8,9].

Acceptance models alone do not capture implementation fully. Reviews of SARs in health and social care identify insufficient training, unclear role boundaries, additional workload, technical unreliability, privacy and safety concerns, weak customization, and the absence of local champions or maintenance structures as recurring barriers [10,11,19]. Stakeholder-oriented design studies similarly indicate that acceptability improves when robot functions are visibly useful, easy to understand, robust in busy environments, and aligned with actual care needs [11].

Acceptance is also multiactor. Children, parents, frontline nurses, administrative staff, technical staff, and managers may evaluate the same robot differently because they encounter different benefits, burdens, and risks. Early interest may coexist with reservations about accountability, privacy, maintenance, and role boundaries [10,13,19]. Qualitative work with nursing professionals likewise indicates that robots gain legitimacy when attached to concrete support tasks and when organizational conditions for integration are explicit [12,20].

Analytical Framework and Study Aim

Drawing these strands together, the study used 4 sensitizing domains: role clarity, interaction affordance, social mediation, and organizational embedding. Role clarity concerns whether

XSL•FO
RenderX

the robot has an obvious and locally meaningful purpose. Interaction affordance concerns whether visitors and staff can understand how to start, continue, benefit from, and recover from interaction. Social mediation concerns whether signage, demonstration, and visible endorsement reduce uncertainty and invite use. Organizational embedding concerns workflow, training, support, maintenance, content management, and responsibility. The 4 domains were developed as an author synthesis of TAM, UTAUT, the Almere model, and implementation evidence [10,11,16-20]. They guided deductive coding and mixed methods interpretation, while inductive analysis remained open to recurrent issues raised by staff. The framework is provisional and is not presented as a validated scale.

The aim of the study was to examine how contextual support shaped observable visitor interaction with Temi V3 and which sociotechnical conditions staff associated with its practical acceptance in pediatric outpatient care. The following research questions (RQs) guided the study:

- RQ1: How did the 4 observed interaction indicators (looking at the robot, approaching it, touching it, and using the robot's screen or app) vary across baseline, poster support, and staff facilitation conditions?
- RQ2: Which sociotechnical factors did hospital staff identify as enabling or constraining practical acceptance and integration of the robot?

Methods

Research Design and Integration

The study used a mixed methods pilot design with separate quantitative and qualitative analyses and convergent interpretation [21]. The quantitative strand examined spontaneous visitor interaction with the robot under 3 sequential implementation conditions. The qualitative strand explored how staff perceived the robot's usefulness, limits, and implementation requirements through small focus group interviews. The strands were analyzed separately and then interpreted together to relate observable visitor behavior to staff perspectives on implementation, without treating staff accounts as evidence of visitors' own perceptions.

This design was selected because unobtrusive observation was needed to document what people actually did near the robot under routine hospital conditions, while interviews were needed to understand how staff interpreted the robot's role, usefulness, support burden, and fit with everyday work. The study did not link the 2 datasets at the person level. It used them as complementary evidence about immediate visitor uptake and organizational implementation.

The participant groups were intentionally different. The observation strand included naturalistic visitor passages, whereas the qualitative strand included employees with direct experience of using, observing, supporting, or coordinating the robot. Children, parents or guardians, and other adult visitors were not interviewed; claims about these groups are therefore limited to observed behavior and staff perceptions. RQ1 was addressed through condition-level behavior counts, observer-assigned age categories, and contemporaneous field notes, and RQ2 through staff interviews. Integration occurred after the separate analyses.

Setting and Robot

The study took place in the second-floor outpatient lobby at Tallinn Children's Hospital. Families in this area navigate several specialties and may encounter uncertainty about room locations, changing numbering systems, and waiting routines. The robot's intended public-facing functions included wayfinding, basic hospital information, and child-friendly engagement during waiting.

The robot's pilot interface combined hospital-specific and generic elements in the national language (Estonian). Mapped hospital destinations and selected hospital information were available alongside generic Temi content, including news and simple entertainment. The robot's hardware, mapped destinations, screen content, and physical location remained unchanged across the 3 observation conditions. The experiment's conditions differed only in the external contextual support provided.

Quantitative Strand: Structured Behavioral Observation

Participants, Sampling, and Eligibility

The unit of analysis was a visitor passage through a predefined lobby observation zone during the scheduled observation periods. Consecutive passages were included when a person entered the zone within an observation window. Passages outside the zone or outside the scheduled period were excluded. No further demographic or behavioral eligibility criteria were applied because the purpose was to document naturalistic visitor flow. Observation was kept nonintrusive and anonymous, so repeat passages by the same person could not be identified or excluded. The reported totals represent passages and cannot be interpreted as counts of unique visitors. Observers assigned a descriptive age category of child or adult. Age category was not linked to each individual behavior code in the aggregated dataset. Age-related findings are therefore based on phase-level counts and field notes, and no statistical comparison of child and adult engagement rates was possible.

Implementation Conditions and Observation Protocol

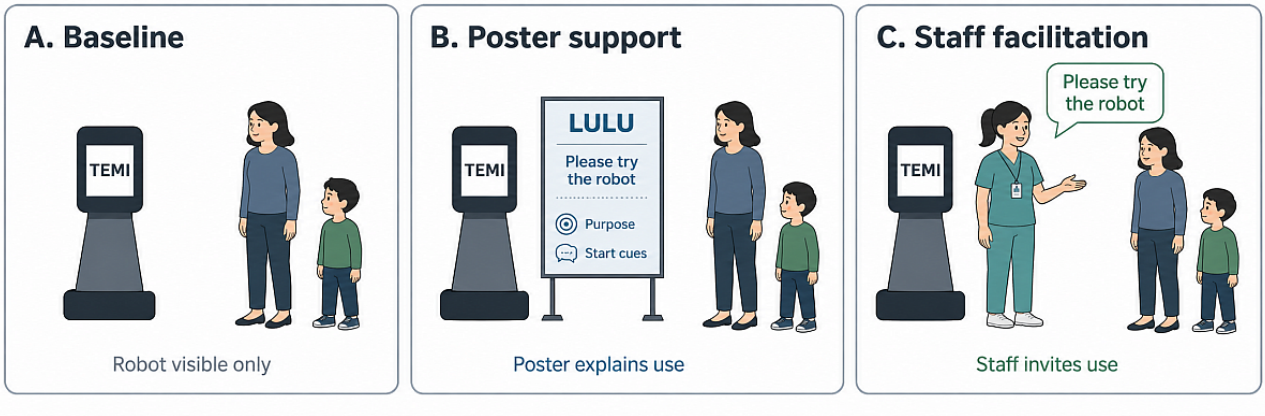
Observation occurred under 3 sequential conditions (Figure 1): baseline (May 22 - 23, 2025), poster support (June 5 - 6, 2025), and staff facilitation (June 12 - 13, 2025). At baseline, the robot was available without added contextual support. In the poster condition, a sign introduced the robot's name, purpose, and simple interaction cues. In the staff facilitation condition, a hospital intern stood near the robot and verbally invited visitors to interact.

Two observers monitored the lobby using a predefined protocol. One focused on entries and exits, while the other recorded visible interaction near the robot. Engagement was operationalized as 4 progressive but separately tallied indicators: looking at the robot, approaching it, touching it, and using the screen or app. A passage could contribute to more than 1 indicator, so the indicators were not mutually exclusive. The 4

XSL•FO
RenderX

indicators were retained as separate outcomes and were not averaged across conditions.

Figure 1. Three sequential implementation conditions in the pediatric outpatient lobby. The robot and its location remained unchanged; only the contextual support differed.



Quantitative Analysis

Frequencies and percentages were calculated separately for each behavior indicator and condition. As an additional descriptive summary, the 4 indicator counts were summed within each condition, divided by the number of visitor passages in that condition, and multiplied by 100. This summary represents recorded indicators per 100 passages, not the percentage of passages with any interaction or a measure of acceptance, because the indicators may overlap. The analysis was descriptive because the field conditions were sequential, nonrandomized, and conducted on different dates. No inferential tests or *P* values were calculated. The quantitative results were compared with staff themes during mixed methods interpretation.

Qualitative Strand: Staff Focus Groups

Participants and Sampling

The qualitative strand comprised 12 health care professionals interviewed in 4 small focus group sessions. Sampling was purposive and criterion based. Inclusion required employment at Tallinn Children's Hospital and direct exposure to the robot through use, observation, support, or deployment-related coordination. The available role information represented frontline clinical and nursing perspectives as well as coordination-level and senior organizational perspectives. Staff without direct robot exposure, hospital visitors, children, and parents or guardians were excluded.

To protect employees from possible deductive identification within the relatively small hospital workforce, detailed demographic data, including exact age, gender, tenure, and job title, were not collected. The manuscript therefore reports only the available broad role-level information. The limited demographic detail is acknowledged as a study limitation, and a structured summary of the available sample characteristics is provided in Multimedia Appendix 1.

Interview Procedure and Analysis

The interview guide covered opportunities for robot use, factors supporting and hindering use, suitable functions in interaction with children and parents, and organizational needs connected to implementation. Participation was voluntary. Sessions were audio-recorded, transcribed using automated speech recognition, manually corrected, deidentified, and stored on institutional infrastructure for the approved research purpose.

A deductive-inductive qualitative content analysis was used [22,23]. Deductive coding drew on the interview topics and the 4 sensitizing domains. Inductive coding remained open to recurrent issues raised by participants. Codes and themes were reviewed iteratively in line with established qualitative analysis guidance [24]. The focused sample was considered information rich for exploratory analysis because all participants had direct exposure to the robot and its organizational setting [25]. A supplementary overview of the resulting themes is provided in Multimedia Appendix 1.

Mixed Methods Integration

After separate analysis, condition-level observation patterns were compared with staff perceptions of usefulness, interaction difficulty, human mediation, and organizational support. This comparison informed the provisional 4-domain framework presented in the Results section. The datasets involved different participant groups and were not linked at the person level; staff accounts were used to contextualize observed patterns, not to establish visitors' own perceptions or reasons for interacting.

Ethical Considerations

The study comprised anonymous, noninvasive observation and voluntary staff interviews. Ethics approval for observation was granted by the Tallinn Children's Hospital Research Ethics Committee (TCH/2025-01), and the interview component was reviewed by Tallinn University of Technology's Academic Ethics Research Committee. Because observation involved no recordings or identifiable personal data, individual consent was waived for the public space observation. Staff participants provided informed consent. Interview recordings and transcripts were stored securely, and the manuscript reports only aggregated observations and anonymized themes. All procedures complied with the General Data Protection Regulation (European Union 2016/679).

XSL•FO
RenderX

Results

Visitor Engagement Across the 3 Field Conditions

Across the 3 observation phases, 675 visitor passages were recorded, comprising 398 child and 277 adult passages. Each of the 4 separately recorded behavior indicators increased from baseline to poster support and again to staff facilitation (Table 1). The sums of the 4 indicator counts were 14, 52, and 156, respectively, corresponding to 5.4, 32.5, and 60.5 recorded indicators per 100 passages. These descriptive summaries include overlapping behaviors and are not percentages of passages with any interaction.

Table . Structured observation results by implementation condition, with a descriptive sum of overlapping indicator counts[a].

Condition	Looked, n (%)	Approached, n (%)	Touched, n (%)	Screen or app use, n (%)	Indicator sum[b], n (per 100 passages)
Baseline (n=257)[c]	8 (3.1)	2 (0.8)	2 (0.8)	2 (0.8)	14 (5.4)
Poster support (n=160)[d]	30 (18.8)	12 (7.5)	6 (3.8)	4 (2.5)	52 (32.5)
Staff facilitation (n=258)[e]	80 (31.0)	44 (17.1)	18 (7.0)	14 (5.4)	156 (60.5)

[a]Counts represent visitor passages, not confirmed unique individuals.

[b]The final column sums the 4 overlapping indicator counts within each condition. Values in parentheses in this column are counts per 100 passages. Because a single passage can contribute to multiple indicators, the sums do not represent distinct interaction episodes, passages with any interaction, or unique visitors.

[c]Includes 108 child and 149 adult passages.

[d]Includes 120 child and 40 adult passages.

[e]Includes 170 child and 88 adult passages.

In the baseline phase (n=257), the robot was largely ignored. Occasional automatic voice output did not translate into sustained use. Looking was recorded for 8 (3.1%) passages, while approaching, touching, and screen or app use were each recorded for 2 (0.8%) passages.

When the poster was added, visible interaction increased (n=160). Looking was recorded for 30 (18.8%) passages, approaching for 12 (7.5%) passages, touching for 6 (3.8%) passages, and screen or app use for 4 (2.5%) passages. The sign provided the robot's name, purpose, and interaction cues without changing its technical capabilities; whether visitors found these cues clearer or more useful was not assessed.

The highest counts occurred when a staff mediator stood near the robot and explicitly invited use (n=258). Looking was recorded for 80 (31.0%) passages, approaching for 44 (17.1%) passages, touching for 18 (7.0%) passages, and screen or app use for 14 (5.4%) passages. In descriptive terms, the phases represented progressively stronger levels of contextual and social support.

Contemporaneous field notes indicated that children were the most visibly responsive users in each phase, initiating contact through looking, approaching, touching, or play, while adults more often engaged in a supervisory or instrumental manner, for example, by watching a child interact or asking what the robot was for. Because age category was not linked to person-level behavior codes, these observations do not establish age-specific engagement rates and are not presented as a statistical comparison.

Staff Perspectives on Robot Use and Adoption

Staff Participant Background

The 4 focus groups included 12 staff members with direct exposure to the robot and represented both frontline clinical and organizational perspectives. The available role-level information is summarized in Multimedia Appendix 1. The absence of exact age, gender, tenure, and job title counts is considered when interpreting the transferability of the qualitative findings.

The interviews produced 5 interrelated themes. Multimedia Appendix 1 provides a compact overview; the main findings and illustrative quotations are presented below.

The Robot Was Accepted When Its Role Was Concrete and Visible

Staff were generally positive about the robot and did not describe strong resistance to its presence. Staff described a concrete and visible function as important for the robot's acceptance. The most frequently mentioned use case was wayfinding: helping families locate rooms, navigate the outpatient floor, or reduce confusion related to changing numbering systems and unfamiliar layouts. One participant explained:

> *If you have two children with you and you are looking for a four-digit room number, that is quite difficult for the parent…for the parent it is easier, for the child it is fun, so it is a win-win.* [P11]

Another participant described the robot as a practical alternative when staff could not leave their own work area:

> *When someone asks me where room A2071 is, I have no idea. Then I say, let's look at Lulu, and they*

XSL•FO
RenderX

disappear with the robot. I do not have to run with them myself. [P12]

Several participants contrasted this concrete usefulness with a more diffuse sense that the robot was present but not yet integrated. As one nurse put it:

For me it is actually still a bit unfamiliar. [P3]

Child Engagement Created Indirect Value for Parents and Staff

Participants consistently described children as drawn to the robot's movement, voice, and games. Even when parents were initially reserved, children often initiated contact, and parents then followed. One participant described the robot as "something very interesting" for children and added:

The child's attention goes there, not, for example, to the procedure. [P5]

Another described a similar impression:

For a moment they forget. [P4]

Staff framed the benefit as indirect as well as direct: engaging the child could make waiting easier for parents and reduce some low-level interaction burden for staff. These accounts describe perceived value; the study did not directly measure anxiety, distraction, or satisfaction.

Human Mediation and Local Prompting Were Critical

Staff accounts were consistent with the higher interaction counts observed in the staff facilitation condition. Staff said the robot was used more when someone invited families to try it, explained what it could do, or modeled its use. A coordination-level participant stated:

When there was a person next to it, inviting people to use it, then it was used more. [P2]

The same participant identified the practical constraint:

There was no resource to put anyone next to it to introduce it. [P2]

Frontline staff described a similar need from another angle. One nurse noted:

Medical staff do not guide Lulu use. [P3]

Another participant preferred a dedicated introduction:

There is a specific time, no one interrupts, and then you know exactly. [P5]

Staff perceived invitations and demonstrations as helpful for initiating interaction. Visitors' own reasons for engaging or not engaging were not assessed.

Limited Functionality Constrained Broader Acceptance

Although participants valued the idea of the robot, they also identified practical limitations. These included slow movement, charging and troubleshooting problems, unclear audibility, uneven content quality, and the need for multilingual support. One participant summarized the technical issues:

More speed. Sometimes it does not charge. [P4]

Another suggested that the robot display a support contact when problems occurred [P6]. Content was also a recurring issue. A senior participant argued:

I would remove the news. No one reads the news, children do not want to read them, and parents mostly watch the children play. [P12]

The participant proposed short, relevant "tips and tricks for parents" instead. A nurse noted that many visitors were Russian speakers while the device did not support Russian-language interaction [P3]. Positive attitudes therefore did not remove routine usability requirements.

Sustainable Use Required Organizational Ownership

A recurring issue was responsibility. Staff did not always know whom to contact when the robot malfunctioned, who could update its content, or who was expected to coordinate its place in daily work. One participant stated:

We need to choose Lulu's owner. [P12]

They added that if the hospital became the owner, responsibility should be assigned within the hospital. Another participant emphasized staff capacity:

The main problem here is time resource…people do not do it because they simply do not have the time. [P2]

Participants were open to broader use and viewed local responsibility for support, updates, and integration as important to its implementation.

Integrated Interpretation and Provisional Framework

Observed indicator counts were higher with poster support and highest with staff facilitation. Staff accounts were consistent with the potential importance of clear functions and human prompting but did not establish why visitors interacted or whether they experienced less uncertainty. The strands provide complementary evidence about observable behavior and staff perceptions of implementation, rather than direct confirmation of visitors' acceptance.

Field notes described children's visible responsiveness, while staff accounts identified perceived benefits and organizational needs. Together, these findings informed the provisional *mediated acceptance framework* in Table 2. The framework organizes possible links between visitor-facing support and staff-identified implementation conditions; it has not been validated against direct visitor feedback or sustained service outcomes.

XSL•FO
RenderX

Table . Provisional mediated acceptance framework linking observed visitor behavior with staff perceptions and proposed implementation priorities.

Domain	Evidence in this study	Proposed implementation priority
Role clarity	All 4 recorded indicators were higher with poster support than at baseline; staff emphasized wayfinding and room finding.	Define a bounded service task and make its purpose visible at the point of use
Interaction affordance	Screen or app use remained limited; staff-identified content, language, speed, audibility, charging, and recovery constraints.	Provide site-relevant multilingual content and clear start, use, and recovery cues
Social mediation	All 4 observed indicators were highest with staff facilitation; staff also described prompting and onboarding as necessary.	Consider early facilitation and brief staff onboarding and assess their usefulness directly with visitors and staff
Organizational embedding	Staff requested a named owner, support contact, content process, maintenance, troubleshooting, and protected time.	Assign ownership and integrate support, updates, maintenance, and evaluation into workflow

Discussion

Main Findings

This mixed methods pilot examined how a socially assistive robot was received in a pediatric outpatient setting and which conditions shaped practical acceptance. Across all 4 observed indicators (looking at the robot, approaching it, touching it, and using the robot's screen or app), interaction was lowest when the robot was present without explanation, increased when a poster clarified its role, and was highest when an intern invited and demonstrated use. Staff accounts provided a complementary implementation perspective rather than a direct explanation of visitors' experiences. Staff evaluated the robot as a service tool whose value depended on a concrete role, relevant and multilingual content, technical support, workflow fit, and visible organizational ownership.

Taken together, observed visitor behavior and staff perceptions informed the provisional mediated acceptance framework. The robot's hardware, location, and interface did not change across the 3 field phases; the surrounding support did. The results should nevertheless be read as descriptive associations. The sequential naturalistic design does not establish that the poster or facilitator alone caused the observed differences.

Interpretation in Light of Prior Research

The stepwise increase in recorded behavior can be interpreted in light of technology acceptance theory. Clearer information and human facilitation are possible mechanisms consistent with perceived usefulness, ease of use, social influence, and facilitating conditions [16,17]. These constructs were not measured among visitors, so this interpretation remains theoretical rather than an empirical explanation of their behavior.

The stronger visible responsiveness of children also aligns with pediatric and hospital robotics literature. Earlier reviews suggest that children often engage readily with social robots when interaction is concrete, emotionally legible, and easy to interpret [3,8]. Staff interviewees described movement, voice, games, and distraction during waiting as meaningful for children. Staff evaluated the robot in terms of usefulness, reliability, and support burden, while field notes described adult visitors more often taking supervisory or instrumental roles. These observations and staff accounts do not establish differences in perceptions between children and adult visitors, although they suggest questions relevant to research on the broader care ecosystem [7,9]. The present age-related interpretation remains cautious because it is based on field notes and lacks linked age-by-behavior counts.

The findings also support a sociotechnical interpretation of acceptance. Staff did not ask for a more spectacular or more human-like robot. They asked for clearer functions, stronger content, multilingual capacity, easier troubleshooting, training, and responsibility. These priorities mirror barriers and enablers identified in implementation reviews, where maintenance, workflow fit, reliability, and stewardship influence whether robots move beyond pilot visibility [10,11,19,20]. Staff descriptions of child engagement and usability constraints raise hypotheses relevant to enjoyment, social presence, trust, and ease of use in the Almere model [18]. These constructs were not assessed directly among visitors and should not be treated as observed acceptance outcomes.

The integrated framework adds a practical bridge between these traditions (Table 2). TAMs explain why a person may approach and try the robot, while implementation research explains whether the organization can sustain the service. In this study, observation documented behavior under different support conditions, while staff accounts identified perceived longer-term needs such as content management, troubleshooting, and ownership. The contribution is a provisional synthesis of these complementary perspectives, not evidence of visitors' perceived usefulness, trust, or acceptance.

Practical Implications

The descriptive observation patterns and staff accounts suggest that hospitals could evaluate clear signage, explicit interaction cues, and early human facilitation when introducing SARs. These approaches may support visitor interaction, but their effects on visitors' understanding, uncertainty, and perceived usefulness were not assessed in this study. Staff-identified wayfinding, waiting time engagement, parent guidance, and simple procedural information as potential use cases. These suggestions are consistent with roles discussed in the broader health care robot literature [4,5], but the present observations did not establish successful task completion or benefits perceived by visitors. Such functions should therefore be

XSL•FO
RenderX

evaluated with children, parents or guardians, and other adult visitors before broader implementation.

The interviews also suggest that a child-facing function may create indirect value for other people in the care environment. A robot that holds a child's attention during waiting may make navigation or preparation easier for parents and may reduce some routine interaction demands on staff. Future deployments should assess this potential chain of benefit directly; visible engagement alone is insufficient. Hospitals should therefore define expected outcomes for each use case, such as successful wayfinding, reduced uncertainty, or perceived helpfulness, and select measures that correspond to those outcomes. Staff accounts emphasized the need for organizational anchoring. A hospital should, with the help of relevant stakeholders, define what the robot does, where its task begins and ends, who updates content, who responds to faults, and how value will be assessed. Staff onboarding should include routine use and common recovery situations. The framework in Table 2 summarizes staff-informed priorities for implementation; whether these priorities improve visitor experience or sustained use requires further evaluation.

Limitations and Future Research

This was a single-site pilot, and the quantitative strand used descriptive observation without inferential testing. The 3 conditions were sequential and conducted on different dates, so temporal variation in visitor flow or other unmeasured contextual factors may have contributed to the differences. Counts represented anonymous passages, and repeat passages could not be identified. The 4 behavior indicators could overlap; their summed count per 100 passages does not measure the proportion of visitors who interacted, the depth or duration of engagement, or acceptance. Age category was not linked to individual behavior codes; therefore, age-related observations were limited to descriptive interpretation of the field notes.

The qualitative strand involved a small, staff-focused sample that differed from the observed visitor population. Children, parents or guardians, and other adult visitors were not interviewed. Although staff had direct exposure to the robot, their accounts reflect their own perceptions and cannot establish visitors' perceived usefulness, ease of use, reasons for interacting, or acceptance. Because the datasets were not linked at the person level, integration cannot establish that staff-identified factors explain the observed behavioral differences. Exact staff age, gender, tenure, and job title distributions were not collected to protect the identity of employees. The study also did not directly measure child emotion, anxiety, trust, or satisfaction, although affect is central to social HRI [8,26]. These limitations restrict transferability and indicate that the 4-domain framework remains provisional.

Future research should combine in situ observation with ethically appropriate direct feedback from children, parents, adult visitors, and staff. Repeated or counterbalanced implementation periods and privacy-preserving person-level behavior coding would permit stronger comparison of contextual and age-related patterns. Longer follow-up should examine whether multilingual content, stronger interaction design, staff onboarding, explicit ownership, and maintenance arrangements support sustained use after novelty effects diminish.

Comparative work should also examine different robot-mediated tasks and settings. Wayfinding may require clear movement, stopping, and destination cues, while waiting time engagement may require age-appropriate content, audibility, and careful limits on distraction. Multisite studies could distinguish context-specific findings from more general implementation requirements and test whether the 4-domain framework remains useful across pediatric hospitals with different layouts, languages, staffing models, and patient populations.

Conclusions

In this single-site descriptive pilot, observed visitor interaction with Temi V3 was higher under poster support and staff facilitation than at baseline. Staff perceived potential value in wayfinding, child engagement during waiting, and basic guidance, and they identified content, language coverage, reliability, support, and workflow fit as important implementation considerations.

These complementary findings informed a provisional mediated acceptance framework comprising role clarity, interaction affordance, social mediation, and organizational embedding. The framework is an interpretation of observed behavior and staff perspectives, not a demonstration of visitors' acceptance or of the effects of specific support mechanisms. Direct feedback from children, parents or guardians, and other adult visitors, together with longer-term evaluation, is needed to assess these proposed conditions and their contribution to sustained use.

Acknowledgments

The authors thank the staff of Tallinn Children's Hospital for their collaboration.

Generative AI Use

Generative AI tools (ChatGPT 5.6; OpenAI) were used to assist with language refinement and schematic figure preparation. The authors checked all scientific content, numerical values, references, and interpretations.

Funding

The authors received no financial support for the research, authorship, or publication of this article.

XSL•FO
RenderX

Data Availability

Anonymized behavioral data are stored on secure servers at Tallinn University of Technology and are available from the corresponding author upon reasonable request, subject to institutional data sharing agreements. Interview transcripts are not publicly available because the small institutional sample may permit reidentification; access would require additional ethical and institutional approval.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supporting methodological and qualitative summary information.
[DOCX File, 30 KB - nursing_v9i1e95702_app1.docx]

References

1. Feil-Seifer D, Mataric MJ. Defining socially assistive robotics. In: 9th International Conference on Rehabilitation Robotics (ICORR 2005): IEEE; 2005:465-468. [doi: 10.1109/ICORR.2005.1501143]
2. Henschel A, Laban G, Cross ES. What makes a robot social? A review of social robots from science fiction to a home or hospital near you. Curr Robot Rep 2021;2(1):9-19. [doi: 10.1007/s43154-020-00035-0] [Medline: 34977592]
3. González-González CS, Violant-Holz V, Gil-Iranzo RM. Social robots in hospitals: a systematic review. Appl Sci 2021 Jun 27;11(13):5976. [doi: 10.3390/app11135976]
4. Holland J, Kingston L, McCarthy C, et al. Service robots in the healthcare sector. Robotics 2021 Mar 11;10(1):47. [doi: 10.3390/robotics10010047]
5. Ragno L, Borboni A, Vannetti F, Amici C, Cusano N. Application of social robots in healthcare: review on characteristics, requirements, technical solutions. Sensors (Basel) 2023 Jul 31;23(15):6820. [doi: 10.3390/s23156820] [Medline: 37571603]
6. Silvera-Tawil D. Robotics in healthcare: a survey. SN Comput Sci 2024 Jan 11;5(1):189. [doi: 10.1007/s42979-023-02551-0]
7. Zhu T, Ahn HS, Broadbent E, MacDonald BA, Gasteiger N. A deep dive into human-robot interaction in hospitals: scoping review on the services provided, engagement behaviours and interaction outcomes. Int J Soc Robot 2025 Aug;17(8):1541-1561. [doi: 10.1007/s12369-025-01299-2]
8. Moerman CJ, van der Heide L, Heerink M. Social robots to support children's well-being under medical treatment: a systematic state-of-the-art review. J Child Health Care 2019 Dec;23(4):596-612. [doi: 10.1177/1367493518803031] [Medline: 30394806]
9. Kabacińska K, Teng KA, Robillard JM. Social robot interactions in a pediatric hospital setting: perspectives of children, parents, and healthcare providers. Multimodal Technol Interact 2025 Feb 11;9(2):14. [doi: 10.3390/mti9020014]
10. Papadopoulos I, Koulouglioti C, Lazzarino R, Ali S. Enablers and barriers to the implementation of socially assistive humanoid robots in health and social care: a systematic review. BMJ Open 2020 Jan 9;10(1):e033096. [doi: 10.1136/bmjopen-2019-033096] [Medline: 31924639]
11. Bradwell HL, Aguiar Noury GE, Edwards KJ, Winnington R, Thill S, Jones RB. Design recommendations for socially assistive robots for health and social care based on a large scale analysis of stakeholder positions. Health Policy Technol 2021 Sep;10(3):100544. [doi: 10.1016/j.hlpt.2021.100544]
12. Klebbe R, Klüber K, Dahms R, Onnasch L. Caregivers' perspectives on human–robot collaboration in inpatient elderly care settings. Machines 2023;11(1):34. [doi: 10.3390/machines11010034]
13. Aymerich-Franch L, Gómez E. Public perception of socially assistive robots for healthcare in the EU: a large-scale survey. Comput Hum Behav Rep 2024 Aug;15. [doi: 10.1016/j.chbr.2024.100465] [Medline: 39263424]
14. Introducing temi robot V3 - in education. temi USA. URL: https://www.robotemi.com/product/temi-sales-contact-1/ [accessed 2026-10-29]
15. Leoste J, Strömberg-Järvis K, Robal T, Marmor K, Kangur K, Rebane AM. Testing scenarios for using telepresence robots in healthcare settings. Comput Struct Biotechnol J 2024 Jan 17;24:105-114. [doi: 10.1016/j.csbj.2024.01.004] [Medline: 38314026]
16. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. Manag Inf Syst Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
17. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
18. Heerink M, Kröse B, Evers V, Wielinga BJ. Assessing acceptance of assistive social agent technology by older adults: the Almere model. Int J Soc Robot 2010 Dec;2(4):361-375. [doi: 10.1007/s12369-010-0068-5]
19. Servaty R, Kersten A, Brukamp K, Möhler R, Mueller M. Implementation of robotic devices in nursing care. Barriers and facilitators: an integrative review. BMJ Open 2020 Sep 21;10(9):e038650. [doi: 10.1136/bmjopen-2020-038650] [Medline: 32958491]

XSL•FO
RenderX

20. Ohneberg C, Stöbich N, Eberl I. Possible applications and conditions for the development and integration of robotic assistance systems in the acute inpatient setting from the perspective of nursing professionals: a qualitative study. Int J Soc Robot 2025 Sep;17(9):1821-1835. [doi: 10.1007/s12369-025-01289-4]
21. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research, 3rd edition: Sage; 2017.
22. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
23. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
24. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2021 Jul 3;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
25. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
26. Stock-Homburg R. Survey of emotions in human–robot interactions: perspectives from robotic psychology on 20 years of research. Int J Soc Robot 2022 Mar;14(2):389-411. [doi: 10.1007/s12369-021-00778-6]

Abbreviations

HRI: human-robot interaction
RQ: research question
SAR: socially assistive robot
TAM: technology acceptance model
UTAUT: unified theory of acceptance and use of technology

Edited by E Borycki; submitted 19.Mar.2026; peer-reviewed by CY Yang, XC Ye; revised version received 27.Sep.2026; accepted 21.Oct.2026; published 06.Nov.2026.

Please cite as:
Leoste J, Rakić S, Marmor K, Krinal K, Soostar LH, Vink KM, Käver A, Aasma K, Beatty K, Leoste LB
Evaluating Human-Robot Interaction With a Socially Assistive Robot in Pediatric Health Care: Mixed Methods Pilot Study
JMIR Nursing 2026;9:e95702
URL: https://nursing.jmir.org/2026/1/e95702
doi:10.2196/95702

© Janika Leoste, Slavko Rakić, Kristel Marmor, Kadi Lubi, Liina-Harjus Soostar, Katariina Maribel Vink, Andres Käver, Katrin Aasma, Katie Beatty, Leo Benjamin Leoste. Originally published in JMIR Nursing (https://nursing.jmir.org), 6.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Social Media Use and Digital Professionalism in Nursing Practice: Cross-Sectional Questionnaire Study

Melita Peršolja[1], PhD, RN; Anton Grmšek Svetlin[2], RN, MSN

[1]Faculty of Health Sciences, University of Primorska, Polje 42, Izola, Izola-Isola, Slovenia

[2]Izola General Hospital, Izola, Izola-Isola, Slovenia

Corresponding Author:
Melita Peršolja, PhD, RN
Faculty of Health Sciences, University of Primorska, Polje 42, Izola, Izola-Isola, Slovenia

Abstract

Background: The rapid growth of social media has transformed communication, with more than 5.66 billion users worldwide in 2025 who spent an average of 141 minutes daily on platforms such as Facebook, Instagram, WhatsApp, YouTube, and TikTok. Nurses frequently use social media to exchange work-related information, communicate with others, and relieve stress. While social media can support professional development, excessive or inappropriate use at work may compromise patient care. Despite widespread use, research on nurses' demographic factors and actual workplace social media behaviors remains limited.

Objective: The aim of the study was to determine how nurses use social media at work and how their demographic characteristics (eg, gender, age, and number of years working in health care) are related to the use of social media at work.

Methods: This study used a quantitative, descriptive, exploratory, and nonexperimental empirical research design. Data were collected via a survey administered to a convenience sample of 132 nurses. The sample comprised of 98 women (74.2%) and 34 men (25.8%). The questionnaire consisted of 17 statements related to social media use in the workplace. A statistical significance level of $P<.05$ was applied. Descriptive statistics and inferential statistical tests were used.

Results: Nurses used Facebook the most and spent between 45 minutes and 1 hour per day on social media. Study participants reported that they primarily used social media for communication with colleagues and acquiring new knowledge. They also used social media for its technical functions, such as file sharing. Nevertheless, a substantial number of nurses reported using social media during working hours, and many had experienced creating an uncomfortable situation after declining social media interactions with patients or colleagues.

Conclusions: This study highlights the important topic of social media use at work by nurses. The study found that nurses frequently used social media, primarily for professional communication and knowledge acquisition. While most demonstrated positive use patterns, the findings indicate that social media use during work may pose potential challenges to professionalism. These results highlight the need for clear guidelines and training to ensure that nurses use social media responsibly, maintaining both professional boundaries and personal integrity. Encouraging mindful and regulated use can enhance the benefits of social media while minimizing risks in the health care setting.

(JMIR Nursing 2026;9:e92923) doi:10.2196/92923

KEYWORDS

nursing; hospitals; digitalization; social media; informatics

Introduction

We are witnessing rapid technological development in the 21st century, particularly in relation to the internet and social media. Social media is defined as a group of online applications that enable the creation and sharing of content generated by users [1]. More than a decade ago, Kaplan and Haenlein [2] defined five main categories of social media: (1) collaborative projects (eg, Wikipedia), (2) blogs or microblogs (eg, X), (3) content communities (eg, YouTube), (4) social networking sites (eg, Instagram and Facebook), and (5) virtual gaming worlds (eg, Second Life).

Unlike traditional media, where users are usually passive recipients of information, social media platforms enable individuals to actively participate in creating and sharing content with others, which is precisely why they are becoming an increasingly important means of communication and connection among people [3]. In recent years, the use of social media has thus become an indispensable part of everyday life for many around the world. Many people spend several hours a day on platforms such as Facebook, Messenger, and Instagram, whether to maintain contacts, exchange information, or for entertainment and relaxation [4]. The data show that the number of social media users worldwide had reached 5.66 billion in October 2025, or approximately 70.8% of the global population [5].

XSL•FO
RenderX

Statistics indicate that in 2025, users spent an average of 141 minutes per day on social media. Compared with data from 2015, this number is on the rise, as daily social media use at that time was just 111 minutes [6]. The data further reveal that the most frequently used social media platforms are Facebook (Meta Platforms), Instagram (Meta Platforms), WhatsApp (Meta Platforms), YouTube (Alphabet Inc), and TikTok (ByteDance) [7].

According to research by Lefebvre et al [8], 87% of the 397 nurse respondents had an active user account on at least one social media platform. It was also found that older age is associated with less frequent use of social media. A study conducted by Wang et al [9] using a sample of 658 nurses revealed that approximately 56% of participants spend between 1 and 3 hours per day using social media. The majority (67.2%) reported that they use social media to exchange work-related information with colleagues. At the same time, 50.3% stated that they had witnessed inappropriate posts by their colleagues. In their study of a sample of 43 nurses at the tertiary health care level, Al-Rumhi et al [10] found that all the participants used social media, most commonly WhatsApp. They also found that most nurses spend 1 to 2 hours a day on social media. The authors state that the main motives for nurses to use social media were to acquire knowledge, communicate, and establish contact with patients, thus reducing the number of hospital visits.

To maintain professional integrity and ensure the responsible use of social media, the American Nurses Association has issued the following recommendations for its use [11]: (1) maintain professionalism online, (2) avoid sharing patient posts, (3) exercise caution when posting photos from the workplace, (4) avoid establishing contact with patients and their families, (5) refrain from disparaging one's employing organization, and (6) avoid posts that cast the nursing profession in a negative light. The growing use of social media in health care has highlighted the importance of digital professionalism, defined as the expression of professional values and behaviors in digital environments. As nurses increasingly engage with social media, understanding how these platforms are used in workplace contexts has become an important aspect of maintaining professional boundaries and ethical practice [12-14].

Through work, people satisfy numerous needs—biological, social, economic, and psychological—including the need for success, belonging, recognition, respect, and self-fulfillment. At the same time, work involves a system of social interactions that shape people, connect them with others, and significantly influence their quality of life [15,16]. Nurses who are exposed to high workloads are much more prone to burnout, absenteeism, and errors in patient care [17]. This is why mobile phones and social media are digital tools that health care workers often use in their work to reduce stress and divert their attention from work-related burdens [8,18]. However, excessive use of social media during working hours can compromise the quality of patient care [19] because it can reduce the time and energy devoted to patient-focused tasks, thereby reducing their work performance [20].

Since nurses are frequent users of social media [10,18], it is crucial that they recognize the associated dangers, especially when it comes to using it at work. Although evidence from numerous countries shows that nurses frequently use social media, there is still limited research examining how demographic factors relate to its use within the nursing profession, and there is a lack of data describing how such behaviors occur and what factors influence them. Moreover, most of the existing research focuses on general attitudes and perceptions, whereas few studies offer empirical quantitative data on actual patterns of social media use by nurses at work. Additionally, research from many regions focuses primarily on nursing students rather than practicing nurses, leaving a knowledge gap about the experiences of nurses already engaged in direct patient care.

The purpose of this research was to describe social media use among nurses. The specific aim of the study was to examine how nurses use social media in the workplace and to explore how their demographic characteristics (eg, gender, age, and years of experience in health care) are associated with social media use at work. We thus set the following research questions: Which social media platforms do nurses use most frequently? How much time do nurses spend on social media each day? For what purposes do nurses use social media at work? How are the demographic characteristics of nurses related to their use of social media at work?

Methods

A descriptive, exploratory, and nonexperimental method of empirical research was used with a survey technique.

Research Instrument Description

The research instrument was developed by Landers and Callan [21]. To ensure conceptual and linguistic accuracy in the Slovenian context, the original English instrument was translated independently by 4 translators. Each translated version was subsequently back-translated into English and compared with the original instrument to identify potential discrepancies and ensure semantic consistency. The translated versions were reviewed and discussed by the research team, and any differences were resolved through consensus.

The questionnaire assessed work-related social media use broadly, including behaviors and interactions connected to the workplace context. It comprised 17 items rated on a 5-point Likert scale ranging from 1 (disagree) to 5 (agree), organized into 2 conceptually distinct subscales. The first subscale assessed the positive, purposeful dimensions of work-related social media use and contained 8 items (eg, "I use social media to learn how to do my job better" and "I use social media to contact colleagues when I cannot reach them in any other way"). The second subscale consisted of 9 items measuring negative or problematic social media behaviors in the workplace (eg, "My coworkers have been offended by something I posted on social media" and "I discuss negative feelings about patients or colleagues on social media").

Before computing the composite work-related social media use score, all 9 items from the negative-behavior subscale were reverse-scored to ensure directional consistency across all 17 items. The composite score was then calculated as the mean of

XSL•FO
RenderX

all items, with higher scores reflecting greater overall work-related social media use. Although the questionnaire originally consisted of 2 conceptually distinct subscales (beneficial work-related use and potentially harmful workplace behaviors), the overall composite score was calculated as the mean of all items to capture a broader continuum of overall work-related social media behavior. This approach is consistent with the conceptual framework proposed by Landers and Callan [21], where work-related social media use may range from professional and beneficial engagement to problematic or inappropriate workplace behaviors. Therefore, higher composite scores reflected a more favorable overall pattern of work-related social media use.

Internal consistency was satisfactory for both subscales, with a Cronbach α of 0.73 for the positive-purpose subscale and 0.74 for the negative-behavior subscale [22].

The third section of the questionnaire collected demographic data, including gender, age, years of nursing work experience, level of health care, and field of employment. Participants were additionally asked whether they held a management position, which social media platforms they used most frequently, and their average daily time spent on social media.

Prior to the main study, the Slovenian version of the questionnaire was pilot-tested among 20 nurses to evaluate clarity, comprehensibility, and cultural appropriateness. Minor linguistic adjustments were made based on participant feedback before the final version was distributed to the study sample. The responses obtained from the pilot testing were not included in the final analysis of this study.

Research Sample Description

The target population of this study consisted of nurses employed in clinical practice across different health care settings in Slovenia. Eligible participants were individuals currently employed in nursing or direct health care practice who were actively involved in patient care or clinical work environments. To be included in the study, participants had to be currently employed within the health care system and willing to voluntarily complete the online questionnaire. Individuals not currently employed in clinical health care practice, health care students, retired health care workers, and respondents who did not complete the questionnaire in full were excluded from the final analysis.

A convenience sample was used in this study, and recruitment was conducted through snowball sampling, with the survey completed and distributed via online communication channels and social media. The snowball sampling approach was chosen because the study aimed to recruit nurses working across different health care settings in Slovenia, while online recruitment allowed efficient access to geographically dispersed participants. The survey link was initially distributed through professional and personal online communication channels, including social media platforms commonly used by health care professionals. Participants who accessed the questionnaire were subsequently encouraged to share the survey link with other nurses within their professional networks, thereby enabling broader recruitment through peer-to-peer distribution. To reduce potential sampling bias, the survey was distributed through multiple communication pathways (eg, social media platforms, email communication, professional networks, and direct person-to-person sharing) rather than relying on a single online platform, allowing access to nurses from different health care institutions and professional backgrounds.

The introductory page of the online questionnaire was viewed by 303 people. A total of 145 respondents began completing the questionnaire, and we included the responses of 132 respondents who completed the entire online questionnaire in our analysis. The research included 98 (74.2%) women and 34 (25.8%) men. The respondents were on average aged 34.3 (SD 10.33) years. The youngest respondent was aged 19 years and the oldest was 60 years. A total of 17 (12.9%) respondents were employed in primary health care, 88 (66.7%) in secondary health care, and 27 (20.5%) in tertiary health care. Most respondents were employed in hospital wards, for example, surgical (n=33, 25%), internal medicine and pediatrics (n=14, 10.6%), and intensive care (n=13, 9.8%). Respondents who did not select one of the offered work areas most often cited emergency medical services, radiology, and oncology as their areas of expertise.

Twenty-eight (21.2%) respondents had completed secondary education, 86 (65.2%) had completed higher education, 17 (12.9%) had completed postgraduate master's education, and 1 (0.8%) respondent had obtained a doctorate. Most nurses (n=92, 69.7%) who participated in our study are employed as registered nurses. Thirty-one (23.5%) respondents stated that they performed a managerial function, while 101 (76.5%) respondents stated that they did not.

The largest group of nurses had been employed in health care for 1 to 4 years (n=26), followed by those employed for 20 years or more (n=25).

Description of the Research Process and Data Processing

The survey was conducted using the open-source online survey application 1KA (version 26.09.01; Faculty of Social Sciences, Ljubljana, Slovenia). The questionnaire was distributed via online media and other communication channels using a snowball sampling method. Respondents had the opportunity to complete the questionnaire from July 9, 2025, to November 9, 2025.

Only fully completed questionnaires were included in the data analysis. The data were analyzed using SPSS (version 26.0; IBM Corp). Basic descriptive statistics were calculated (frequency, range, mean, and SD), and statistical tests were performed: the Kolmogorov-Smirnov test and 1-way ANOVA, independent-samples *t* test, 1-sample *t* test and the Tukey post hoc test. The level of statistical significance was set at $P<.05$. During the analysis, we created a composite variable, "Use of social media at work," which combined all 17 statements about the use of social media at work into a single variable.

Ethical Considerations

This study was conducted under Umbrella Ethical Approval for Research granted by the University of Primorska (approval

XSL•FO
RenderX

4264-19-6/23). Before completing the questionnaire, respondents gave their consent to participate in the survey. All procedures in this research were conducted in accordance with ethical guidelines outlined by the Declaration of Helsinki [23] and followed institutional protocols to ensure quality, integrity, and ethical responsibility.

Results

Overview

In our study, we found that Facebook was the most frequently used social media platform (n=113), followed by Instagram (n=105), Viber (Rakuten Viber; n=61), TikTok (n=55), and WhatsApp (n=34). Respondents who did not select one of the options provided most frequently cited LinkedIn (Microsoft), Pinterest (Pinterest, Inc), and BeReal (Voodoo).

Most respondents (45/132, 34.1%) spend between 45 minutes and 1 hour per day on social media, followed by those who use social media from 1 to 2 hours a day (37/132, 28%), 30 to 45 minutes a day (23/132, 17.4%), more than 2 hours a day (20/132, 15.2%), and less than 30 minutes (7/132, 5.3%).

The highest average value in the first set of questions (4.11, SD 1.09) was obtained for the statement that nurses use social media to contact colleagues when they cannot reach them in any other way. This indicates that one of the most important aspects of social media use is communication with colleagues. Nurses also rated relatively highly (mean 2.98, SD 1.24) the statement that they use social media to learn how to do their job better, which means that one of the purposes of using social media is education. In this section, they also expressed a high level of agreement (mean 2.95, SD 1.31) with the statement that they use the technical functions of social media, such as file sharing, to perform their work tasks, which reveals an additional aspect of their social media use. The lowest average value in this section of the questionnaire was for the statement that the respondents establish contact with potential new patients via social media (mean 1.40, SD 0.84). The average beta coefficient for the variability of the first set of statements was β=0.48. The full sociodemographic characteristics of the respondents are shown in Table 1.

The statements in the second set of questions mostly describe negative situations related to the use of social media at work, which means that lower average values for the statements indicate a lower level of agreement with the statements offered, suggesting less problematic use of social media. Nurses rated the statements in this section of the questionnaire low on average. Nevertheless, 2 statements received higher ratings, namely that nurses used social media when they should have been working (mean 3.08, SD 1.27) and that they had caused an unpleasant situation in the past by refusing to connect with colleagues, superiors, or patients via social networks (mean 2.75, SD 1.41). The lowest rating (mean 1.22, SD 0.59) was given to the statement that patients had posted information about respondents on social media that damaged their reputation at work, which means that most respondents had not yet encountered this situation. In this part of the questionnaire, the average beta coefficient measuring the variability of the statements was β=0.43.

XSL•FO
RenderX

Table . Sociodemographic characteristics of the participants (N=132).

Characteristics	Participants, n (%)
Gender	
Man	34 (25.8)
Woman	98 (74.2)
Health care setting	
Primary level (eg, primary health center and nursing home)	17 (12.9)
Secondary level (eg, general hospital)	88 (66.7)
Tertiary level (eg, clinical center)	27 (20.5)
Area of work	
Surgical ward	33 (25)
Internal medicine ward	14 (10.6)
Pediatric ward	14 (10.6)
Gynecological ward	9 (6.8)
Specialist outpatient clinic	8 (6.1)
Primary care clinic	3 (2.3)
Nursing home	3 (2.3)
Home care nursing	5 (3.8)
Intensive care unit	13 (9.8)
Emergency room	7 (5.3)
Other (eg, radiology ward)	23 (17.4)
Education level	
High-school diploma	28 (21.2)
Bachelor’s degree	86 (65.2)
Master’s degree	17 (12.9)
Doctoral degree	1 (0.8)
Professional role	
Practical nurse	2 (1.5)
Professional nurse	19 (14.4)
Registered nurse	92 (69.7)
Other (eg, managers)	19 (14.4)
Years of experience in nursing	
<1	20 (15.2)
1-4	26 (19.7)
5-9	23 (17.4)
10-14	19 (14.4)
15-19	19 (14.4)
≥20	25 (18.9)

The 2 statements that had the biggest impact (β=0.61) on the variability of the entire questionnaire were that respondents use the technical functions of social media to perform work tasks and that they ask for help via social media when they encounter problems at work (Table 2). The *P* values for all 17 statements were obtained with a 1-sample 2-tailed *t* test.

XSL•FO
RenderX

Table . Values of statements related to the use of social media at work.

Statement	Score, mean (SD; range)	β[a]	*P* value
I use social media to contact my coworkers when I am unable to reach them by other means.	4.11 (1.09; 1-5)	0.49	<.001
If I find something on social media that could harm the reputation of my coworkers or our organization, I let people know.	3.61 (1.21; 1-5)	0.21	.01
I have used social media when I should have been working.	3.08 (1.27; 1-5)	0.54	<.001
I have used social media to learn how to perform better in my job.	2.98 (1.24; 1-5)	0.47	<.001
I have taken advantage of the technical features of social media (such as file sharing or scheduling functions) to accomplish work tasks.	2.95 (1.31; 1-5)	0.61	<.001
I have created an uncomfortable situation by refusing connections with coworkers, supervisors, or patients via social media.	2.75 (1.41; 1-5)	0.55	<.001
I access social media while I am doing other work.	2.38 (1.14; 1-5)	0.53	<.001
I request help from people on social media when I am having trouble solving a problem at work.	2.33 (1.19; 1-5)	0.61	<.001
I use my organization's official social media presence to network.	2.32 (1.43; 1-5)	0.42	<.001
I communicate with patients via social media.	1.62 (1.04; 1-5)	0.50	<.001
Other people at work have been offended by something I posted on social media.	1.61 (1.07; 1-5)	0.40	<.001
Through social media, I have entered a personal relationship with a patient or a colleague that I should not have had.	1.53 (0.93; 1-4)	0.43	<.001
I have stolen information or other content from social media and used it as if it were my own work.	1.42 (0.81; 1-4)	0.31	<.001
I connect with potential new patients through social media.	1.40 (0.84; 1-4)	0.50	<.001
I have discussed negative feelings toward patients or coworkers on social media.	1.37 (0.77; 1-4)	0.34	<.001
I have done poor quality work using my organization's social media accounts.	1.25 (0.57; 1-4)	0.46	<.001
Patients posted information about me on social media that harmed my reputation at work.	1.22 (0.59; 1-4)	0.28	<.001

[a]β: beta coefficient

Inferential Statistics

Due to the normal distribution of the data, which we determined using the Kolmogorov-Smirnov test (Kolmogorov-Smirnov statistic=0.07; *P*=.20), we used parametric tests for further analysis.

Using the independent-samples *t* test, we did not find any statistically significant differences in the use of social media

XSL•FO
RenderX

between genders (t_{130}=0.177; P=.86; Cohen d=0.04). Additionally, we did not find statistically significant differences in the use of social media according to the managerial function of nurses (t_{130}=–0.294; P=.77; Cohen d=–0.06).

A 1-way ANOVA showed statistically significant differences ($F_{2,129}$=7.399; P=.001; η^2=0.103) in the use of social media at work among nurses at the primary (mean 40.47, SD 7.61), secondary (mean 38.75, SD 8.44), and tertiary (mean 32.18, SD 8.08) levels of health care. The Tukey post hoc test showed statistically significant differences (P=.001) between nurses at the secondary and tertiary levels of health care. Nurses employed at the tertiary level achieved lower (more favorable) values, which indicated a more professional and thoughtful use of social media compared with nurses at the secondary level of health care.

By contrast, ANOVA did not detect statistically significant differences in the use of social media based on professional role ($F_{4,127}$=1.268; P=.29; η^2=0.03) and years of nursing experience ($F_{5,126}$=0.883; P=.49; η^2=0.03).

Discussion

Principal Findings

The purpose of our research was to describe the use of social media at work among nurses. We found that the most frequently used platform among the respondents was Facebook, with 113 nurses using it. Other frequently used social media platforms were Instagram, Viber, TikTok, and WhatsApp. Gamor et al [24] reported similar results, finding that health care workers most frequently used Facebook, WhatsApp, Twitter (now X), YouTube, and Instagram. In line with our results, Van Ravenswaay et al [25] found that Facebook was the most frequently used social media platform among health care professionals.

In our study, we found that nurses are frequent users of social media, using it for an average of 45 minutes to 1 hour per day. Surani et al [26] came to a similar conclusion, finding that most nurses spent between 31 and 60 minutes per day on social media. Ahmead et al [27] found that 33.5% of nurses spent more than 1 hour per day on social media, while 45% of them used social media for 1 to 3 hours per day. They also found that 76.6% of nurses spent more than 1 hour per day on social media at work. In our study, we found that nurses often used social media when they should have been working. The use of social media in the workplace can be potentially harmful to nurses' work. Wilson [28] identified 5 key risks associated with the use of social media at work, namely (1) a decline in employee productivity, (2) information leaks, (3) damage to the institution's reputation, (4) cybercrime, and (5) data loss due to insecure passwords. Geraghty et al [29] concluded that the use of social media at work can be detrimental to nurses, as it reduces their level of attention and perception of the situation at hand, which are both crucial for the safe care of patients.

We found that nurses primarily used social media to communicate with colleagues when they could not reach them by other means. This was also reflected in the more frequent use of social media platforms whose primary purpose is communication (eg, Viber and WhatsApp). Gamor et al [24] also found that one of the primary purposes of social media use for nurses was to communicate with colleagues. In addition, we found that nurses often used social media to learn how to do their jobs better. The results were consistent with the findings of other studies, which concluded that nurses often used social media for professional development, knowledge exchange, and education [24,30,31].

In the second part of the questionnaire, the respondents agreed least with the statement that they use social media to establish contact with patients. It is very important for nurses to be professional in their work, to have clear boundaries, and to maintain a professional relationship with their patients to protect their integrity. At the beginning of the rise of social media, Aylott [32] noted that the line between personal and professional relationships between nurses and patients was quickly blurred, especially with the help of social media, which facilitates communication and accessibility. Valente [33] notes that nurses may receive friend requests from patients on social media, but that crossing professional boundaries in this way can significantly damage their professionalism. At the same time, such behavior undermines patients' trust in the work of nurses and damages the therapeutic relationship.

Because nurses often do not respond to friend or follow requests on social media or other calls to establish contact with patients or colleagues, they can quickly create an uncomfortable situation. Most of the participants in our study agreed that they had been in such a situation before. Unhjem [34] found that nurses often received friend requests from patients on Facebook, and that most nurses decided to accept them, even though this negatively affected their professionalism and integrity and represented a serious violation of the therapeutic relationship with the patient. The author concluded that most nurses accepted the friend requests because otherwise they felt uncomfortable and guilty about rejecting patients. Importantly, this discomfort is not merely interpersonal but is grounded in ethical and patient safety concerns. Professional standards emphasize that nurses must respect defined boundaries when interacting with patients, whether in clinical settings or on social media [35]. From an ethical and patient safety standpoint, several key risks are associated with social media interactions involving patients, such as the blurring of professional boundaries [34] and compromising confidentiality and privacy [36]. It is thus crucial for all nurses to maintain their professionalism outside the clinical settings where they work and not to enter into personal relationships with patients, thus preserving their professionalism.

In our study, we found that nurses employed at the tertiary level of health care used social media in the most professional manner. This could be attributed to clearer policies and guidelines in tertiary health care institutions, which encourage the professional use of social media in clinical practice [37]. This suggests that environments with more established governance provide a clearer framework for social media use that mitigates any potential risks (such as confidentiality concerns and unprofessional behavior). However, this study did not find any statistically significant differences in social media use based on age (P=.99), gender (P=.11), or job position (P=.29). In contrast,

XSL•FO
RenderX

Nelson-Brantley et al [31] found that the use of social media differed statistically significantly among nurses based on demographic data such as age, job position, and type of employment. Lefebvre et al [8] found that age affected the frequency of social media use and that the use declined with increasing age. This could be attributed to the fact that younger nurses generally have easier and more frequent access to mobile phones.

This study contributes to the growing body of literature on digital professionalism and social media use in health care by providing empirical evidence on how nurses use social media within clinical work environments. While previous research has often focused on nursing students or general attitudes toward social media, our study offers insight into actual workplace behaviors among practicing nurses. The findings have relevance beyond the Slovenian health care setting, as health care institutions worldwide increasingly face challenges related to digital professionalism, maintaining professional boundaries, and ensuring the responsible use of social media in clinical practice. Our results thus contribute to the broader international discussion on how digital technologies are reshaping professional behavior within modern health care systems.

Study Limitations and Suggestions for Future Research

While nurses in clinical practice participated in the study, the study relied on nonprobability sampling methods and included a small, uneven sample and spanned different health care settings, which limited the ability to generalize the findings to the broader global nursing workforce. Future research should therefore include a larger sample of nurses across more diverse regions of the world and systems to enable meaningful comparisons across different countries and contexts. Through multicountry comparative studies, researchers could examine how cultural norms affect nurses' social media use in the workplace.

Another limitation of this study relates to the use of snowball sampling and online recruitment through social media and digital communication channels. Because participation depended on voluntary engagement with online distribution channels, the study may have disproportionately reached nurses who are more familiar with or more active users of social media. This may have introduced self-selection bias and potentially overrepresented participants with higher levels of digital engagement. Consequently, the findings cannot be generalized to the broader nursing population.

Another limitation is the low number of respondents holding master's or doctoral degrees, which prevents the findings from being extended to highly educated nurses. In the future, it is imperative that more educated nurses participate in the study.

Another limitation of the study is that no a priori sample size calculation or power analysis was performed before data collection. Consequently, the adequacy of the sample size for detecting smaller effects cannot be fully determined and should be considered when interpreting the findings.

Furthermore, the cross-sectional design of the study limits the ability to draw causal conclusions regarding the relationships between demographic characteristics and social media use. The findings represent a snapshot of participants' behaviors and perceptions at a single point in time. Moreover, data were collected using a self-report questionnaire, which may have introduced response bias and social desirability bias. Participants may have underreported potentially inappropriate social media behaviors or overreported professionally desirable practices, which could have influenced the accuracy of the reported findings.

Additional qualitative research is recommended using in-depth individual interviews or multiple focus groups to gain a more comprehensive understanding of how nurses use social media in clinical practice, and experimental research is also recommended. In this way, researchers could test how specific types of social media influence nurses' decision-making and how they affect their work performance.

Conclusions

Our study demonstrates that nurses in clinical practice are frequent users of social media, primarily for professional communication with colleagues and for acquiring new knowledge. However, our findings also highlight potential challenges, including the use of social media during work hours, which may distract nurses from patient care. It is essential that nurses maintain professional boundaries and adhere to ethical guidelines when using social media in the workplace, both to safeguard patient safety and to uphold the integrity of the nursing profession. By promoting responsible and purposeful social media use, health care institutions can support nurses in leveraging digital tools to enhance professional development, collaboration, and the overall quality of patient care.

Funding

The internal research program SiZDRAV – Synergy between Health, Work and Education is cofunded by the University of Primorska within the framework of the Internal Call for Internal Program Groups at the University of Primorska and by the Public Agency for Scientific Research and Innovation of the Republic of Slovenia (ARIS) through stable funding under the heading Development Pillar of Funding.

Authors' Contributions

Conceptualization: MP (lead), AGS (equal)
Data curation: AGS
Formal analysis: AGS (lead), MP (supporting)

XSL•FO
RenderX

Funding acquisition: MP
Investigation: AGS
Methodology: MP
Project administration: AGS (lead), MP (equal)
Resources: MP
Supervision: MP
Validation: MP
Visualization: MP (lead), AGS (supporting)
Writing—original draft: AGS (lead), MP (supporting)
Writing—review and editing: MP (lead), AGS (supporting)

Conflicts of Interest

None declared.

References

1. Hamm MP, Chisholm A, Shulhan J, et al. Social media use by health care professionals and trainees: a scoping review. Acad Med 2013 Sep;88(9):1376-1383. [doi: 10.1097/ACM.0b013e31829eb91c] [Medline: 23887004]
2. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of social media. Bus Horiz 2010;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]
3. Nam SJ, Hwang H. Consumers' participation in information-related activities on social media. PLoS One 2021;16(4):e0250248. [doi: 10.1371/journal.pone.0250248] [Medline: 33891643]
4. Bartosik-Purgat M, Filimon N, Kiygi-Calli M. Social media and higher education – an international perspective. Econ Sociol 2017;10(1):181-191. [doi: 10.14254/2071-789X.2017/10-1/13]
5. Most popular social networks worldwide as of October 2025, by number of monthly active users (in millions). Statista. 2025. URL: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ [accessed 2025-11-11]
6. Daily time spent on social networking by internet users worldwide from 2012 to 2025. Statista. 2025. URL: https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/ [accessed 2025-11-11]
7. Number of internet and social media users worldwide as of April 2026. Statista. 2026. URL: https://www.statista.com/statistics/617136/digital-population-worldwide/ [accessed 2025-11-11]
8. Lefebvre C, McKinney K, Glass C, et al. Social media usage among nurses: perceptions and practices. J Nurs Adm 2020 Mar;50(3):135-141. [doi: 10.1097/NNA.0000000000000857] [Medline: 32049701]
9. Wang Z, Wang S, Zhang Y, Jiang X. Social media usage and online professionalism among registered nurses: a cross-sectional survey. Int J Nurs Stud 2019 Oct;98:19-26. [doi: 10.1016/j.ijnurstu.2019.06.001] [Medline: 31255853]
10. Al-Rumhi A, Al-Rasbi S, Momani AM. The use of social media by clinical nurse specialists at a tertiary hospital: mixed methods study. JMIR Nurs 2023 Aug 24;6(1):e45150. [doi: 10.2196/45150] [Medline: 37616026]
11. Social media do's and don'ts for nurses. American Nurses Association. 2024 Mar 14. URL: https://www.nursingworld.org/content-hub/resources/workplace/social-media-dos-and-donts-for-nurses/ [accessed 2025-11-11]
12. Gum L, Brown A, Royals R, Matricciani L, Kelly MA. Digital professionalism in preregistration nursing and midwifery students: a scoping review to explore perceptions of professionalism when using social media. Nurse Educ Pract 2024 Oct;80:104128. [doi: 10.1016/j.nepr.2024.104128] [Medline: 39293167]
13. Guraya SS, Guraya SY, Rashid-Doubell F, et al. Reclaiming the concept of professionalism in the digital context: a principle-based concept analysis. Ann Med 2024 Dec;56(1):2398202. [doi: 10.1080/07853890.2024.2398202] [Medline: 39263743]
14. O'Connor S, Zhang M, Honey M, Lee JJ. Digital professionalism on social media: a narrative review of the medical, nursing, and allied health education literature. Int J Med Inform 2021 Sep;153:104514. [doi: 10.1016/j.ijmedinf.2021.104514] [Medline: 34139621]
15. Gorbaniuk J, Chuchra M. The meaning of professional work in the life of nurses − comparative study [Article in Polish]. Rocz Teol 2019;66(1):105-122. [doi: 10.18290/rt.2019.66.1-7]
16. Lisowska E. Professional determinants of job satisfaction among teachers [Article in Polish]. Pedagogical Forum 2017;7(1):227-244. [doi: 10.21697/fp.2017.1.16]
17. Farid M, Purdy N, Neumann WP. Using system dynamics modelling to show the effect of nurse workload on nurses' health and quality of care. Ergonomics 2020 Aug;63(8):952-964. [doi: 10.1080/00140139.2019.1690674] [Medline: 31696791]
18. Boukari ZI, Elseesy NA, Felemban O, Alharazi R. Between clicks and care: investigating social media addiction and work engagement among nurses in Saudi Arabia. Nurs Rep 2025 Feb 28;15(3):84. [doi: 10.3390/nursrep15030084] [Medline: 40137656]
19. Javed A, Yasir M, Majid A, et al. Evaluating the effects of social networking sites addiction, task distraction, and self-management on nurses' performance. J Adv Nurs 2019 Nov;75(11):2820-2833. [doi: 10.1111/jan.14167] [Medline: 31385324]

XSL•FO
RenderX

20. Jeyaraman M, Ramasubramanian S, Kumar S, et al. Multifaceted role of social media in healthcare: opportunities, challenges, and the need for quality control. Cureus 2023;15(5):e39111. [doi: 10.7759/cureus.39111] [Medline: 37332420]
21. Landers RN, Callan RC. Validation of the beneficial and harmful work-related social media behavioral taxonomies: development of the Work-related Social Media Questionnaire. Soc Sci Comput Rev 2014;32(5):628-646. [doi: 10.1177/0894439314524891]
22. Taber KS. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ 2018 Dec;48:1273-1296. [doi: 10.1007/s11165-016-9602-2]
23. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
24. Gamor N, Dzansi G, Konlan KD, Abdulai E. Exploring social media adoption by nurses for nursing practice in rural Volta, Ghana. Nurs Open 2023 Jul;10(7):4432-4441. [doi: 10.1002/nop2.1685] [Medline: 36840611]
25. Van Ravenswaay L, Parnes A, Nisly SA. Clicks for credit: an analysis of healthcare professionals' social media use and potential for continuing professional development activities. Med Educ Online 2024 Dec 31;29(1):2316489. [doi: 10.1080/10872981.2024.2316489] [Medline: 38359156]
26. Surani Z, Hirani R, Elias A, et al. Social media usage among health care providers. BMC Res Notes 2017 Nov 29;10(1):654. [doi: 10.1186/s13104-017-2993-y] [Medline: 29187244]
27. Ahmead M, Hamamadeh N, Iram IA. The effects of internet and social media use on the work performance of physicians and nurses at workplaces in Palestine. BMC Health Serv Res 2022 May 12;22(1):633. [doi: 10.1186/s12913-022-07934-2] [Medline: 35550072]
28. Wilson J. Social networking: the business case. Eng Technol 2009 Jun;4(10):54-56. [doi: 10.1049/et.2009.1010]
29. Geraghty S, Hari R, Oliver K. Using social media in contemporary nursing: risks and benefits. Br J Nurs 2021 Oct 14;30(18):1078-1082. [doi: 10.12968/bjon.2021.30.18.1078] [Medline: 34645353]
30. Kazemi SS, Tavafian SS, Hidarnia A, Montazeri A. Exploring nurses' experiences of social media and in-person educational interventions for professional development: a qualitative study. BMC Nurs 2022 May 24;21(1):126. [doi: 10.1186/s12912-022-00903-4] [Medline: 35610638]
31. Nelson-Brantley HV, Parchment J, Chipps E, et al. The relationship between social media use, nurse characteristics, and job decision-making. J Nurs Adm 2024;54(10):544-553. [doi: 10.1097/NNA.0000000000001487] [Medline: 39265034]
32. Aylott M. Blurring the boundaries: technology and the nurse-patient relationship. Br J Nurs 2011;20(13):810-812. [doi: 10.12968/bjon.2011.20.13.810] [Medline: 21841689]
33. Valente SM. Managing professional and nurse-patient relationship boundaries in mental health. J Psychosoc Nurs Ment Health Serv 2017 Jan 1;55(1):45-51. [doi: 10.3928/02793695-20170119-09] [Medline: 28135391]
34. Unhjem JV. Accept or decline? Nurses and patients as Facebook friends. Sykepleien 2015. [doi: 10.4220/Sykepleienf.2020.83092en]
35. McCartney PR. Social networking principles for nurses. MCN Am J Matern Child Nurs 2012;37(2):131. [doi: 10.1097/NMC.0b013e3182430380] [Medline: 22357076]
36. Guraya SS, Guraya SY, Yusoff MS. Preserving professional identities, behaviors, and values in digital professionalism using social networking sites; a systematic review. BMC Med Educ 2021 Jul 12;21(1):381. [doi: 10.1186/s12909-021-02802-9] [Medline: 34247617]
37. Mariano MC, Maniego JC, Manila H, et al. Social media use profile, social skills, and nurse-patient interaction among registered nurses in tertiary hospitals: a structural equation model analysis. Int J Nurs Stud 2018 Apr;80:76-82. [doi: 10.1016/j.ijnurstu.2017.12.014] [Medline: 29353713]

Edited by E Borycki; submitted 05.Feb.2026; peer-reviewed by D Gandasari, O Cohen; revised version received 31.Jul.2026; accepted 07.Sep.2026; published 21.Oct.2026.

Please cite as:
Peršolja M, Grmšek Svetlin A
Social Media Use and Digital Professionalism in Nursing Practice: Cross-Sectional Questionnaire Study
JMIR Nursing 2026;9:e92923
URL: https://nursing.jmir.org/2026/1/e92923
doi:10.2196/92923

© Melita Peršolja, Anton Grmšek Svetlin. Originally published in JMIR Nursing (https://nursing.jmir.org), 21.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Radar-Based Contactless Monitoring in Different Inpatient Nursing Care Settings: Mixed Methods Feasibility Study

Madeleine Ritter-Herschbach[1*], Dr rer med; Sebastian Hofstetter[1,2*], PhD; Dominik Behr[1], BSc; Melina Hager[3], BSc; Burkhard Mueller[4], RN; Patrick Jahn[1], Prof Dr

[1]Health Service Research Working Group, Faculty of Medicine, Martin Luther University Halle-Wittenberg, Ernst-Grube-Str. 40, Halle, Saxony-Anhalt, Germany

[2]Dorothea-Erxleben-Lernzentrum, Faculty of Medicine, Martin Luther University Halle-Wittenberg, Halle, Saxony-Anhalt, Germany

[3]University Clinic and Polyclinic for Geriatric Medicine, University Hospital in Halle, Halle, Saxony-Anhalt, Germany

[4]University Clinic and Polyclinic for Internal Medicine II, University Hospital in Halle, Halle, Saxony-Anhalt, Germany

[*]these authors contributed equally

Corresponding Author:
Madeleine Ritter-Herschbach, Dr rer med
Health Service Research Working Group, Faculty of Medicine, Martin Luther University Halle-Wittenberg, Ernst-Grube-Str. 40, Halle, Saxony-Anhalt, Germany

Abstract

Background: Monitoring of vital signs is a critical component of nursing care, facilitating the early detection of complications, enabling the observation of the progression of health status, and providing a foundation for diagnosis and therapeutic decision-making. It is instrumental in ensuring quality of care. Contactless measurement methods, such as radar-based sensors, have the potential to enable continuous monitoring of respiration, heart rate, and bed activity without disturbing or stressing vulnerable patients.

Objective: The study investigates the feasibility, usability, acceptance, and implementation of a contactless digital vital-signs monitoring system. The objective is to ascertain the suitability of the device for incorporation into the workflow of nursing personnel and to assess implementation factors in an inpatient setting. Additionally, the study examines user acceptance and satisfaction among nursing staff.

Methods: A total of 13 nursing staff members from geriatric and dialysis departments participated in this mixed methods pilot study. Following the installation of the technology, the participants underwent a 30-minute training period to become acquainted with its operation. Thereafter, the system was subjected to a 4-week trial period. Quantitative data were collected at 3 points in time using the Technology Usage Inventory and the System Usability Scale. Furthermore, measurement protocols were evaluated to compare the time required for the standard and contactless measurement methods. Qualitative data were collected subsequent to the conclusion of the trial through the implementation of focus group interviews. Ethical approval and informed consent were obtained.

Results: The study results suggest that the average intention to use among the study participants (n=13) reached a relatively high level after 4 weeks (mean 85.5, SD 107.7). However, variation was observed when the 2 nursing wards were considered separately (geriatrics: 28.4 vs dialysis 199.8). The monitoring system was evaluated based on the System Usability Scale, which yielded a mean score of 74.5 (SD 10.5), indicating good usability. The nurses reported a perceived enhancement in their sense of security and monitoring oversight during routine care. Additionally, potential time savings were observed due to a reduction in disruptions to established processes. However, compared with conventional real-time monitoring, disadvantages were identified and technical requirements were defined.

Conclusions: The use of contactless measurement in the assessment of vital signs holds considerable promise for enhancing the efficacy of nursing care, particularly in contexts where manual measurements have been the prevailing standard. In order to further explore technical capabilities in a nursing context, additional research needs to be conducted.

Trial Registration: German Clinical Trials Register DRKS00036910; https://drks.de/search/en/trial/DRKS00036910/details

(JMIR Nursing 2026;9:e93624) doi:10.2196/93624

KEYWORDS

vital signs monitoring; radar-based sensors; patient safety; inpatient care; nursing care process; usability

XSL•FO
RenderX

Introduction

Contactless Monitoring Systems in Hospital Care

The continuous monitoring of vital signs constitutes a fundamental element of professional nursing care and clinical decision-making. The regular assessment of parameters such as heart rate, respiratory rate, and physical activity enables early identification of clinical deterioration, supports timely interventions, and contributes to patient safety [1]. Beyond its clinical relevance, systematic monitoring and documentation of vital signs also play a crucial role in quality assurance processes and provide legal protection for nursing staff by ensuring transparency and traceability of care interventions.

In many inpatient settings, vital signs are still manually measured at predetermined intervals [2,3]. Conventional monitoring methods often require direct physical contact with patients, which can impose a significant burden or elicit feelings of stress. Furthermore, this contact has been identified as a risk factor for the development of delirium [4], particularly among vulnerable populations, including older patients, individuals with cognitive impairment [5], patients who had a stroke [6], and those with fragile skin conditions [7]. Repeated manual measurements can disrupt patients' rest and sleep. This disruption may, in turn, have a deleterious effect on the recovery and overall well-being of the individual.

Advancements in digital health technologies have precipitated the development of contactless monitoring systems, including radar-based sensors capable of detecting micromovements of the chest wall [8]. These systems provide continuous monitoring of respiration, heart rate, and bed activity without the use of electrodes, cuffs, or wearable devices. In addition to enhancing patient comfort, contactless monitoring technologies have the potential to reduce nursing workload, support early detection of critical events, and enhance safety within care processes. However, the extent of evidence regarding the usability, acceptance, and integration of these tools into nursing workflows remains limited [9].

Aim

The present study explores the feasibility and usability of implementing a contactless digital monitoring system (Neteera 130H-Plus) [10] within a hospital environment. The aim of this study is to determine the extent to which the device integrates into the workflow of nursing personnel, to assess the device's acceptance among users, and to identify potential barriers to implementation. From the perspective of the users, the following assessments will be formulated: time savings, work simplification, satisfaction, and usability in everyday clinical practice.

Methods

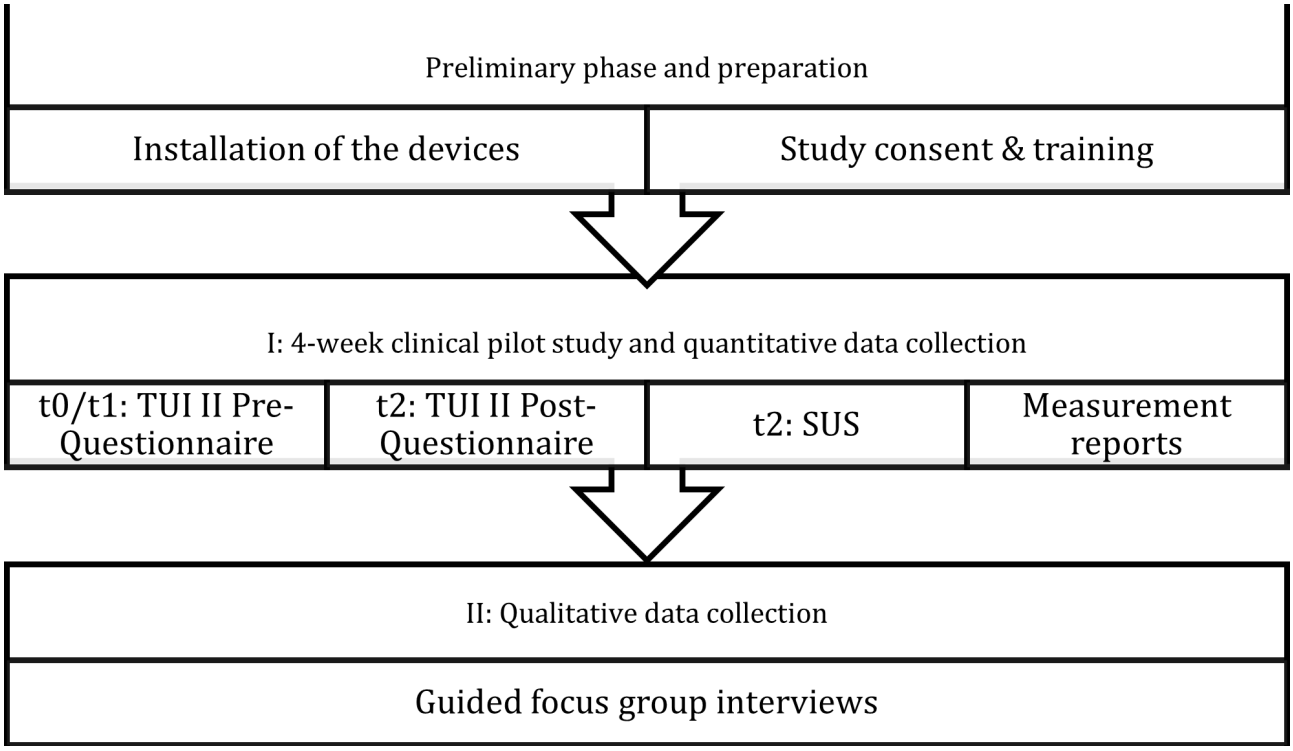
Study Design

This study was designed as a feasibility study to assess the usability and feasibility using a mixed methods approach [11,12]. The study used a sequential explanatory design, combining quantitative and qualitative data collected in 2 consecutive phases. The qualitative phase followed the quantitative phase to further explain and contextualize the results. This approach was selected to comprehensively evaluate the feasibility, usability, and acceptance of the Neteera 130H-Plus system in routine clinical nursing care. Methodological reporting was carried out using the GRAMMS (Good Reporting of a Mixed Methods Study) checklist [13] (Checklist 1).

The study was conducted under real-world conditions without altering standard clinical workflows. Patients were monitored using both the Neteera system and standard-of-care monitoring devices, enabling direct comparison without influencing clinical decision-making. Following a preparatory phase that included device installation, consent procedures, and staff training, a 4-week clinical pilot phase was conducted with quantitative data collection. The Technology Usage Inventory II (TUI II) [14] was used to assess acceptance at baseline (T0/T1) and postintervention (T2). System usability was measured using the System Usability Scale (SUS) [15] at T2. Concurrently, measurement reports and time records were generated. Subsequently, qualitative data were gathered through guided focus group interviews to contextualize and expand upon the quantitative findings (Figure 1).

XSL•FO
RenderX

Figure 1. Study design and procedure. SUS: System Usability Scale; TUI: Technology Usage Inventory.



In this study, *usability* refers to the ease of use of the system, *feasibility* to its practical implementation under routine clinical conditions, *usefulness* to the perceived benefit to nursing work, *acceptance* to the overall evaluation of the technology, and *intention to use* to the willingness to continue using the system in clinical practice.

Setting and Participants

The study was conducted at a German university hospital in 2 inpatient clinical departments: an acute dialysis department and a geriatric ward. The wards were selected to reflect different clinical settings, each with distinct monitoring requirements and workflows. The study population comprised registered nurses and nursing professionals involved in the routine monitoring of vital signs in inpatient, predominantly bedbound patients. The inclusion criteria comprised active involvement in vital sign monitoring and direct use of the Neteera system during clinical shifts. Participation in the study was voluntary and conducted during regular working hours. The study sample included all nursing staff available at the time of recruitment in the 2 wards. The sample size was selected pragmatically, in line with the exploratory nature of a feasibility study and without a formal power calculation. Patients were not actively recruited for study participation; however, they were provided with study-specific information about the technology being tested. Monitoring was performed exclusively in the context of routine clinical care. All patients continued to receive standard monitoring, independent of the study, ensuring that no clinical decisions were based on data generated by the investigational system.

Intervention

The Neteera 130H-Plus is a contactless, radar-based remote patient monitoring system designed to continuously capture health parameters (heart rate, heart rate variability, respiration rate, depth, and amplitude) and activity tracking (micromovements, time in bed, and bed exits) without direct patient contact. The device is mounted above the patient's bed (on the wall or headboard) and operates through clothing and bedding (Figure 2). All data are transmitted in encrypted form to a cloud-based platform hosted on European Union–based servers. The nursing staff underwent a structured, 30-minute training and onboarding session prior to system implementation. A live demonstration of the available functions was provided at the nursing station, followed by immediate system testing. The research staff were on site at all times to answer any questions.

XSL•FO
RenderX

Figure 2. Central positioning of the "Neteera" sensor (Neteera product photo).



Quantitative Data Collection and Analysis

Acceptance, Usefulness, and Usability

The TUI II [14] was used to assess technology acceptance and perceived usefulness at 3 distinct time points: baseline (T0, following the training), mid-study (T1, after 2 weeks of use), and end of study (T2, after 4 weeks of use). The TUI methodically assesses numerous dimensions. The TUI II is a parallel version of the TUI and is used to minimize learning effects or measurement repeatability errors when participants complete the questionnaire multiple times (eg, for preinteraction and postinteraction comparisons with a technology). At T0, the "TUI-II Pre" was administered. The questionnaire comprises 8 items, which relate to both the new technology and the general attitude toward technologies, and it allows statements to be made on the "curiosity" and "technology anxiety" scales. At T1 and T2, the "TUI-II Post" was used to assess 6 scales: "interest," "accessibility," "usefulness," "ease of using a technology," "skepticism," and "intention to use (ITU)," comprising a total of 18 items [14]. The ITU scale is a 10-centimeter-long visual analog scale used to assess agreement with given statements. The end points of the scale (eg, "agree" vs "disagree") serve as anchors for assessing the degree of agreement. The distance between the scale end point representing complete disagreement and the participant's mark is measured in millimeters for each of the 3 items. The 3 values are then summed to obtain an ITU score ranging from 0 to 300. The ITU scale assessed the ITU, intention to purchase, and desire for access to the technology. Lower scores indicate higher ITU, willingness to purchase, and readiness for access. Usability was additionally assessed using the SUS [15]. The SUS was administered after 4 weeks of system use (T2) and consisted of 10 items rated on a Likert scale.

Time Required for Vital Sign Measurement

The potential time savings were assessed by documenting the time required (in seconds) to monitor pulse and respiratory rate, both with and without the Neteera device. The documentation was carried out by a nurse in the geriatric ward, who followed a standard protocol. The following components were documented: preparation, measurement, documentation, and deviations or delays during measurement.

The quantitative data were analyzed descriptively using IBM SPSS Statistics version 2. The results were reported as mean scores and SDs according to the respective scoring manuals.

Qualitative Data Collection and Analysis

Following the testing phase and completion of the quantitative surveys, focus group interviews were conducted with nursing staff from both wards to identify experiences, perceived benefits, barriers, and implementation requirements. The focus groups were conducted using a semistructured interview guide. The guide was developed in accordance with established standards of qualitative research and tailored to the research objectives (see Multimedia Appendix 1). The interview guide explored participants' initial impressions and practical experiences with the Neteera system, with a particular focus on usability and its integration into daily nursing workflows. The investigation further examined the perceived benefits and limitations regarding patient safety, quality of care, and the influence of specific clinical settings on system use. Finally, participants were asked to reflect on future roles, necessary improvements, and overall expectations for the implementation of contactless monitoring technologies in nursing practice.

All discussions were audio-recorded, transcribed verbatim, and pseudonymized prior to analysis. Qualitative content analysis was conducted in accordance with the structured approach

XSL•FO
RenderX

proposed by Kuckartz and Rädiker [16] using the MAXQDA 2022 software (VERBI Software GmbH), integrating deductive and inductive category development. This method is particularly well-suited for the systematic analysis of extensive qualitative data. It integrates deductive category formation based on the research questions with inductive further development from the material. The transcripts were initially coded according to the interview guidelines and research objectives by one researcher and subsequently by a second researcher. Subsequently, the discrepancies in the coding were reviewed and discussed. Additional subcategories were inductively added based on the material. The transcripts were coded thematically to identify recurring patterns, perceived benefits, and challenges related to system use and integration into nursing care processes.

The qualitative findings were used to explain and contextualize the quantitative results, particularly differences in usability, acceptance, and perceived usefulness across the 2 clinical settings.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines. Ethical approval was obtained from the responsible institutional ethics committee prior to study initiation. All participating nursing professionals provided written informed consent. Data were pseudonymized and stored in encrypted form in accordance with the General Data Protection Regulation. The study received ethics approval from the ethics committee of the Medical Faculty of Martin Luther University Halle-Wittenberg (approval number 2025 - 047, April 30, 2025) and was registered at the German Clinical Trials Register (DRKS00036910). Written informed consent was obtained from all participants, and no financial compensation was provided.

Results

Pilot Study Outcomes

A total of 16 devices were installed in the 2 wards, distributed across 7 patient rooms (n=3 geriatrics, n=4 dialysis). The study staff administered 3 training sessions on site, each lasting approximately 30 minutes. A total of 21 nurses were trained in the system and its use (n=14 geriatrics, n=7 dialysis). The 4-week clinical trial phase began immediately afterward.

Quantitative Results

In the context of geriatric care, a total of 9 (64.3%) nurses completed the questionnaire survey. In the dialysis care setting, 4 (57.1%) nurses completed the survey.

Technology Acceptance

High scores on the TUI II scale indicate a high degree of expression within the designated dimension. Conversely, low scores indicate a low degree of expression [14]. Table 1 presents the mean values obtained from the scales at the 3 distinct measurement time points. For scales with 3 items, the total scale value ranges between 3 and 21 (b), and for scales with 4 items, between 4 and 28 (a). The highest values were recorded on the scales “interest” (T1=22.1; T2=20.9) and “usefulness” (T1=22.6; T2=23.6), with the former decreasing as expected and perceived usefulness being highest at the end of the trial. Skepticism decreased over the 4-week observation period (T1=8.3; T2=7.1), although it was already very low at the baseline. Due to missing responses, not all scales could be calculated in full. Consequently, the total sample size varies and is reported separately for each measurement point in column N.

Table . Technology Usage Inventory II subscale scores across measurement time points (T0-T2).

Scale	Range	T0		T1		T2	
		n	Mean (SD)	n	Mean (SD)	n	Mean (SD)
Curiosity	4 - 28	15	17.6 (5.1)	—[a]	—	—	—
Fear of using technology	4 - 28	16	8.0 (3.2)	—	—	—	—
Interest	4 - 28	—	—	14	22.1 (3.9)	15	20.9 (3.7)
User-friendliness	3 - 21	—	—	11	18.9 (1.8)	13	18.2 (2.3)
Usefulness	4 - 28	—	—	14	22.6 (5.5)	14	23.6 (3.2)
Skepticism	4 - 28	—	—	10	8.3 (3.6)	10	7.1 (3.0)
Accessibility	3 - 21	—	—	5	12.0 (3.4)	6	11.8 (3.6)

[a]Not applicable.

The ITU subscale was also examined using a group-specific evaluation (geriatric ward and dialysis) of the 3 subscales. This finding explains the substantial disparities observed, the origin of which warrants further investigation through a subsequent qualitative survey of nursing personnel. The maximum total score that can be attained is 300, with higher scores indicating a lower willingness to use. As demonstrated in Table 2, a clear discrepancy in scores is observed between the 2 wards. The geriatric ward demonstrated low total scores of 40.8 (T1) and 28.4 (T2) on the ITU scale, indicating a high willingness to use. In contrast, the dialysis respondents showed a heightened degree of caution, as evidenced by a diminished propensity to use the device with scores of 195.5 (T1) and 199.8 (T2).

XSL•FO
RenderX

Table . Intention-to-use (TUI[a] II) scores in geriatric and dialysis wards at T1 and T2[b].

TUI II item	T1, mean (SD)		T2, mean (SD)	
	Geriatrics	Dialysis	Geriatrics	Dialysis
Would you want to use this technology?	12.6 (16.0)	66.3 (37.3)	12.6 (16.0)	12.6 (16.0)
Would you purchase this technology?	19.7 (28.4)	66.8 (37.8)	16.2 (24.9)	65.0 (35.3)
Would you want to have access to this technology?	8.6 (10.1)	62.5 (34.4)	5.1 (8.1)	66.2 (37.4)
Total intention to use	40.8 (51.2)	195.5 (109.0)	28.4 (41.8)	199.8 (111.3)

[a]TUI: Technology Usage Inventory.

[b]Lower values indicate a higher intention to use.

System Usability

The mean overall SUS score was 74.5 (SD 10.5), with values ranging from 58 to 98. According to the SUS manual, a system is considered usable if it attains a score of 68, which was achieved in this study. A comparison of ward-specific scores showed that the SUS score was 76.2 (SD 12.1) for the geriatric ward and 71.0 (SD 5.7) for the dialysis ward. This indicates slightly lower perceived usability in the dialysis ward compared with the geriatric ward. The nursing staff (n=15) evaluated a series of statements using a scale ranging from 1 (“strongly disagree”) to 5 (“strongly agree”). These are presented in Multimedia Appendix 2. The system is described as straightforward to operate (ie, ease of use), with minimal training requirements. The values align with the ITU results on the TUI II scale.

Time Saving

A total of 31 measurement situations were meticulously observed and documented in the geriatric ward using a standard measurement protocol. The vital parameters of heart rate and respiratory rate were determined through the measurement of time in seconds. As shown in Table 3, radar-based measurement was considerably faster, with an average completion time of 20.3 (SD 16.2) seconds compared with 133.8 (SD 69.8) seconds for the conventional method. In 7 (22.6%) cases, delays or challenges in using the Neteera system were described. In 5 instances, the values were not displayed in real time at the desired moment. In 2 cases, it was necessary to reenter the user credentials due to a system logout that occurred in the interim.

Table . Comparison of the standard measurement method and the Neteera system (in seconds) (N=31).

Measurement method	Mean (SD, range) (s)
Manual measurement	133.8 (69.8, 50-380)
Radar-based measurement	20.3 (16.2, 6-60)

Qualitative Results

Findings from the Focus Groups

Following the completion of the trial period, 2 focus groups were conducted, comprising 4 participants in the geriatric ward and 3 in the dialysis unit. The interviews lasted between 26 and 32 minutes. Initially, primary categories were delineated in accordance with the research interests. In an iterative process, these categories were refined during the analysis and supplemented with subcategories. The coding of the 8 primary categories was conducted in close proximity to the data and was grounded in the perspectives of the nurses interviewed. The coded text passages were subsequently summarized by category and condensed for content.

Handling and Usability

The category of handling and usability includes statements on the user-friendliness of the dashboard, training, and practical limitations in everyday use. The operation of the system was characterized as uncomplicated and intuitive. Even those with limited technological expertise reported that the system was straightforward to use. The training was deemed adequate:

> *Well, I have to say, I'm from a generation that didn't grow up with computers and I don't know a whole lot about them, but I got along just fine. I think it’s easy to use. I think we can all manage it.*

However, it was noted that external factors, including bed position, sensor placement, or interference (eg, bed gallows, elevated positioning), can substantially impact transmission stability. Repeated display failures resulted in additional work and frustration, as the nursing staff were compelled to examine the cause of the missing measurements directly at the patient’s bedside. These findings are consistent with the quantitative usability results, as reflected in the overall SUS score of 74.5, indicating good usability and ease of use.

Integration Into Nursing Workflows

Integration of the system into existing workflows emerged as a pivotal category. The ability to collect data without directly disturbing patients, especially during nocturnal hours, was highlighted as a positive feature. Caregivers reported that

XSL•FO
RenderX

nighttime routine activities, such as medication preparation, were interrupted less frequently and that there was a greater sense of security because vital signs could be monitored continuously:

> *It simply means that we can collect data at night without having to wake patients.*

The system was perceived as beneficial due to its capacity to offer an overview of patients' conditions, thereby obviating the necessity for repeated in-person visits. From the participants' perspective, these measures contributed to enhanced work organization and a subjective sense of safety.

This perceived advancement in workflow integration is consistent with the high ratings achieved in the TUI II dimensions of "usefulness" and "user-friendliness" (Table 1).

Positive Usage Situations

This category comprises specific clinical application scenarios in which the Neteera system has been demonstrated to be particularly beneficial. The aforementioned participants primarily included medical and nursing issues in which progress data contributed to clinical decision-making.

One example cited was its use in patients with respiratory diseases, where changes in breathing could be monitored over time:

> *Well, I once had a patient with COPD who we had newly prescribed a new spray for, and she was always experiencing acute shortness of breath. The doctor came and asked how her breathing was, so we looked at the progression and saw what the curve looked like, and that helped us at that moment. Because we could see in black and white how things had progressed. I think that's great for things like that.*

The system was also described as helpful in situations requiring continuous monitoring in general wards, as it enabled selective, close observation without requiring continuous physical presence.

The findings from these cases align with the high perceived usefulness reported in the quantitative data (TUI II), particularly in the geriatric ward, where there was a high level of ITU.

Negative Usage Situations

Within this category, critical feedback is summarized. A large number of complaints concerned the lack of real-time measurements and unclear update intervals for vital signs. This was perceived as a significant limitation in nursing areas that performed circulation-intensive treatment procedures:

> *But we couldn't see on the tablet how often the heart rate was updated. So we couldn't tell, because it wasn't a real-time measurement, it was a fixed number that stayed the same. Our [dialysis] ECG monitoring [electrocardiogram] really shows the current heart rate, and that changes constantly.*

In addition, participants reported frequent measurement failures and significant deviations from established standard procedures (eg, electrocardiogram monitoring). The observed discrepancies have led to a decline in confidence in the accuracy of the measured values. As a result, interest in using the system decreased.

> *Everyone stood there comparing their patients, but then something didn't work and interest waned. Yes. You could tell.*

These limitations are reflected in the lower ITU values at the dialysis unit (Table 2). This finding is consistent with the independent benefit because lower ITU values can be interpreted as a prerequisite for subsequent use. The qualitative results provide a valuable explanation for the quantitative differences observed between the units, particularly the lower acceptance levels in contexts requiring real-time monitoring and high measurement reliability.

Opportunities and Benefits for the Nursing Process and Outcomes

A wide range of potential benefits of the system for the nursing process was identified. Wireless monitoring and the reduction of physical constraints were perceived as key benefits for patients. These features were expected to enhance safety and reduce agitation, particularly in patients at risk of delirium, restlessness, or falls.

> *Actually, all of the patients we have here [geriatric ward] are at risk of delirium. Patients who have delirium definitely benefit from it. Patients who have a disturbed day-night rhythm definitely benefit from it. Hypoactive delirious patients.*

The ability to perpetually assess the respiratory rate, a parameter that has not been systematically documented in conventional nursing practice, has furnished nurses' novel perspectives into the early identification of stress, anxiety, or pain:

> *Without this technology, we would not do this, of course. We would not measure respiratory rate at night. If it can be collected in this way, we can then draw conclusions and take action. Perhaps the patient is in pain, perhaps they are anxious, and so on and so forth.*

Earlier detection of nocturnal events may be possible without interrupting patients' sleep. These perceived benefits are consistent with the high ratings for "usefulness" and "interest" found in the quantitative results. The findings are complemented by the qualitative data, which illustrate how continuous monitoring can improve patient safety and support clinical decision-making under specific conditions.

Nursing Workload and Stress

The system received a mixture of favorable and unfavorable reviews with regard to its workload. In situations involving high levels of stress, the implementation of automated vital sign monitoring has been shown to result in a substantial reduction in workload, primarily by eliminating the need for time-consuming manual measurements. The use of this technology has been demonstrated to reduce the necessity for direct observation, thereby minimizing disruptions and enhancing the perception of security:

XSL•FO
RenderX

> *I would have had to check in every five minutes. We were still there, regularly, but still, you felt… you felt more at ease.*
>
> *But for patients at risk of falling who now need some kind of bedside attendance, this could of course lead to a certain reduction in staff numbers in the short term.*

At the same time, participants emphasized that these benefits depend on the system functioning reliably and consistently. These observations are consistent with the time savings identified in the quantitative data, which point to reduced measurement times. The qualitative results indicate that the benefits are contingent on the system's reliability and on further advancements in its functionality, such as blood pressure measurement.

Future Potential and Economic Aspects

The participants formulated explicit expectations for the subsequent evolution of the system. The integration of additional vital signs, especially blood pressure, was identified as an important requirement. The implementation of comprehensive digital documentation, accompanied by a direct link to the hospital information system, would also constitute a considerable development. Furthermore, reliable real-time data transmission and bedside alerting were identified as key requirements for future system development:

> *It's also ideal for some areas of nephrology. It's particularly ideal for adjusting blood pressure medication. It would be great to be able to see what happens or would happen an hour after administering medication.*
>
> *Well, I think that in order to implement this permanently, you would need improved programme stability so that there are fewer disruptive factors, and a HIS interface.*

The system was perceived as potentially cost-effective, especially compared with conventional wired monitoring systems:

> *That's very affordable, because I recently replaced some cables on a monitor and that cost €400 for just one monitor. That would be quite interesting for us.*

These expectations for further system development are consistent with the high overall interest and willingness to engage with the technology reported in the TUI II results (Tables 1 and 2). The qualitative data suggest that future adoption depends on meeting the identified technical and structural requirements.

Challenges and Barriers to Implementation

Finally, technical, organizational, and structural barriers were discussed. While nurses were generally open to digital systems and did not perceive a risk of dehumanized care, they expected some patients to be skeptical of the technology.

> *This is also extremely important for nursing staff. It has nothing to do with dehumanization or anything like that. It's simply about having the security of having data available at night as well.*

A perceived need for additional patient information was identified:

> *Many [patients] were initially worried that it was a camera, that they were being filmed. That was always their first thought, but then they were actually quite open. No one said, "No, for God's sake, I don't want that".*

The primary barriers to implementation identified included a high susceptibility to malfunction, an absence of real-time capability, inadequate interfaces, and an insufficient trial period. It was also emphasized that the system's full potential could only be realized if all rooms were fully equipped and seamlessly integrated into existing IT structures.

The identified barriers provide a contextual framework for understanding the variability in quantitative acceptance scores, particularly the lower ITU in more technically demanding settings (Table 2). Technical reliability and system integration emerged as critical factors for successful implementation.

Discussion

Principal Findings

The objective of this study was to assess the feasibility and applicability of implementing a contactless, radar-based monitoring system in an inpatient setting from the perspective of nursing professionals.

The findings indicate that the Neteera 130H-Plus system is perceived as usable and potentially beneficial for nursing care, particularly in settings where vital signs are traditionally measured manually and intermittently. While quantitative results support the system's usability, qualitative data provide additional insights into the integration of nursing workflows, perceived safety gains, and setting-specific limitations. The findings indicated notable discrepancies in acceptance and ITU between the geriatric and dialysis wards. The results underscore the pivotal role of contextual factors in the implementation of digital monitoring technologies within nursing practice.

Technology Acceptance and ITU

The quantitative results indicate a high level of acceptance of and willingness to use radar-based, contactless monitoring among nursing staff, particularly in the geriatric ward.

The SUS score indicates that the system is usable and exceeds the commonly accepted threshold [15], suggesting that the system can be handled effectively by nursing staff with limited training. Participants described the system as easy to use, suggesting a high level of usability. This finding is consistent with previous studies on digital assistive technologies in inpatient care. The extant literature demonstrates that usability is a prerequisite for sustained adoption in clinical practice [17], but it is not sufficient on its own.

The TUI II subscales "interest" and "usefulness" showed consistently high scores throughout the observation period. However, perceived interest demonstrated a slight decline, while perceived usefulness reached its peak at the conclusion of the 4-week test phase. This finding suggests that the nursing staff

XSL•FO
RenderX

regarded the implementation of the system as both practical and conducive to the effectiveness of nursing processes. Concurrently, the degree of skepticism remained minimal, underscoring a fundamental openness to the technology. A thorough analysis of the TUI II subscales reveals variations between the wards. In the geriatric ward, the values were found to be particularly low, which correspond to a high level of willingness to use the system. In contrast, the dialysis ward exhibited elevated values, indicating a more cautious approach among nursing staff. The observed discrepancy can be attributed to context-specific implementation factors. The qualitative findings indicate that acceptance was shaped not only by system usability but also by the degree to which the technology aligned with existing monitoring practices, workflow requirements, and expectations regarding real-time information. Workflow and monitoring requirements were in place due to the standard practice of continuous monitoring and real-time feedback in the dialysis ward (eg, via electrocardiogram and blood pressure measurement). Consequently, nursing staff required immediate access to relevant data that are continuously updated, particularly in situations involving medical decisions and emergency situations. The lack of reliable real-time data and concerns about measurement accuracy diminished the perceived relevance of the data for nursing decision-making. A recent systematic review described deviations in radar-based measurement methods. The review emphasizes that measurement accuracy is highly dependent on the measurement environment and conditions. Interfering factors, such as movement or device positioning, may influence measurement accuracy. The present findings are consistent with the observations made in the study [8].

Moreover, the system enables continuous vital sign monitoring without disturbing patients, a novel feature particularly relevant during nocturnal hours in the geriatric ward. The participants reported that continuous monitoring without disturbing patients contributed to a perceived sense of security among nursing staff. This finding is consistent with previous literature, which demonstrates that contactless monitoring is particularly well-received in settings involving vulnerable patients, as it enhances comfort and safety [18].

Time Savings and Impact on Workload

During the pilot phase, a key benefit identified was the reduction in time required to measure selected vital signs using radar-based monitoring. However, the potential for greater time savings remains limited in the absence of radar-based blood pressure measurements. The findings should be interpreted with caution, as the calculations were based on data from a single nurse and were not performed under controlled conditions. Further comparative measurements under controlled conditions are needed to draw reliable conclusions. The qualitative findings indicate that a reduction in interruptions, a decline in visits to the bedside, and the implementation of continuous background monitoring contributed to a subjective feeling of relief and enhanced safety, particularly during night shifts. However, nursing staff also emphasized that the efficacy of these benefits is contingent upon the stability of the system. The occurrence of technical issues, including delayed data display and forced logins, reduced the initially observed time savings. These issues also increased workload and frustration among nurses. Consequently, the efficacy of time savings appears to be contingent on the reliability of system performance [19] and the availability of blood pressure measurements.

Strengths and Limitations

This design enabled the contextualization and interpretation of quantitative usability and acceptance data through qualitative insights. The qualitative findings helped explain why most participants perceived the system as easy to learn and use despite occasional technical difficulties. The integration of 2 distinct clinical wards allowed the exploration of different nursing workflows and monitoring requirements, which is particularly relevant for feasibility research [19]. Conducting the study under real-world conditions without modifying standard care processes strengthened ecological validity and provided realistic insights into everyday nursing practice. It is imperative to acknowledge the active involvement of academically trained nursing staff, particularly those specializing in geriatrics, in the clinical setting. Consequently, while the time expenditure measurements were conducted on a limited scale and are not generalizable, they were executed under real-world conditions. In light of a single observer, one ward, and the limited number of observations, the findings related to time-measurement should be interpreted as exploratory observations rather than estimates of implementation-related efficiency gains.

Several methodological limitations should also be considered. The sample size was limited and based on pragmatic considerations, which is common in feasibility studies but restricts generalizability. Furthermore, the 4-week intervention period may not have been sufficient to establish stable routines or to assess long-term acceptance, particularly in technically demanding clinical environments.

Implications for Nursing Practice

The observed discrepancies among nursing wards underscore the notion that the efficacy of implementation is contingent upon the prevailing context. In settings characterized by acute demands for real-time data, such as dialysis nursing care, the necessity for robust and frequently updated systems is paramount to ensure trust and use. In less intensively monitored areas, such as geriatrics, the implementation of a system that provides vital parameters continuously, albeit not necessarily in real time, has been demonstrated to effectively alleviate workflow burdens, support monitoring processes, and reduce workflow interruptions from the perspective of nursing staff [20]. From the perspective of the nurses surveyed, blood pressure measurements play a decisive role in the complete mapping of basic parameters. Therefore, the potential time savings in vital sign measurement can only be partially evaluated, as blood pressure and pulse are usually measured in separate cycles. While intuitive design and ease of use are frequently associated with positive user perceptions, broader contextual factors such as workflow integration, organizational support, training, and alignment with clinical needs fundamentally influence whether technologies are sustainably adopted in nursing care environments [17,21].

Despite the system's capability for activity tracking, this feature had little influence on nurses' evaluations. The use of activity

XSL•FO
RenderX

data for clinical decision-making was limited, as these data were often perceived as less relevant than vital parameters such as respiratory rate and heart rate. In some acute care settings, nursing staff have reported that automated activity tracking is of secondary importance compared with direct clinical observation, which may reduce the added value of such systems in these contexts. The quantitative and qualitative findings suggest that activity tracking may be more relevant in long-term care or settings with lower staff-to-patient ratios, where continuous direct observation is less feasible. In these contexts, changes in movement patterns may serve as early indicators of functional decline, delirium, or fall risk [22]. The limited use of this feature in the present study does not necessarily reflect a lack of potential value but highlights the importance of adapting technological functionalities to the specific needs and constraints of the care environment.

Implications for Future Research

In order to validate these findings and investigate acceptance and usability across a broader range of clinical contexts, larger, multicenter studies are necessary. A comparative analysis of contactless monitoring systems and established monitoring technologies could provide further insights into the added value of radar-based approaches. The potential time savings observed in this study should be further evaluated using systematic cost-benefit analyses. In addition, subsequent research should concentrate on outcome-oriented end points, including the impact of continuous contactless monitoring on the early detection of clinical deterioration, the incidence of delirium, fall rates, and care-sensitive outcomes. The integration of activity tracking warrants further investigation. Finally, greater emphasis must be placed on the implementation of research findings in clinical care. A comprehensive understanding of how training concepts, technical support, interoperability with hospital information systems, and participatory implementation strategies influence long-term acceptance is imperative for translating technological feasibility into sustainable clinical practice [21,23].

Conclusions

The outcomes of this study indicate that, from the perspective of nursing personnel, contactless monitoring may be both feasible and acceptable in selected inpatient settings. A comparison of geriatric and dialysis care reveals the necessity of implementing strategies tailored to local workflows and requirements. The findings from this study, which employed a mixed methods approach, suggest that the clinical context and specific application conditions influence how new technology is perceived and used. These findings underscore the importance of aligning implementation approaches with technical performance, clinical workflows, and nursing needs. In light of the exploratory nature of the study and the methodological limitations imposed on it, further research is necessary. This additional research should be conducted in larger and more diverse clinical settings. Future studies should validate these findings and investigate long-term usability, acceptance, and integration into routine care.

Acknowledgments

The authors would like to thank the nursing professionals who supported this study and tested the monitoring system in routine clinical practice. Their engagement and practical expertise were essential for the successful implementation of the study. Furthermore, the authors acknowledge the technical staff for their support in integrating the system into the clinical infrastructure. They are grateful to the nursing directorate, the nursing leadership of the participating departments, and the staff council for their support and for enabling the implementation of this study. The authors gratefully acknowledge Ms Mary Land for her careful English editing and for improving the clarity and quality of the manuscript. AI-based language tools (DeepL Write/DeepL Academic, DeepL SE, Cologne, Germany; accessed March 2025; ChatGPT, OpenAI, GPT-4; accessed March 2025) were used for the section-by-section review of the manuscript to support language quality, clarity, and readability. All suggestions were critically evaluated, revised where necessary, and approved by the authors.

Funding

The financial support for this study was obtained from the Open Access Publication Fund of the Martin-Luther-University Halle-Wittenberg. All study-related activities were financed using institutional resources of the corresponding academic chair. The manufacturer, Neteera Technologies Ltd, provided the monitoring devices free of charge for the purpose of application testing. The manufacturing company of Neteera had no role in the study design, data collection, analysis, or interpretation; manuscript preparation; or the decision to submit the manuscript for publication.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: MR-H, SH, PJ
Data curation: MR-H, SH
Formal analysis: MR-H, SH
Investigation: MR-H, SH, BM, MH

XSL•FO
RenderX

Methodology: MR-H, SH, PJ, MH
Project administration: MR-H
Supervision: PJ
Validation: MR-H, SH, PJ
Visualization: MR-H, SH
Writing – original draft: MR-H, SH
Writing – review and editing: MR-H, SH, PJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Guiding questions for the focus groups.
[DOCX File, 22 KB - nursing_v9i1e93624_app1.docx]

Multimedia Appendix 2
System usability after 4 weeks (T2), n=15.
[PNG File, 123 KB - nursing_v9i1e93624_app2.png]

Checklist 1
GRAMMS checklist.
[PDF File, 154 KB - nursing_v9i1e93624_app3.pdf]

References

1. Cardona-Morrell M, Prgomet M, Turner RM, Nicholson M, Hillman K. Effectiveness of continuous or intermittent vital signs monitoring in preventing adverse events on general wards: a systematic review and meta-analysis. Int J Clin Pract 2016 Oct;70(10):806-824. [doi: 10.1111/ijcp.12846] [Medline: 27582503]
2. Day T, Oxton J. The National Early Warning Score in practice: a reflection. Br J Nurs 2014;23(19):1036-1040. [doi: 10.12968/bjon.2014.23.19.1036] [Medline: 25345453]
3. Downey CL, Chapman S, Randell R, Brown JM, Jayne DG. The impact of continuous versus intermittent vital signs monitoring in hospitals: a systematic review and narrative synthesis. Int J Nurs Stud 2018 Aug;84:19-27. [doi: 10.1016/j.ijnurstu.2018.04.013] [Medline: 29729558]
4. Kuhlmann J, Alhammadi E, Mevissen A, Möllmann H. Delirium and sleep disturbances-a narrative review. Z Gerontol Geriatr 2023 Nov;56(7):539-544. [doi: 10.1007/s00391-023-02228-6] [Medline: 37665376]
5. Fabrazzo M, Russo A, Camerlengo A, et al. Delirium and cognitive impairment as predisposing factors of COVID-19 infection in neuropsychiatric patients: a narrative review. Medicina (Kaunas) 2021 Nov 14;57(11):1244. [doi: 10.3390/medicina57111244] [Medline: 34833462]
6. Middleton S, McElduff P, Drury P, et al. Vital sign monitoring following stroke associated with 90-day independence: a secondary analysis of the QASC cluster randomized trial. Int J Nurs Stud 2019 Jan;89:72-79. [doi: 10.1016/j.ijnurstu.2018.09.014] [Medline: 30342326]
7. Kiser A, Korkomaz T, Ong B, Kumar AB, França K. Geriatric psychodermatology: current concepts and therapeutic strategies. Dermatol Ther (Heidelb) 2026 Feb;16(2):855-873. [doi: 10.1007/s13555-025-01638-1] [Medline: 41543823]
8. Liebetruth M, Kehe K, Steinritz D, Sammito S. Systematic literature review regarding heart rate and respiratory rate measurement by means of radar technology. Sensors (Basel) 2024 Feb 4;24(3):1003. [doi: 10.3390/s24031003] [Medline: 38339721]
9. Hassanpour A, Yang B. Contactless vital sign monitoring: a review towards multi-modal multi-task approaches. Sensors (Basel) 2025 Aug 4;25(15):4792. [doi: 10.3390/s25154792] [Medline: 40807957]
10. Shifron R. Neteera. 2024. URL: https://www.neteera.com/product [accessed 2024-11-28]
11. Ivankova NV, Creswell JW, Stick SL. Using mixed-methods sequential explanatory design: from theory to practice. Field methods 2006 Feb;18(1):3-20. [doi: 10.1177/1525822X05282260]
12. Tashakkori A, Teddlie C. Handbook of Mixed Methods in the Social & Behavioral Sciences: SAGE Publications; 2002.
13. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
14. Kothgassner OD, Felnhofer A, Hauk N, et al. TUI: technology usage inventory manual. : FFG Forschungsförderungsgesellschaft; 2012 URL: https://ucrisportal.univie.ac.at/en/publications/tui-technology-usage-inventory-manual/ [accessed 2026-08-22]
15. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry: CRC Press; 1996. [doi: 10.1201/9781498710411-35]

XSL•FO
RenderX

16. Kuckartz U, Rädiker S. Analyzing Qualitative Data with MAXQDA Text, Audio, and Video: Springer International Publishing; 2019.
17. Walzer S, Armbruster C, Mahler S, Farin-Glattacker E, Kunze C. Factors influencing the implementation and adoption of digital nursing technologies: systematic umbrella review. J Med Internet Res 2025 Jul 31;27:e64616. [doi: 10.2196/64616] [Medline: 40743516]
18. Driscoll A, Grant MJ, Carroll D, et al. The effect of nurse-to-patient ratios on nurse-sensitive patient outcomes in acute specialist units: a systematic review and meta-analysis. Eur J Cardiovasc Nurs 2018 Jan;17(1):6-22. [doi: 10.1177/1474515117721561] [Medline: 28718658]
19. van Zeist-de Jonge B, de Man-van Ginkel J, Olvers M, van den Berge K, Kooij L, Rood PJT. Nurses' experiences with inhospital continuous monitoring of vital signs in general wards: a systematic review. PLOS Digit Health 2025 Aug;4(8):e0000949. [doi: 10.1371/journal.pdig.0000949] [Medline: 40845039]
20. Donabedian A. The quality of care: how can it be assessed? JAMA 1988 Sep 23;260(12):1743. [doi: 10.1001/jama.1988.03410120089033]
21. Kraft B, Rothmann L, Zawatzki S, Hofstetter S, Jahn P. Practical training for nursing and allied health workers for a reflective approach to digitalization processes: a mixed methods study [Article in German]. Pflege 2026 Jun;39(3):140-149. [doi: 10.1024/1012-5302/a001063] [Medline: 41346189]
22. Bezold J, Krell-Roesch J, Eckert T, Jekauc D, Woll A. Sensor-based fall risk assessment in older adults with or without cognitive impairment: a systematic review. Eur Rev Aging Phys Act 2021 Jul 9;18(1):15. [doi: 10.1186/s11556-021-00266-w] [Medline: 34243722]
23. Hofstetter S, Zilezinski M, Behr D, et al. Integrating digital assistive technologies into care processes: mixed methods study. JMIR Med Educ 2024 Oct 9;10:e54083. [doi: 10.2196/54083] [Medline: 39383526]

Abbreviations:

GRAMMS: Good Reporting of a Mixed Methods Study
ITU: intention to use
SUS: System Usability Score
TUI: Technology Usage Inventory

Edited by E Borycki; submitted 17.Feb.2026; peer-reviewed by A Hari, S Walzer; revised version received 08.Jul.2026; accepted 16.Jul.2026; published 28.Aug.2026.

Please cite as:
Ritter-Herschbach M, Hofstetter S, Behr D, Hager M, Mueller B, Jahn P
Radar-Based Contactless Monitoring in Different Inpatient Nursing Care Settings: Mixed Methods Feasibility Study
JMIR Nursing 2026;9:e93624
URL: https://nursing.jmir.org/2026/1/e93624
doi:10.2196/93624

© Madeleine Ritter-Herschbach, Sebastian Hofstetter, Dominik Behr, Melina Hager, Burkhard Mueller, Patrick Jahn. Originally published in JMIR Nursing (https://nursing.jmir.org), 28.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Changes in Nurse-Patient Communication Through Health Technologies and Nursing Practices to Recognize and Support Limited Digital Health Literacy: Qualitative Study

Eline Mariosé Dijkman[1,2], MSc; Jobke Wentzel[3], PhD; David Edvardsson[4,5], PhD; Carine Doggen[2,6], PhD; Constance Helène Christine Drossaert[7], PhD

[1]Department of Surgery, Isala Hospital, Zwolle, The Netherlands
[2]Health Technology & Services Research, University of Twente, Enschede, The Netherlands
[3]Department of Health and Social Studies, Windesheim University of Applied Sciences, Zwolle, The Netherlands
[4]Department of Nursing, Swinburne University of Technology, Melbourne, Victoria, Australia
[5]Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
[6]Clinical Research Centre, Rijnstate Hospital, Arnhem, The Netherlands
[7]Department of Psychology, University of Twente, Drienerlolaan 5, Enschede, The Netherlands

Corresponding Author:
Constance Helène Christine Drossaert, PhD
Department of Psychology, University of Twente, Drienerlolaan 5, Enschede, The Netherlands

Abstract

Background: In the past decade, the use of health technologies, such as telemonitoring, video consultations, and patient portals, has increased. However, it remains unclear how these technologies have influenced nurse-patient communication. Additionally, little is known about the role nurses play in recognizing and supporting limited (digital) health literacy patients.

Objective: This study aimed to explore which health technologies are currently being used in a hospital context and how nurse-patient communication has changed as a result. Furthermore, we sought to identify the practices nurses use and the barriers they experience in recognizing and supporting patients with limited digital health literacy.

Methods: This is a qualitative descriptive study that used semistructured interviews with nurses working in a hospital (n=21). The interview guide was partly based on the 6-function model of medical communication by de Haes and Bensing. All interview transcripts were analyzed by 2 independent coders using a combination of deductive and inductive approaches.

Results: According to the nurses, health technologies have impacted all 6 functions of nurse-patient communication. They noted improvements in gathering information, providing information, enabling disease and treatment management, and responding to patients' emotions. In contrast, technology made fostering the relationship more difficult, and technologies were seldom used in shared decision-making. Nurses identified limited digital health literacy through intuition, observation of verbal and nonverbal cues, and direct questioning. To support patients with limited digital health literacy, nurses relied on building trust, involving the social network, tailoring communication, and offering additional support. High workload and limited knowledge were the main barriers to applying these practices.

Conclusions: Our findings show that health technologies have significantly influenced nurse-patient communication in the hospital setting. The results highlight the need for tailored training programs to strengthen nurses' competencies in identifying and supporting patients with limited digital health literacy. This is essential to ensure more comprehensible and accessible care and promote equitable patient engagement with health technologies.

(JMIR Nursing 2026;9:e82272) doi:10.2196/82272

KEYWORDS

health literacy; digital health literacy; nurse-patient communication; hospitals; tailored communication; eHealth literacy; nursing; hospital

Introduction

One of the essential elements of nursing care is communication with patients and their relatives [1]. Communication between nurses and patients serves different functions. The 6-function model of de Haes and Bensing [2] distinguishes the following functions from the viewpoint of the care professional: (1) fostering the relationship, (2) gathering information, (3) providing information, (4) decision-making, (5) enabling disease and treatment-related behavior, and (6) responding to the

XSL•FO
RenderX

emotions of patients. In practice, these functions overlap and interact during nurse-patient communication. Nevertheless, they are important for positive patient outcomes, including patient satisfaction, adequate diagnosis and treatment, treatment adherence, and even health [3,4].

Over the past decade, the rapid rise of the internet and digital health technologies has begun to reshape the context in which these communication functions take place. Nowadays, patients have access to health information via the internet and social media and to specific information about their own health status and treatment via the hospitals' patient portal. Apps exist for self-management, and the telemonitoring of vital functions may provide care professionals with continuous insight into bodily functions, which was formerly not possible. In addition, nursing consultations are increasingly being conducted digitally, via email, chat, or video conferencing, and health education is frequently supported by websites and videos.

These developments offer various benefits, such as improved patient insight into their health and reduced care costs and patient burden due to fewer hospital visits [5]. However, the benefits may not exist for all patients, and patients with limited health literacy may not experience the benefits from these health technologies. Patients with limited health literacy find it difficult to find, comprehend, or apply health information; they are not able to take responsibility for their health and find communication difficult [6]. The aforementioned digitalization of care requires even more skills of patients, such as the ability to operate devices and navigate on the web, understand and use health apps, and apply digitalized nursing consultations. These skills are called digital health literacy or eHealth literacy [7]. Digital health literacy is not only a technical skill set but the ability to engage with digital technologies in effective, safe, and helpful ways to achieve health goals [8]. It is influenced by sociodemographic factors, such as age, educational level, income, perception of the internet as a health source, and social support [9-11]. In the Dutch context, these inequalities are substantial: it is estimated that 1 in 3 people have a limited health literacy [12], and approximately 20% experience difficulties with digital skills [13]. It can be expected that, in a hospital setting, this number may even be higher because patients may experience tension. For example, in acute or critical care settings, patients may be sedated or intubated, limiting direct interaction, and nurses often need to communicate primarily with family members rather than the patients themselves.

Given these challenges, nurses could play an important role in the care of patients with limited (digital) health literacy, starting with their ability to recognize these patients. However, recognizing these patients is difficult, and studies have shown that nurses often overestimate the health literacy of their patients [14,15]. In addition, limited evidence exists regarding nurses' knowledge and awareness of patients' (digital) health literacy, and little is known about the practices nurses currently use to identify patients with limited (digital) health literacy in clinical practice. Once patients with limited (digital) health literacy have been identified, nurses need practices to support them. Although previous qualitative research has examined how nurses assess and support patients with limited health literacy, attention to the digital component is mostly lacking. In addition, concept descriptions of (digital) health literacy are available in the literature, and evaluations of nursing practices or interventions remain scarce. Given the increased use of health technologies, an understanding of recognizing and subsequently supporting patients with limited (digital) health literacy by nurses in hospitals is needed.

In conclusion, the increased availability of health technologies is likely to have changed nurse-patient communication. Yet, to date, only a few studies have examined how the rapid digitalization of health care specifically reshapes the 6 communication functions between nurses and patients in hospital settings. In addition, little is known about how nurses can recognize or support patients who experience difficulties in using these health technologies. Therefore, in this qualitative study, we aim (1) to explore which health technologies are used by nurses within a hospital context and how the 6 functions of nurse-patient communication have changed through the use of these technologies and (2) to examine nurses' current practices, as well as the barriers they experience, in recognizing and supporting patients with limited (digital) health literacy.

Methods

Design and Setting

A qualitative descriptive study was conducted, using thematic analysis of semistructured interviews. The interviews were conducted in Isala, a teaching hospital in Zwolle, and at Windesheim University of Applied Sciences in Zwolle, and a vocational education program at Drenthe College in Meppel, all in the Netherlands. This study is reported in concordance with the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines [16].

Participants

Purposive sampling was used to obtain a heterogeneous sample of nurses, intentionally selecting participants with diverse specialisms and experiences. Inclusion criteria were current employment at Isala, a teaching hospital in Zwolle, the Netherlands, either as a registered nurse or a nurse in training; a minimum of 1 year of clinical experience or completion of at least 1 hospital-based internship; and currently using at least 1 health technology.

Procedures

Potential participants (17 registered nurses and 6 nurses in training) were invited via email, either directly by the researchers or through their supervisor about the study's aim, interview topics, and the implications of participation. Of those invited, 2 did not respond or were unable to participate due to time constraints, while the remaining individuals agreed to participate.

Interviews were conducted a few weeks after the invitations were sent, either face-to-face or via video call. Before each interview, participants provided written informed consent.

Interviews took place between November 2022 and January 2023 at the hospital and were conducted by 2 researchers (EMD and JW), both trained in qualitative research. One researcher

XSL•FO
RenderX

was employed as a nurse researcher at the hospital, and the other was a lecturer at the University of Applied Sciences. A pilot interview was conducted to assess the clarity and structure of the interview guide, leading to minor adjustments in question sequencing.

A total of 21 interviews were conducted, with 17 held in person and 4 online. All interviews were audio-recorded with participants' permission and transcribed verbatim. During most interviews, a student assistant was present to help manage time and ensure all questions from the interview guide were addressed. Interviews lasted between 25 and 54 minutes.

Interview Guide

The interviews were conducted using an interview guide, including 3 main topics.

The Use of Health Technologies and Their Impact on the 6 Functions of Patient-Nurse Communication

The data were gathered using the 6-function model of communication (Table 1) [2].

Table . The 6-function model of communication [2].

Six functions	Explanation
Fostering the relationship	Emphasizes the importance of mutual trust, understanding, and commitment, as well as agreement about each other's roles and expectations
Gathering information	Aims to ascertain a correct interpretation of symptoms and establish an adequate nursing diagnosis
Providing information	Aims to provide understandable, accurate, sufficient, and timely information in a manner appropriate to the patient
Decision-making	Aims to make sure decisions are being based on the information and preferences of patients
Enabling disease and treatment-related behavior	Aims to support patients in treatment adherence and a healthy lifestyle
Responding to emotions	Requires nurses to elicit patients' emotional distress, communicate an understanding of the patient's emotions, and respond with empathy and support

First, nurses were asked how each function was part of the nurses' work and if any health technologies were used in performing this function. Health technology was defined as the application of organized knowledge and skills in the form of devices, medicines, vaccines, procedures, and systems developed to solve a health problem and improve quality of life [17]. An infographic with examples of health technologies (such as video calling, patient portals, and digital leaflets) used in nursing care was presented to stimulate discussions about their own experiences with these health technologies and experiences with other technologies they may have used. Furthermore, nurses were asked if and how these technologies have changed the specific function. Each participant was interviewed about at least 2 functions. If there was enough time, nurses were asked about more than 2 functions. These 2 functions were rotated between interviews so that all functions were addressed equally.

Current Practices in Recognizing and Supporting Patients With Limited (Digital) Health Literacy

The critical incident technique [18] was used to gain insight into current practices in recognizing and supporting patients with limited (digital) health literacy. Nurses were asked to describe a situation where the communication with a patient with limited (digital) health literacy failed, followed by several follow-up questions aimed at elaborating on the situation (eg, "What happened?," "What went wrong?," and "What did you do to fix it?"). Then, the opposite was asked to describe a situation wherein the communication was very effective. Again, nurses were encouraged with follow-up questions to elaborate on the situation. Furthermore, some additional questions were asked to gain insight into their current practices regarding recognition, such as, "Can you assess whether a patient you are seeing for the first time can work with that technology?," and support, such as, "How do you ensure that a patient with limited digital health literacy can work with this technology?"

Barriers to Recognizing and Supporting Patients With Limited (Digital) Health Literacy

Nurses were asked which barriers they experienced in recognizing and supporting patients with limited (digital) health literacy. For example, the following questions were asked: Which barriers do you experience in recognizing patients with limited digital health literacy? And how is this for supporting patients with limited digital health literacy?

Data Analysis

The data were analyzed using a thematic analysis [19] using Atlas.ti 22. Two researchers (EMD, JW) first independently read all transcripts to familiarize themselves with the data and identified relevant text fragments. Coding was conducted independently, and coders met after 3, 6, and 12 interviews to discuss their findings and refine the coding scheme. Any discrepancies were resolved through discussion, and a third researcher (CHCD) was consulted when necessary to reach consensus. After each round of discussion, the coding scheme was further refined. No formal intercoder reliability metrics were calculated, as this is not standard for thematic analysis, in

XSL•FO
RenderX

which the focus is more on finding patterns of meaning rather than on numerical agreement. We combined deductive and inductive approaches: deductive coding was guided by the main research questions, while inductive coding allowed subcategories to emerge from the data. Text fragments were first sorted (deductively) into categories in line with the research questions: changes in the 6 functions of communication, practices to recognize limited (digital) health literacy, practices to support patients with limited (digital) health literacy, and experienced barriers in recognizing and supporting patients with limited (digital) health literacy. Within each category, subcategories were developed inductively, allowing themes to emerge from the data.

Data saturation was assessed following Hennink et al [20], who distinguish between code saturation, the point at which no new codes are identified, and meaning saturation, the point at which a richly textured understanding of the themes is achieved. In our study, no new codes or themes emerged after 16 interviews, indicating that code saturation had been reached. A total of 21 interviews were conducted, ensuring that all research questions were comprehensively addressed and meaning saturation was also achieved.

Ethical Considerations

The study was conducted following the ethical principles of the Declaration of Helsinki. The Medical Research Ethics Committees confirmed that the study (reference number 20221009) is not subject to the Medical Research Involving Human Subjects Act (WMO). Participants provided written informed consent prior to the interview. Participants were able to decline to answer any questions and withdraw from the study at any time up to the time of data analysis. All participant-related data were deidentified and kept by the researcher. Only research team members had access to the data. No compensation was provided, apart from reimbursement for the time participants spent participating in the interview.

Results

Characteristics of Participants

The characteristics of the participants (n=21) are presented in Table 2. One nurse worked at the virtual monitoring center, where patients at home are monitored remotely. A chief nursing information officer and 2 nurse information officers also participated; these roles aim to support the integration of health technologies into nursing practice. Additional specialized roles included an electronic medical record (EMR) key user, who acts as a mediator between EMR developers and clinical staff, and a digital health technology coach (“digicoach”) who supports nurses in using digital tools. No compensation was provided, apart from reimbursement for the time participants spent participating in the interview.

XSL•FO
RenderX

Table . Characteristics of participating nurses (n=21).

Characteristics	Participants, n
Age (y)	
20 - 29	10
30 - 39	5
40 - 49	5
≥50	1
Sex	
Male	4
Female	17
Working in	
Clinical wards	
Abdominal and vascular	2
Urology	1
Oncology	1
Dialysis	1
Intensive care	2
Cardiology	2
Nurses in training across various clinical wards	6
Outpatient specializations	
Pulmonary	1
Wound	1
Oncology	1
Ostomy	1
Virtual monitoring center	1
Other	
Chief nursing information officer	1

Health Technologies in Patient-Nurse Communication

Nurses indicated using various health technologies, both in the hospital and in the home environment. For example, at the hospital, nurses are using the EMR to report about the patient's condition, use videos to support their education for patients, and use a projector with images (eg, nature, animals, or cities) to calm patients. The mentioned health technologies were divided into 5 categories: e-interaction, e-monitoring, e-education, e-recording, and e-support (Table 3).

Table . Overview of health technologies used by the hospital nurses.

Category of health technologies	Health technology
E-interaction	Video consultation, e-coach, translator, email contact (including sending photos), phone call, and notifications for nurses via the app
E-monitoring	Distant monitoring of measurement directly transferred to EMR[a], online questionnaire about self-reported symptoms, monitoring exercise behavior, and camera surveillance
E-education	Images, digital leaflets, videos, webinars, and websites
E-recording	Electronic medical record, patient portal, and digital diary for family
E-support	Projector with calming images, virtual reality, music, online church services, and VR[b] cycling: stationary bike combined with a screen

[a]EMR: electronic medical record.

[b]VR: virtual reality.

XSL•FO
RenderX

Health Technology Influence on Patient-Nurse Communication

Overview

We analyzed nurses' responses regarding whether and how health technologies have affected the 6 functions of communication. According to the nurses, health technologies influenced all 6 functions of nurse-patient communication. They reported positive effects on gathering information, providing information, supporting disease and treatment management, and responding to patients' emotions. In contrast, technology was perceived as making the development of the nurse-patient relationship more challenging, and it was rarely used to support shared decision-making. The results are summarized in Table 4, with + indicating positive changes, – negative changes, and ± changes that were neither clearly positive nor negative, reflecting neutral effect.

Table . How health technologies have positive, negative, or neutral changed the 6 functions of communication.

Functions	Changes
Fostering the relationship	• Increases patients' sense of safety (eg, through camera surveillance) (+[a]) • Lack of eye contact or nonverbal communication (±[b]) • Creates more distance between patient and caregiver (eg, video consultations) (–[c]) • Requires more time to build a relationship (–) • Patients feel less listened to (–) • Less privacy for a patient (eg, through camera surveillance) (–)
Gathering information	• Data are more structured in EMR[d] (+) • Allows measurements to be recorded in the EMR with the patient present (+) • Increases safety through automatic transfer of vital signs into the EMR (+) • Quick and complete data collection via online questionnaires before or during admission (+) • Continuous data collection at home (+) • Reduces number of hospital visits through video consultations (+) • More insight into the home environment (+) • Email consultations offer the opportunity to respond whenever suitable (+) • Enables quick and adequate treatment in the home situation through, eg, sending photos of ostomy (+) • Missing "the clinical view" (±)
Providing information	• Increases patient insight into medical data through, for example, portal (+) • Provides relatives insight into patient well-being via "digital diary" (+) • Increases options for multimodal education, which enhance patient understanding (eg, videos) (+) • Can be sent in advance to enhance patient preparation (eg, digital leaflets) (+) • Check patient understanding in e-education (±) • Check patient understanding in e-education (±) • Some images are too complex and reduce clarity (–)
Decision-making	• Access to (personal) data can contribute to informed decision-making (+) • Retrieving information from EMR is simplified (+) • Online decision aid can contribute to shared decision-making (+)
Enabling disease and treatment-related behavior	• Exercise options to stimulate the patient to healthy behavior (+) • Requires patients to take a more active role through remote monitoring (±)
Responding to emotions	• Can contribute to distraction, for example, through VR[e] (which can reduce pain) (+) • Enables emotional expression via iPad emoticons (+) • Builds reassurance through better understanding of condition via e-education (+) • Enables nurses to inform their colleagues about patients' emotions fostering coordinated care (+) • Enhanced comfort in home situation (+) • Emotional support after admission through, for example, messages (+) • Requires attentive listening to the emotions expressed, as visual cues are not available (±) • Technology can create disruptions, for example, nurse paging systems (–)

[a]+: positive change.

[b]±: changes that were neither clearly positive nor negative.

[c]–: negative change

[d]EMR: electronic medical record.

[e]VR: virtual reality.

XSL•FO
RenderX

Fostering the Relationship

Building trust and fostering the relationship between nurse and patient constitutes the basis for successful nurse-patient communication. According to participants, this function has changed considerably with the increased use of health technologies, such as video consultation and notifications via patient portals and apps, and mostly in a negative way. It causes a lack of eye contact and nonverbal communication, creates more distance between patients and nurses, and requires more time to build a relationship. Some nurses mentioned that their patients feel less listened to because the nurse is typing during the conversation to add results into the EMR. Regarding e-monitoring, nurses thought that this would increase the patient's sense of safety. However, a negative consequence was also noted, namely that (camera) surveillance could compromise privacy. Finally, the nurses indicated that a good relationship could be established if it starts with a face-to-face interaction before the technology is being used, as is illustrated in the following quote:

> *It also does bring a bit of distance but I think you can reduce that distance if you have created trust during a [face-to-face] consultation at the outpatient clinic. And, if you have taken time for that.* [R10]

Gathering Information

The second function of nurse-patient interaction is gathering information: nurses need to gather information about symptoms and health status to reach an adequate diagnosis and treatment (care) plan. According to the nurses, this function has mostly positively changed with the availability of health technology. For example, e-recording provides more structured data of patients and allows a quick search for information, and the possibility to report the provided information into the EMR in the presence of the patient reduces the chances of mistakes. Automatic measurement and transfer of vital signals into the EMR increases safety and reduces workload for nurses. The use of online questionnaires allows for quick and complete gathering of data before or during hospital admission and continuous data collection also at home. The positive changes of e-interaction (eg, video consultations) included more insight into the home environment of the patient and that patients have to come less often to the hospital. In addition, nurses indicated that online consultations are less often interrupted, and email consultations or messages offer the opportunity to respond whenever suitable. For patients with ostomy or wounds, especially, the opportunity to send photos enables quick and adequate treatment in the home situation.

As a downside, nurses emphasized that by e-interaction, they sometimes miss "the clinical view." This makes listening skills particularly important, as is outlined in this quote:

> *It is very different to have a clinical view over the phone or email than live, where you can also watch someone. Yes, you have to listen very carefully and you have to start asking a lot more questions, so you don't just look at how someone is sitting, but you also ask, what did you just do when someone is out of breath? Because I didn't see all that. Where are you right now? Are you standing? So I have to start asking very specific questions about a measurement.* [R8]

Providing Information

The third function of nurse-patient interaction is information provision: nurses have to provide clear information tailored to the patient, for example, information about symptoms, diagnosis, disease, and treatment. Nurses felt that this function was mostly changed in a positive way. Regarding e-recording, the patient portal provides more insight for patients into their personal medical data, and a "digital diary on the ward" provides patients and their relatives insight into nurses' interpretation of patients' well-being and daily activities. Regarding e-education, websites, visualizations, and videos provide options for multimodal education and thereby enhance patients' understanding, as is illustrated in the following quotation:

> *A video with illustrations is a much better tool and easier than a leaflet, I think. Many leaflets that I hand out, turn out to be still unread at the bedside table by the end of the admission.* [R5]

Moreover, education via the internet can easily be sent in advance, enhancing patient preparation and improving personalized communication. On the other hand, one nurse indicated that some images are so difficult that they make the information less understandable. When oral education is replaced by e-education, for example, digital leaflets, it is important to check if the information is found and understood by the patient.

Decision-Making

Nurses play a pivotal role in facilitating shared decision-making by providing information and taking into account preferences. Currently, nurses use a limited number of health technologies in the decision-making process. Regarding e-recording, patient access to (personal) data can contribute to informed decision-making, and retrieving information by nurses from EMR is simplified. E-support, such as online decision aids, can contribute to shared decision-making. Although most interviewed nurses had not yet incorporated online patient decision aids into their daily practice, these technologies were recognized as potentially valuable additions for shared decision-making with patients. This potential valuable was described by one of the nurses as follows:

> *They exist [Patient decision aids], but are not used very often in practice. I was introduced to a digital colorectal cancer decision aid, explaining types of cancer and the advantages and disadvantages of different treatments, help patients choose what they want. I liked that, but in practice, it is hardly used.* [R11]

Enabling Disease and Treatment-Related Behavior

The fifth function includes enabling disease and treatment-related behavior, which involves supporting patients in treatment adherence and promoting a healthy lifestyle. Nurses described different changes in this function resulting from the increased use of health technologies. About e-support, the

XSL•FO
RenderX

implementation of virtual cycling increases exercise options to stimulate the patient to healthy behavior.

Regarding e-interaction, the use of health technologies, such as apps and sending photos by email, facilitates a more active role of patients as they are required to make choices and take actions on their own. For example, patients have to make medication intake changes or have to make changes in the wound care treatment, such as stopping flushing the wound. This more active role is mostly regarded as positive but also requires extra effort and skills of the patient, as becomes clear from the following quote:

> *I always say [to patients]; care at a distance is good care, but it is care in which you are in charge, you have to take the steps yourself. But the patient also needs to understand those steps.* [R10]

Responding to Emotions

Patients often experience a range of emotions, such as anxiety, anger, and sadness, as a result of their disease. Nurses can play a crucial role in identifying and responding to these emotions, which can provide patients with a sense of support, but is also a basis for other functions, such as adequate information exchange, decision-making, and self-management. Nurses mentioned that health technologies mainly caused positive changes in this function. Technologies mentioned under e-support included music, virtual reality, video calling with relatives, and online church services that can provide distraction and support to patients. For example, one nurse explained how projected images could distract and calm patients:

> *For example, you find yourself in a forest that was really an experience, and you notice that patients are really happy with that, so in that way, yes, offering something with images and sound, brings distraction. Being in another world for a moment.* [R2; an image projector was used]

In a similar way, virtual reality can lower the experienced pain, for example, during wound treatment.

In addition, if patients are not able to talk, an app was used that allows them to digitally indicate emotions by selecting an emoticon on an iPad. This technology makes it possible to express emotions without using voice. E-education can establish reassurance because a better understanding of their condition or treatment can reduce emotional stress. Using e-interaction can contribute to enhanced comfort in the home situation and create the opportunity for emotional support after an admission. Additionally, as a result of new opportunities to communicate digitally with patients (such as via messages, phone consultations, and emails), nurses are required to attentively listen to the emotions expressed by patients as visual cues are not available. Finally, e-recording, such as the EMR, enables nurses to inform their colleagues about patients' emotions and feelings, fostering comprehensive and coordinated care. On a downside, nurses also admitted that interruptions caused by health technologies, such as nurse paging systems, can affect the emotional support provided by nurses.

Recognizing Patients With Limited Digital Health Literacy

In the interviews, nurses describe different practices to recognize a patient with limited digital health literacy, which could be divided into three categories: (1) intuition, (2) observation of verbal and nonverbal signals, and (3) explicitly ask and register in EMR (Table 5).

XSL•FO
RenderX

Table . Current practices in recognizing patients with limited (digital) health literacy.

Current practices and subthemes	Examples
Intuition	
Intuition	• Gut feeling, intuition
Observe verbal and nonverbal signals	
Patient characteristics	• Advanced age • Low educational level • Poor lifestyle • Comorbidities
Depending on network	• Social network • Care network
Behavior	• Non-compliant with advice • Not having a critical attitude • No motivation or not taking initiative • Asking less or irrelevant questions • Use of language: choice of words, writing style • Not understanding or not reading health information • Only using a few health technologies or difficulties with using them • Not willing to use health technologies • Excuses or avoiding information or do not want any information
Explicitly ask and register in EMR[a]	
Ask questions to identify limited digital health literacy	• Explicitly ask about digital skills and health literacy • Apply a validated assessment tool
Ask patients to demonstrate how they use health technologies	• Instructions are not carried out sufficiently • Ask the patient to demonstrate and observe difficulties
Register in the EMR	• Low literacy registered

[a]EMR: electronic medical record.

Intuition

One frequently used answer to the question on how nurses identify patients with limited (digital) health literacy was intuition. Twelve nurses mentioned that they use their gut feeling to assess if patients have limited or adequate (digital) health literacy. It was difficult for them to explain what this was based on, as is illustrated in the following quote: “You have a gut feeling that the patient is going to perform it properly, but on what do I base that ... I don’t know” (R3). Several nurses acknowledged that this nursing practice is quite uncertain and can lead to incorrect assessments.

Observe Verbal and Nonverbal Signals

Other nurses were more explicit about the cues they used for inferring limited literacy or digital literacy in their patients. Certain patient characteristics were mentioned, including advanced age, low educational level, multiple comorbidities, and an unhealthy lifestyle. Some of these characteristics were observed, such as being overweight, advanced age, smoking, and shabby clothing. Other characteristics were indirectly deduced, such as a low educational level and the presence of comorbidities.

Nurses also identify patients with limited digital health literacy based on the level of dependence on their social or care network. They either asked or observed during a hospital admission the extent to which patients rely on their family or acquaintances, for example, with reading or understanding health-related information or completing online questionnaires. Hospital nurses can collaborate with home care nurses or health care professionals from other departments to gather additional information of the patient.

Another nursing practice to identify patients with limited (digital) health literacy is by observing certain behaviors or verbal cues, such as lack of critical attitude, noncompliance with advice, asking irrelevant questions, and not being motivated. The following quote explains how a lack of a critical attitude is being used as a cue for limited (digital) health literacy:

> *I observe how critical a patient is, or whether he will let everything happen to him. For example, do they take their medication willingly, or do they ask me, what is it or what is it for? I take that into account in my assessment.* [R1]

Explicitly Ask and Register in EMR

Finally, nurses sometimes explicitly asked or assessed (digital) health literacy levels with questions or observed the patient while using the technology. An example of a question that was typically used by nurses to identify patients with limited health

XSL•FO
RenderX

literacy was as follows: “Do you feel overwhelmed by the abundance of health information?”

Digital skills were mostly identified with straightforward questions about the patients’ use of email, WhatsApp, or social media. Another nursing practice to identify limited digital literacy was to observe how a patient works with health technologies:

> *If we have the equipment on the ward to let them do it themselves and then I will observe them: How does the patient handle it, does the patient know how it works, what does someone do with the explanation? Does it then go right at once or does the patient need more frequent explanations? That does say something about the skills.* [R5]

Formal assessment tools to measure limited (digital) health literacy were never used, and most of the interviewed nurses did not know about the existence of such tools. During the interviews, it also became clear that low literacy is not or only scarcely registered in the EMR. Yet, several nurses indicated that it would be helpful to report about (digital) health literacy in the EMR to facilitate awareness and enable them to tailor care accordingly. Documentation in the EMR would also ensure that other involved health care providers would be informed so that the patient does not have to be repeatedly questioned about these skills.

Supporting Patients With Limited Digital Health Literacy

In daily practice, nurses use certain practices to support patients with limited (digital) health literacy in nursing care. Four main practices were identified: (1) create trust, (2) tailor communication, (3) provide additional support, and (4) use the patient network (Table 6).

Table . Current practices in supporting patients with limited (digital) health literacy.

Current practices	Subthemes
Create trust	• Create trust among patients in the health professional • Support the self-confidence of the patient • Normalize having limited (digital) health literacy
Tailor communication	• Repeat information • Limit the amount of information and use short explanations • Plain or easy language • Visualize • Use additional methods; write information down or use a whiteboard • Check if information given to the patient is understood (teach-back) • Step-by-step explanation • Clear instructions • Ask open questions • Develop tailored health technologies
Provide additional support	• More frequent contact (phone or face to face) • Provide a summary of the provided education and instructions • Create a place where patients can ask questions about the use of health technologies • Offer the nondigital method
Use the patient network	• Involving relatives to support the patient in using health technologies • Asking relatives to use health technology instead of the patient • When various relatives are receiving the information, relatives can also repeat it for the patient • If the patient does not understand the explanation, the nurse can use the patient’s network • Relatives can help the patients formulate questions • If no relative is available, arrange home care • Consult and coordinate with other caregivers

Create Trust

To support patients with limited (digital) health literacy, 5 nurses indicated that it is important to create trust. This practice included addressing patients’ distrust of health technologies, supporting and increasing patients’ self-confidence, and also creating trust in health professionals. Nurses mentioned that these are important skills to support patients in understanding health information and support patients in using health technologies. It is crucial to normalize having limited (digital) health literacy and inform patients that many patients are experiencing difficulties in understanding and applying (digital) health information. This was explained by one of the nurses as follows:

> *Some patients ask me, why are you asking if I can’t read? I always say, there are quite a lot of people in the Netherlands who cannot read at all. And if that is the case, we like to take that into account because then we can also show you information in a different way.* [R10]

XSL•FO
RenderX

Normalizing having limited (digital) health literacy hopefully contributes to less shame about having limited (digital) health literacy and makes it easier for the patient to ask questions or to indicate if something is not fully understood.

Tailor Communication

Different types of practices exist to tailor the communication to patients with limited (digital) health literacy. These include repeating information, using short explanations, using plain language, and visualizing information. Additional methods to support the patient's understanding of health information are to write down information or draw pictures and use a whiteboard in a hospital patient room. Certain nurses indicated using the teach-back method to ensure the patient's understanding. This method asks patients to state in their own words what they understood or remembered from the given health information [21]. Even though this method was mentioned, it was unclear whether all nurses performed it correctly, as several nurses subsequently indicated they were only asking if a patient still had any questions.

Provide Additional Support

Providing additional support is described as a useful nursing practice by the majority (n=16) of the nurses. An example was frequent telephone contact to establish more guidance and more opportunities to practice with health technology. Another suggestion was to create a support team who can be reached by phone, where patients can ask questions related to the use of health technologies.

In addition, nurses described that face-to-face communication instead of digital communication should remain possible, recognizing that not all patients may be capable of using health technologies even with extra support. Some nurses described that patients who were not willing or able to use health technologies were excluded from using health technologies. This means that these patients do not have the opportunity to participate and experience the positive effects of it. One nurse indicated that patients were never excluded from home dialysis, including the use of health technologies due to insufficient skills. Excluding none of the patients was a conscious decision, as long as the patient was willing to learn and try. Consequently, this requires nurses to invest time, patience, and creativity in guiding these patients.

Use the Patients' Network

The majority (n=15) of nurses reported that patients' social networks are also utilized to support patients with limited (digital) health literacy. For example, relatives can assist patients in working with health technologies and repeat important educational instructions: "But you can also check if the son, daughter, or grandchildren might know a little more about technology and could offer help" (R18). Besides, the health care network of a patient (eg, home care) can help if there are no relatives available. Consulting other caregivers about the care for the patient can lead to comprehensive care and appropriate support.

Barriers to Recognize and Support Patients With Limited (Digital) Health Literacy

Several barriers to recognize and support patients with limited (digital) health literacy exist. Regarding the work environment, one frequently mentioned barrier was the experienced workload. The continuous high workload often leaves limited time to identify patients with limited (digital) health literacy and provide tailored support. Moreover, nurses often see patients only for a very short time, which makes it particularly difficult to get to know the patient very well:

> *You don't know the patients when they enter the hospital and just being admitted. You have to get to know them first. That is sometimes difficult.* [R4]

Related to the patient, the nurse described that numerous patients feel ashamed about having limited (digital) health literacy and try to disguise this. This makes recognition more difficult. Another barrier is the family members who speak consistently on behalf of patients instead of giving the patient the opportunity to talk. This may make it difficult to gather the right information because it may be the opinion of the family rather than that of the patient. The last barrier is nurses' own lack of knowledge. Nurses indicate that they have little knowledge about recognizing and supporting these patients. As a result, they often rely on intuition.

Discussion

Principal Findings

The results of this study show that nurses in hospitals use a wide variety of health technologies, and this has affected the nurse-patient communication significantly, mainly in a positive way. Nurses mentioned intuition, observing verbal and nonverbal signals, and explicitly asking as nursing practices to identify patients with limited digital health literacy. Nursing practices to support patients with limited (digital) health literacy included building trust, tailoring communication, providing additional support, and engaging the patient's network. Several barriers were identified in recognizing and supporting patients with limited (digital) health literacy, including a high workload, lack of knowledge, and not knowing the patient.

Our results revealed that according to the nurses, health technologies affected all 6 functions of the de Haes and Bensing's communication model [2]. Positive consequences were seen in gathering and providing health information, enabling disease and treatment management, decision-making, and supporting patients to handle their emotions. These results are in line with previous research reporting benefits such as reduced unnecessary consultations, improved accessibility for patients, and the potential of messaging systems to facilitate nursing care [22. While our study primarily observed positive changes, we also identified certain challenges. These included, for example, verifying patients' understanding and addressing complex images or content that required additional explanation for patients. Similar challenges have been reported in other qualitative studies, including the additional workload associated with training patients [23]. These studies have further highlighted nurses' limited digital skills, as well as the rapid

XSL•FO
RenderX

pace of technology implementation [24]. Although overall changes were largely positive, it was notable that our study identified predominantly negative changes within the function of fostering the nurse-patient relationship. Similar challenges were reported in a study using focus groups with primary health care nurses exploring their views on eHealth tools for patient self-management [24]. Nurses expressed concerns about losing personal contact, perceiving digitalization as a potential threat to the nurse-patient relationship [24]. To overcome this challenge, nurses in our study recommended always starting with a face-to-face meeting and switching to digital communication only after the relationship is built. The preference for initial face-to-face consultations, as well as for combining digital and in-person consultations, has also been noted previously [25]. Our findings suggest that task-oriented functions of the 6-function model, such as providing health information and supporting disease management, benefit from technology through increased efficiency, accessibility, and structure. In contrast, relational functions, such as fostering the nurse-patient relationship, are harder to support digitally, as they rely on nonverbal cues and trust-building [2].

To the best of our knowledge, this study is the first to examine the impact of health technologies on nurse-patient communication using the 6-function model of communication [2]. Although the 6-function model was not specifically designed for use in the nurse-patient communication context, it proved to be highly usable, and nurses could easily relate to the 6 identified functions. This is in line with a previous study in which the model was used to examine nurse-patient communication [26]. Although the model was also not developed to study the impact of health technologies, it turned out to be helpful in providing insight into their impact on the various functions of communication.

In nurse-patient communication, it is important to identify the patients who experience difficulties in using health technologies. Patients with limited health literacy are less likely to use health technologies or perceive them as difficult and useless [27], while these patients can actually benefit from them, for example, from the visualization of health education through videos. To reduce exclusion of patients with limited (digital) health literacy, nurses can play a pivotal role in recognizing and supporting these patients. Nurses in our study mentioned different practices to recognize patients with limited (digital) health literacy. These are intuition, observed verbal and nonverbal signals, and explicitly asking patients about their (digital) health literacy skills or asking them to demonstrate the use of the health technology. Previous studies also showed that nurses often use their intuition to estimate health literacy [28] and use verbal and nonverbal cues, such as patients' use of simple language, facial expressions, or level of interaction [29]. Unfortunately, "intuition" and "observing verbal and nonverbal signals" are very subjective. Relying only on intuitions may not be an appropriate practice, since it is known that nurses frequently overestimate the health literacy of their patients [30]. Some verbal and nonverbal signals, such as lower education, low self-management, less motivation, and low self-efficacy, are indeed associated with limited (digital) health literacy [31,32], but this does not apply to every individual patient and may be difficult to observe. Moreover, for some nonverbal signals, such as looking slobby, evidence is lacking. A potentially more reliable nursing practice is to explicitly ask patients about their (digital) health literacy skills or to demonstrate how they use technology. Although the importance of this was acknowledged by a few of the interviewed nurses, it was not common practice. A reason for being hesitant to ask patients about their health literacy is that many patients feel ashamed about having limited (digital) health literacy [33]. Our results highlight the importance of educating nurses on how to assess patients' (digital) health literacy in a neutral and nonoffensive manner, so as to avoid causing discomfort or feelings of shame.

Interestingly, our results showed that patients' levels of (digital) health literacy are only occasionally documented in the EMR, despite the availability of numerous questionnaires to assess digital health literacy [34]. The systematic documentation of this information could reduce the need for repeated assessment and could enhance continuity of care among health care professionals within hospital settings.

To support patients with limited (digital) health literacy, nurses used different practices: building trust, tailoring communication, providing additional support, and engaging the patient's network. To tailor communication, the "teach-back method," already proven to be effective [21], was often mentioned. Unfortunately, it appears that the teach-back method is not always performed correctly. A recent review including similar health literacy practices as in our study showed positive effects on patient support, including activation, patient comprehension, and engagement [35]. Strategies often included multimedia or technology-based approaches, simplification of written material, and facilitation of in-person sessions. More research is needed to examine if these practices are actually effective and how nurses can be motivated to use these in daily practice.

Our study revealed that many nurses are experiencing multiple important barriers while recognizing and supporting these patients. These include a high workload, lack of knowledge, and not knowing the patient. For some of these barriers, training could be helpful. Training, encompassing various educational programs and online resources, has been identified as an effective intervention that significantly contributes to improving digital health literacy [36]. Several studies have shown that offering training for nurses on how to recognize and support patients with limited health literacy shows positive effects on knowledge, attitude, confidence, and increased usage of strategies to recognize and support patients with limited health literacy [37-43]. Unfortunately, the digital component is often not included (yet) in these existing training. It is recommended that nursing education and existing training programs systematically integrate the digital component, with particular emphasis on enabling nurses to recognize and support patients with limited (digital) health literacy.

Strengths and Limitations

A strength of this study is that a heterogeneous group of nurses from different wards and various specializations was included. This contributed to more generalizable results regarding nurses working in hospitals. This study was conducted in the Dutch hospital context, where the implementation of health

XSL•FO
RenderX

technologies is relatively advanced. However, it remains unclear how these findings translate to other countries, as differences in health care systems, digital developments, digital literacy of nurses, and organizational structures may influence the adoption and impact of health technologies. Further research in diverse international contexts is needed to explore their applicability beyond the Netherlands.

In this study, we selected participants who had at least some experience with health technologies to explore their perspectives based on actual use. As a result, our findings primarily reflect the experiences of nurses with some engagement in digital technologies. Future research including nurses with little or no prior experience could provide additional insights into potential barriers and challenges.

Future Research

Since this study primarily used self-reported data, it would be valuable to observe actual nursing behaviors in practice, for example, how frequently nurses use strategies to recognize and support patients with limited digital health literacy. Furthermore, gaining a deeper understanding of patients' perspectives is equally important. Patients with limited digital health literacy may experience challenges with health technologies that differ from or overlap with those faced by nurses. It is still unclear to what extent patients recognize potential disadvantages or barriers related to technology use. Therefore, future research should explore the specific difficulties patients encounter and the types of support they need to effectively engage with digital health tools.

Besides the patient perspective, future research may explore potential differences across hospital departments, as communication challenges may vary by clinical context. For example, nurses working in intensive care units may face distinct communication challenges, such as needing to communicate primarily with family members or navigating ethical considerations related to digital consent.

Providing training to nurses on how to support patients with limited (digital) health literacy could be beneficial. Such training may help nurses apply structured strategies rather than relying solely on intuition. When developing this training, it is important to cocreate its content with nurses from diverse clinical settings and with varying levels of experience, ensuring that the modules align with current practice. In addition, integrating these competencies into nursing education is essential so that future nurses also develop the necessary skills.

Moreover, involving nurses in the selection and implementation of health technologies is important. Their perspectives can guide the choice and adoption of new technologies, the design of training programs, and the development of user-friendly systems, such as EMRs. This approach helps ensure that digital health innovations are accessible, practical, and responsive to the needs of both nurses and patients. Ultimately, the impact of these technologies depends not only on the tools themselves but also on organizational and systemic factors, including staff readiness, training support, and supportive organizational structures [44-46].

Conclusions

This study shows that health technologies have strengthened nurse-patient communication and extended de Haes and Bensing's model [2] to the digital context. Yet, challenges persist, particularly in fostering relationships. It also reveals that systematic (digital) health literacy assessment remains lacking and that current practices to identify and support patients with limited (digital) health literacy may not always be effective. Once effective approaches are established, they should be integrated into nursing care plans and education.

Acknowledgments

We thank Alex Boshuizen and all students for their assistance in conducting this research and all nurses and nursing students for participating in this study.

Funding

This work was supported by ZonMw, the Netherlands Organisation for Health Research and Development (grant 10040022010001).

Data Availability

The datasets generated or analyzed during this study are not publicly available.

Authors' Contributions

Conceptualization: CHCD, CD, DE, EMD, JW
Data curation: ED, JW
Formal analysis: CHCD, ED, JW
Investigation: ED, JW
Methodology: CHCD, CD, ED, JW
Project administration: ED
Supervision: CD
Writing – original draft: CHCD, CD, DE, ED, JW

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Fleischer S, Berg A, Zimmermann M, Wüste K, Behrens J. Nurse-patient interaction and communication: a systematic literature review. J Public Health 2009 Oct;17(5):339-353. [doi: 10.1007/s10389-008-0238-1]
2. de Haes H, Bensing J. Endpoints in medical communication research, proposing a framework of functions and outcomes. Patient Educ Couns 2009 Mar;74(3):287-294. [doi: 10.1016/j.pec.2008.12.006] [Medline: 19150197]
3. Zolnierek KBH, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. Med Care 2009 Aug;47(8):826-834. [doi: 10.1097/MLR.0b013e31819a5acc] [Medline: 19584762]
4. Sharkiya SH. Quality communication can improve patient-centred health outcomes among older patients: a rapid review. BMC Health Serv Res 2023 Aug 22;23(1):886. [doi: 10.1186/s12913-023-09869-8] [Medline: 37608376]
5. Aapro M, Bossi P, Dasari A, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020 Oct;28(10):4589-4612. [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
6. Murugesu L, Heijmans M, Rademakers J, Fransen MP. Challenges and solutions in communication with patients with low health literacy: perspectives of healthcare providers. PLoS ONE 2022;17(5):e0267782. [doi: 10.1371/journal.pone.0267782] [Medline: 35507632]
7. Kayser L, Kushniruk A, Osborne RH, Norgaard O, Turner P. Enhancing the effectiveness of consumer-focused health information technology systems through eHealth literacy: a framework for understanding users' needs. JMIR Hum Factors 2015 May 20;2(1):e9. [doi: 10.2196/humanfactors.3696] [Medline: 27025228]
8. Milanti A, Norman C, Chan DNS, So WKW, Skinner H. eHealth Literacy 3.0: updating the norman and skinner 2006 model. J Med Internet Res 2025 Mar 11;27:e70112. [doi: 10.2196/70112] [Medline: 40068166]
9. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: a systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124. [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
10. Qiu CS, Lunova T, Greenfield G, et al. Determinants of digital health literacy: international cross-sectional study. J Med Internet Res 2025 Jun 30;27:e66631. [doi: 10.2196/66631] [Medline: 40587845]
11. Zaghloul H, Fanous K, Ahmed L, et al. Digital health literacy in patients with common chronic diseases: systematic review and meta-analysis. J Med Internet Res 2025 Aug 25;27:e56231. [doi: 10.2196/56231] [Medline: 40854231]
12. Heijmans M, Cariot L, Brabers AEM, Rademakers J. Infographic. Eén op de drie nederlanders heeft onvoldoende of beperkte gezondheidsvaardigheden - feiten en cijfers 2023 [Report in Dutch]. : Nivel; 2024 URL: https://www.nivel.nl/sites/default/files/bestanden/1004587.pdf [accessed 2026-03-05]
13. Nederlanders digitaal steeds vaardiger. Centraal Bureau voor de Statistiek. URL: https://www.cbs.nl/nl-nl/nieuws/2023/45/nederlanders-digitaal-steeds-vaardiger [accessed 2026-03-05]
14. Loan LA, Parnell TA, Stichler JF, et al. Call for action: nurses must play a critical role to enhance health literacy. Nurs Outlook 2018;66(1):97-100. [doi: 10.1016/j.outlook.2017.11.003] [Medline: 29331444]
15. Dickens C, Lambert BL, Cromwell T, Piano MR. Nurse overestimation of patients' health literacy. J Health Commun 2013;18 Suppl 1(Suppl 1):62-69. [doi: 10.1080/10810730.2013.825670] [Medline: 24093346]
16. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
17. Medical devices. World Health Organization. URL: https://www.who.int/teams/health-product-policy-and-standards/assistive-and-medical-technology/medical-devices [accessed 2026-03-05]
18. Flanagan JC. The critical incident technique. Psychol Bull 1954 Jul;51(4):327-358. [doi: 10.1037/h0061470] [Medline: 13177800]
19. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
20. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? Qual Health Res 2017 Mar;27(4):591-608. [doi: 10.1177/1049732316665344] [Medline: 27670770]
21. Talevski J, Wong Shee A, Rasmussen B, Kemp G, Beauchamp A. Teach-back: a systematic review of implementation and impacts. PLoS ONE 2020;15(4):e0231350. [doi: 10.1371/journal.pone.0231350] [Medline: 32287296]
22. Navarro Martínez O, Igual García J, Salcedo VT. Nurses' view of benefits, enablers and constraints to the use of digital health tools with patients: a cross-sectional study. Digit Health 2023;9:20552076231197339. [doi: 10.1177/20552076231197339] [Medline: 37675060]
23. Frennert S, Petersson L, Erlingsdottir G. "More" work for nurses: the ironies of eHealth. BMC Health Serv Res 2023 Apr 27;23(1):411. [doi: 10.1186/s12913-023-09418-3] [Medline: 37106404]

XSL•FO
RenderX

24. Öberg U, Orre CJ, Isaksson U, Schimmer R, Larsson H, Hörnsten Å. Swedish primary healthcare nurses' perceptions of using digital eHealth services in support of patient self-management. Scand J Caring Sci 2018 Jun;32(2):961-970. [doi: 10.1111/scs.12534] [Medline: 28960451]
25. Lie SS, Karlsen B, Graue M, Oftedal B. The influence of an eHealth intervention for adults with type 2 diabetes on the patient-nurse relationship: a qualitative study. Scand J Caring Sci 2019 Sep;33(3):741-749. [doi: 10.1111/scs.12671] [Medline: 30866066]
26. Ricci L, Villegente J, Loyal D, Ayav C, Kivits J, Rat AC. Tailored patient therapeutic educational interventions: a patient-centred communication model. Health Expect 2022 Feb;25(1):276-289. [doi: 10.1111/hex.13377] [Medline: 34816546]
27. Mackert M, Mabry-Flynn A, Champlin S, Donovan EE, Pounders K. Health literacy and health information technology adoption: the potential for a new digital divide. J Med Internet Res 2016 Oct 4;18(10):e264. [doi: 10.2196/jmir.6349] [Medline: 27702738]
28. Hogan A, Hughes L, Coyne E. Understanding nursing assessment of health literacy in a hospital context: a qualitative study. J Clin Nurs 2023 Oct;32(19-20):7495-7508. [doi: 10.1111/jocn.16809] [Medline: 37340615]
29. Brooks C, Ballinger C, Nutbeam D, Mander C, Adams J. Nursing and allied health professionals' views about using health literacy screening tools and a universal precautions approach to communication with older adults: a qualitative study. Disabil Rehabil 2020 Jun;42(13):1819-1825. [doi: 10.1080/09638288.2018.1538392] [Medline: 30669896]
30. Hogan A, Hughes L, Coyne E. Nurses' assessment of health literacy requirements for adult inpatients: an integrative review. Health Promot J Austr 2024 Apr;35(2):504-517. [doi: 10.1002/hpja.780] [Medline: 37443427]
31. Milanti A, Chan DNS, Parut AA, So WKW. Determinants and outcomes of eHealth literacy in healthy adults: a systematic review. PLoS ONE 2023;18(10):e0291229. [doi: 10.1371/journal.pone.0291229] [Medline: 37792773]
32. Wilandika A, Pandin MGR, Yusuf A. The roles of nurses in supporting health literacy: a scoping review. Front Public Health 2023;11:1022803. [doi: 10.3389/fpubh.2023.1022803] [Medline: 37663836]
33. Parikh NS, Parker RM, Nurss JR, Baker DW, Williams MV. Shame and health literacy: the unspoken connection. Patient Educ Couns 1996 Jan;27(1):33-39. [doi: 10.1016/0738-3991(95)00787-3] [Medline: 8788747]
34. Dijkman EM, Ter Brake WWM, Drossaert CHC, Doggen CJM. Assessment tools for measuring health literacy and digital health literacy in a hospital setting: a scoping review. Healthcare (Basel) 2023 Dec 20;12(1):11. [doi: 10.3390/healthcare12010011] [Medline: 38200917]
35. Bader M, Zheng L, Rao D, et al. Towards a more patient-centered clinical trial process: a systematic review of interventions incorporating health literacy best practices. Contemp Clin Trials 2022 May;116:106733. [doi: 10.1016/j.cct.2022.106733] [Medline: 35301134]
36. Arias López MDP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. PLOS Digit Health 2023 Oct;2(10):e0000279. [doi: 10.1371/journal.pdig.0000279] [Medline: 37824584]
37. Yang HF, Chang CC, Tseng PL, et al. Effectiveness of innovative instructional module for professional competence in health literacy in medical students. BMC Med Educ 2022 Mar 28;22(1):210. [doi: 10.1186/s12909-022-03252-7] [Medline: 35351115]
38. van der Giessen J, Fransen MP, Spreeuwenberg P, Velthuizen M, van Dulmen S, Ausems M. Communication about breast cancer genetic counseling with patients with limited health literacy or a migrant background: evaluation of a training program for healthcare professionals. J Community Genet 2021 Jan;12(1):91-99. [doi: 10.1007/s12687-020-00497-x] [Medline: 33319336]
39. Hsieh JG, Yu JH, Wang YW, et al. Health literacy training program for community healthcare providers using hybrid online team-based learning in Taiwan. BMC Med Educ 2022 Jul 27;22(1):576. [doi: 10.1186/s12909-022-03646-7] [Medline: 35897091]
40. Gibson C, Smith D, Morrison AK. Improving health literacy knowledge, behaviors, and confidence with interactive training. Health Lit Res Pract 2022 Apr;6(2):e113-e120. [doi: 10.3928/24748307-20220420-01] [Medline: 35522855]
41. Mackert M, Ball J, Lopez N. Health literacy awareness training for healthcare workers: improving knowledge and intentions to use clear communication techniques. Patient Educ Couns 2011 Dec;85(3):e225-e228. [doi: 10.1016/j.pec.2011.02.022] [Medline: 21474264]
42. Allenbaugh J, Spagnoletti CL, Rack L, Rubio D, Corbelli J. Health literacy and clear bedside communication: a curricular intervention for internal medicine physicians and medicine nurses. MedEdPORTAL 2019 Jan 18;15:10795. [doi: 10.15766/mep_2374-8265.10795] [Medline: 30800995]
43. Kaper MS, Reijneveld SA, van Es FD, et al. Effectiveness of a comprehensive health literacy consultation skills training for undergraduate medical students: a randomized controlled trial. Int J Environ Res Public Health 2019 Dec 20;17(1):81. [doi: 10.3390/ijerph17010081] [Medline: 31861918]
44. Walzer S, Armbruster C, Mahler S, Farin-Glattacker E, Kunze C. Factors influencing the implementation and adoption of digital nursing technologies: systematic umbrella review. J Med Internet Res 2025 Jul 31;27:e64616. [doi: 10.2196/64616] [Medline: 40743516]

XSL•FO
RenderX

45. Jeilani A, Hussein A. Impact of digital health technologies adoption on healthcare workers' performance and workload: perspective with DOI and TOE models. BMC Health Serv Res 2025 Feb 18;25(1):271. [doi: 10.1186/s12913-025-12414-4] [Medline: 39966874]
46. Conte G, Arrigoni C, Magon A, Stievano A, Caruso R. Embracing digital and technological solutions in nursing: a scoping review and conceptual framework. Int J Med Inform 2023 Sep;177:105148. [doi: 10.1016/j.ijmedinf.2023.105148] [Medline: 37453178]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
EMR: electronic medical record

Edited by E Borycki; submitted 10.Oct.2025; peer-reviewed by F Fischer, K Connell, VP Jover; revised version received 19.Dec.2025; accepted 23.Dec.2025; published 26.Mar.2026.

Please cite as:
Dijkman EM, Wentzel J, Edvardsson D, Doggen C, Drossaert CHC
Changes in Nurse-Patient Communication Through Health Technologies and Nursing Practices to Recognize and Support Limited Digital Health Literacy: Qualitative Study
JMIR Nursing 2026;9:e82272
URL: https://nursing.jmir.org/2026/1/e82272
doi:10.2196/82272

© Eline Mariosé Dijkman, Jobke Wentzel, David Edvardsson, Carine Doggen, Constance Helène Christine Drossaert. Originally published in JMIR Nursing (https://nursing.jmir.org), 26.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Facilitating Teamwork With Minimal Interaction via a Mesh-Based Communication Device in a Clinical Nursing Setting: Intervention Study

Yujiro Matsuishi[1], RN, PhD; Masakazu Hirokawa[2], PhD; Yoshiaki Inoue[3], MD, PhD; Kenji Suzuki[2], PhD

[1]Science and Engineering, Women's Health Nursing and Midwifery, Institute of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan

[2]Artificial Intelligence Laboratory, University of Tsukuba, Tsukuba, Ibaraki, Japan

[3]Department of Emergency and Critical Care Medicine, Institute of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan

Corresponding Author:
Yujiro Matsuishi, RN, PhD
Science and Engineering, Women's Health Nursing and Midwifery, Institute of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan

Abstract

Background: The prevention of medical errors depends on effective communication and teamwork between nurses, yet traditional nurse call systems create inefficiencies that result in higher staff workloads. Further research is needed to evaluate how digital health care technologies affect nursing workflows and whether they contribute to staff workload.

Objective: This study aimed to evaluate whether a smartphone app using Bluetooth Low Energy (BLE) Mesh–based signal communication could enhance nurse-to-nurse communication, decrease response times, and improve teamwork as measured by the TeamSTEPPS Teamwork Perceptions Questionnaire (T-TPQ).

Methods: We conducted a between-subjects (independent groups) interventional study on an acute care ward, with 3 intervention days (BLE Mesh app use) and 3 nonintervention days administered on alternating days. Thirty nurses (15 per condition) participated, with no nurse contributing observations to both conditions in the final analysis. Response times were tracked, and the T-TPQ was used to evaluate Situation Monitoring and Mutual Support perceptions at the end of each shift.

Results: The app shortened the total response duration from a per-nurse median of 38.5 (IQR 31.0-58.5) seconds in the control condition to 19.5 (IQR 18.5-23.5) seconds in the intervention condition (n=15 per condition; Mann-Whitney *U* test=217.5; *P*<.001; rank-biserial *r*=–0.93). The app reduced the time it took for staff to confirm nursing actions, search for items, and request care assistance. A total of 40 nurses participated, generating 48 questionnaire responses; after exclusions, 30 nurses (15 per condition) were analyzed. T-TPQ confirmed significant improvements in Situation Monitoring (3.04 vs 3.39; *P*=.001) and Mutual Support (3.01 vs 3.31; *P*=.003). Post hoc power ranged from 87% to 92% (Cohen *d*>1.0), supporting the interpretability of these results.

Conclusions: BLE Mesh technology reduces basic signal communication delays, which shortens nurse response times and enhances team performance through improved Situation Monitoring and Mutual Support, according to our research. The study demonstrates how this technology delivers functional benefits that enhance nursing communication to produce superior team-based care outcomes.

(JMIR Nursing 2026;9:e85599) doi:10.2196/85599

KEYWORDS

nursing communication; teamwork; digital health; response time; mobile apps

Introduction

Effective communication and robust teamwork are foundational pillars of high-quality health care delivery, particularly within nursing practice. The intricate nature of patient care necessitates continuous, efficient, and accurate information exchange among health care professionals to ensure comprehensive and holistic patient outcomes [1]. This "professional communication" is paramount for facilitating information sharing, enhancing the appropriateness of medical interventions, and ultimately improving patient safety [2]. Conversely, communication failures represent a pervasive and critical issue in modern health care, frequently leading to medical errors, adverse events, and even preventable patient mortality [1-3]. Indeed, ineffective communication often creates obstacles to information flow among health care teams, a lack of teamwork, organizational inefficiencies, and decreased staff morale. Notably, poor communication has been identified as the sole cause in 13.2% of patient safety incidents and a contributing factor in 24.0% of incidents [4]. Factors such as hierarchical structures, cultural

XSL•FO
RenderX

differences, or even disruptive behaviors in the health care environment can hinder effective communication, highlighting the need for improvements through organizational commitment, training, and standardized policies [2]. This underscores the urgent necessity for innovative solutions to enhance communication and collaboration within health care teams.

Despite the recognized importance of effective communication, nurses frequently encounter significant challenges that impede seamless information transfer and mutual support. For example, traditional nurse call systems are predominantly static and "place-oriented," confining patients to fixed locations for initiating calls and lacking crucial contextual information regarding patient risk factors or staff availability [2,5]. This leads to considerable inefficiencies, including unnecessary interruptions for nurses, wasted time spent locating staff (≈10% of nursing time is spent "looking for someone"), and inappropriate call routing, where highly skilled registered nurses are often engaged in tasks outside their professional scope [2,5]. Such situations can increase nurses' workload and potentially diminish the quality of care.

Furthermore, while digital technologies are increasingly integrated into nursing practice with the aim of improving efficiency and communication, their impact is not always positive. The adoption of information and communication technologies, such as smartphones and wearable devices, indicates a tendency for nurses to use these tools for information retrieval and communication with health care team members. However, their overall impact on nurses' perceived workload has been complex, often proving "mainly negative" [6-8]. This suggests that simply introducing technology, without careful consideration of workflow integration and cognitive load, can inadvertently exacerbate existing burdens [5]. Barriers to the adoption of digital health technologies include infrastructural and technical issues, psychological and personal concerns, and worries about increased working hours and workload. Conversely, training, perceived usefulness, government incentives, and user involvement in development have been identified as facilitators for adoption [5].

Bluetooth Low Energy (BLE) Mesh is a communication solution characterized by its flexibility, reliability, and low power consumption. It offers the advantage of enhanced network coverage and internetwork connectivity, which is especially beneficial in complex hospital environments [8]. While BLE has seen applications in the medical field, such as for contact tracing and patient monitoring systems [8,9], research specifically focused on leveraging this technology for dynamic, peer-to-peer communication and mutual support among nurses, with an emphasis on concretely reducing workload and improving communication efficiency, largely remains unexplored. This app is characterized by simple signals conveying mutual support among nurses. Therefore, to address these significant gaps and challenges, this study developed and validated a novel, simple signal communication app using BLE Mesh network technology.

Methods

Study Design

This study used a between-subjects (independent groups) interventional design, comparing nurse response times (RTs) and teamwork perceptions between an intervention condition (using the BLE Mesh app) and a nonintervention condition (without the app), administered on alternating days. Different nurses participated in each condition; in the final analysis, no individual nurse contributed observations to both groups.

Intervention Design

This research comprises 2 distinct phases: intervention and nonintervention days administered on alternating days, beginning with a nonintervention day. The study included 3 app-use days alongside 3 app-free days for a complete observation period of 6 days.

Participants

Each shift on the ward was staffed by approximately 8 room-based nurses, including the charge nurse, who shared primary responsibility for direct patient care. Because nurses were assigned to shifts according to the standard ward roster rather than being fixed across the study period, a total of 40 unique nurses participated across the 6 study days, generating 24 questionnaire responses per condition (48 responses in total). Different nurses contributed to the intervention and nonintervention conditions, and those who happened to provide responses under both conditions were excluded from the final analysis (see "Demographic Data" section).

Experimental App

This study developed a prototype smartphone app as research support for social signal communication between nurses. Through its core technology, the app builds a mesh network using BLE to allow real-time communication between multiple devices. Users can access the app through smartphones running Android OS (Android 11 or later) with dynamic mesh network configurations. The study used 8 Google Pixel 4a devices running Android 13. The app functions through 2 different call systems that distinguish between critical and nonurgent assistance needs: green call and red call. In green call, the function is intended for nonemergency help requests. The app communicates both call types and sender names to the other nurses' terminals when a user initiates a call. The call is cleared when any recipient presses the designated response button; a confirmation notice is then displayed for 20 seconds on the sender's terminal (10 s on the other terminals). In red call, the function serves as an urgent help request system. The notification persists until 2 responders answer the call or the initiator manually ends the alert.

The app was developed as a research prototype by the study team. The Android app was built using Android Studio (Google LLC) and implements a custom application-layer BLE relay protocol, enabling multihop, infrastructure-free, device-to-device communication. Message delivery was subjectively immediate in prestudy testing, although formal latency measurements were

XSL•FO
RenderX

not performed. The prototype underwent iterative usability testing by the research team prior to deployment.

Experimental Procedure and Data Collection

The nurses who participated in the study received Android devices with the app they needed to complete for this research. The data collection process included both days with app use and days without app use according to the following framework:

- Preliminary explanation and informed consent: the study's objective, along with its details, privacy measures, and ethical principles, was presented verbally and in writing to obtain consent.
- Observation during work shift (6 h): participants' actions were observed during their 6-hour work shift.
- Intervention day: the developed app was used.
- Nonintervention day: the app was not used; participants only carried the device.
- Postshift evaluation: after the work shift, participants answered a questionnaire regarding teamwork perception. The observation of cooperative actions was conducted by observers with nursing qualifications.

Evaluation Items

This research investigates the clinical impact of app use by evaluating the time taken for responses to occur as its main outcome.

Method of Measuring RT

The RT measurement procedure operates identically for both intervention days and nonintervention days. The observer starts timing when the nurse expresses their intention to contact another staff member. The observed period ends when the nurse initiates a work-related dialog with another person, which marks the end of the RT interval.

Measurement Scales

The Agency for Healthcare Research and Quality created TeamSTEPPS Teamwork Perceptions Questionnaire (T-TPQ) [10] as a research-based framework for enhancing team performance to improve health care quality and safety [11], and T-TPQ has already been validated in many languages, including Japanese [12-14]. The T-TPQ consists of 35 self-report items that assess individual perceptions about the 5 core competencies: Team Structure, Leadership, Situation Monitoring, Mutual Support, and Communication [15]. This study focused on the Situation Monitoring and Mutual Support competency scales from the T-TPQ because these 2 dimensions are most directly relevant to the intervention mechanism. Situation Monitoring captures nurses' awareness of their team members' statuses and the surrounding environment, which is directly facilitated by the app's real-time broadcast notifications. Mutual Support captures assistance-seeking and assistance-offering behaviors, which are directly targeted by the app's Green Call and Red Call functions. The remaining 3 T-TPQ competencies (Team Structure, Leadership, and Communication) were considered less directly addressed by this simple signal-based intervention and were therefore excluded from the primary analysis.

Data Analysis

All statistical analyses were performed using IBM SPSS Statistics and Python (*statsmodels*; Python Software Foundation).

Primary Analysis (RT)

The prespecified primary analysis compared per-nurse median RTs between the 15 control and 15 intervention nurses using a 2-sided Mann-Whitney *U* test (exact test). Each nurse's RT events were first aggregated into a single per-nurse median, yielding the 15-vs-15 comparison and bringing the RT analysis into structural symmetry with the T-TPQ analysis (which is inherently per-nurse).

Sensitivity Analyses (RT)

Two prespecified sensitivity analyses on the unaggregated event-level data were performed using linear mixed effects models on log-transformed RT, fitted by restricted maximum likelihood: (model A) log(RT)~Condition+(1 | nurse; Table S1 in Multimedia Appendix 1) and (model B) log(RT)~Condition+(1 | day; Table S2 in Multimedia Appendix 1).

Exploratory Analysis (RT by Request Category)

Event-level Mann-Whitney *U* tests were applied within each of the 6 request categories. Per-nurse aggregation within each category was not used because each nurse contributed only 0 to 3 events per category. To control the family-wise error rate across the 6 categories, a Bonferroni correction was applied (adjusted α=.05/6≈.0083).

T-TPQ Analysis

The Mann-Whitney *U* test (2-sided, with normal approximation and continuity correction) was applied at the per-nurse level (n=15 per condition) to the total subscale scores for Situation Monitoring and Mutual Support and the 7 individual item scores within each subscale. Subscale-total comparisons were treated as single confirmatory tests and were, therefore, not Bonferroni-adjusted. Item-level tests within each subscale were Bonferroni-adjusted across 7 items (α=.05/7≈.0071).

Effect Sizes and Power

Effect sizes were quantified using Cohen *d* on the log scale and rank-biserial *r*. Post hoc power was estimated at α=.05.

Sample Size Calculation

To our knowledge, no prior study has evaluated the effect of BLE Mesh–based real-time communication systems on nurse RT or teamwork perceptions. The absence of equivalent preceding research meant that no empirical basis existed for estimating the expected effect size, and accordingly, a formal a priori sample size calculation was not conducted. The study was designed as a feasibility investigation to generate preliminary effect size estimates that can inform the sample size planning of future adequately powered trials. Post hoc power analysis was, therefore, conducted to retrospectively assess whether the achieved sample provided adequate power to detect the observed effects.

XSL•FO
RenderX

Ethical Considerations

The study was approved by the University of Tsukuba Institutional Review Board (approval 2024R885) and was conducted in accordance with principles equivalent to those of the Declaration of Helsinki. Consent for participation and data use was obtained through an opt-out procedure approved by the institutional review board. A written notice describing the study objectives, the types of data to be recorded, the data protection procedures, and the procedure for declining participation was posted in the ward in advance, and nurses who did not opt out during the notification period were considered to have consented to participation. In addition, prior to each observation session, the research team provided a careful verbal explanation of the study objectives, the data to be recorded, and each participant's right to decline or withdraw at any time, giving every on-duty nurse repeated opportunities to ask questions or withdraw before the session began. To protect staff privacy, nurse identifiers in the behavioral logs and questionnaire records were replaced with pseudonymous participant IDs at the time of data entry; the linkage table was stored separately on an access-controlled institutional server and was not exported with the analysis files. Communication logs recorded only (1) the time interval between request initiation and response, (2) the request location at the room level (not patient identifier), and (3) the request category; no patient identifiers, patient clinical data, or patient outcomes were recorded. Questionnaire responses were collected in a manner that did not link individual respondents to specific patients. Access to raw logs and questionnaire data was restricted to the research team.

Results

Demographic Data

A total of 48 questionnaire responses were collected from 40 unique nurses: 24 responses under the nonintervention condition and 24 under the intervention condition. Because some nurses responded more than once, the following screening steps were applied to ensure unique samples.

First, 5 nurses who participated in both the intervention and nonintervention conditions were excluded to avoid within-subjects contamination, leaving 35 nurses who participated in only 1 condition (18 in the nonintervention group and 17 in the intervention group).

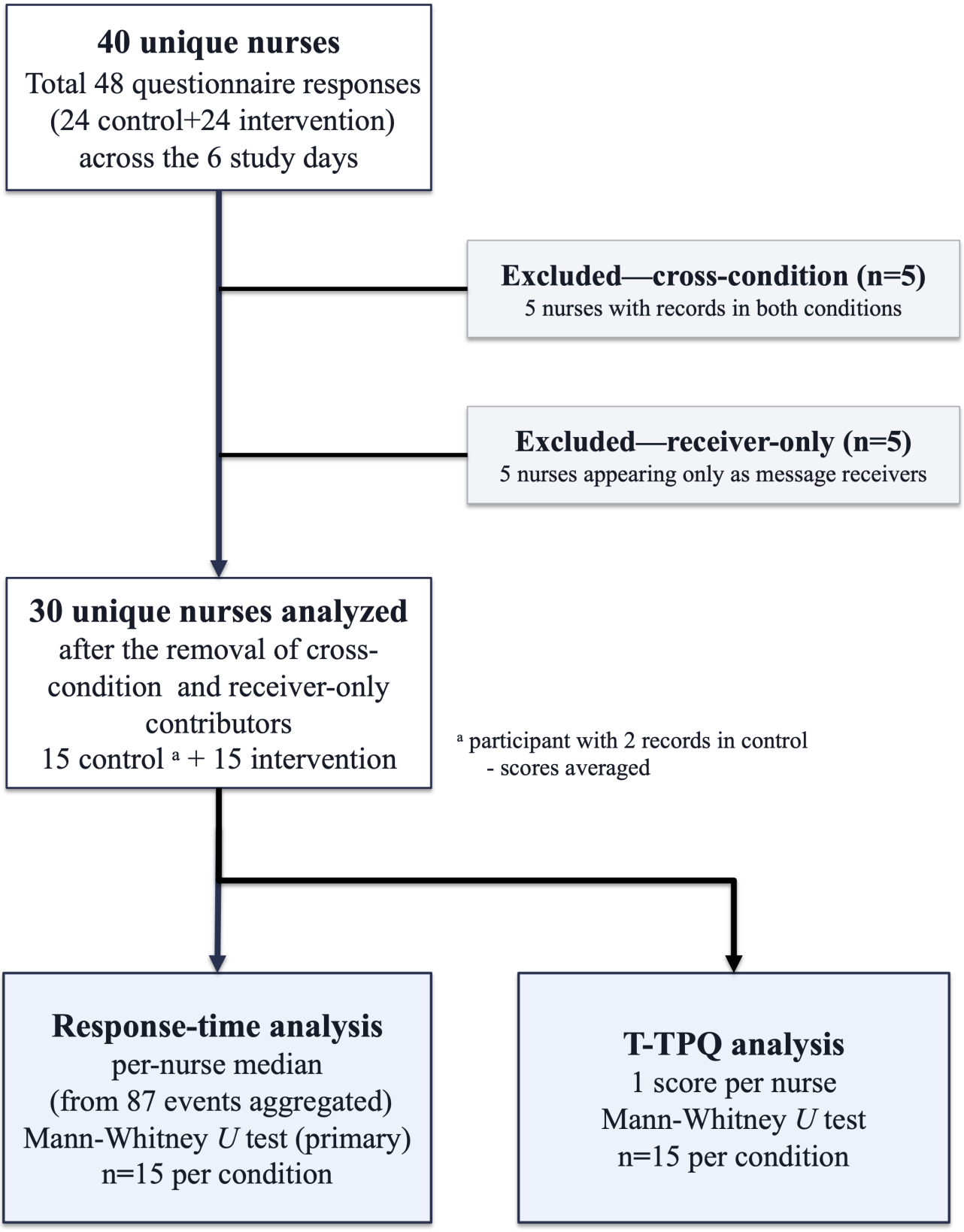
Second, 5 nurses (3 from the nonintervention group and 2 from the intervention group) were excluded because they appeared only as message receivers in the behavioral log and contributed no sender-initiated data, making it impossible to link their questionnaire responses to observed cooperative behaviors.

Finally, 1 nurse, who submitted 2 responses within the nonintervention condition, had the scores averaged to produce a single record.

These steps yielded 15 unique participants per condition (30 total; Figure 1).

XSL•FO
RenderX

Figure 1. Participant flow diagram. Diagram tracing how 48 questionnaire responses from 40 unique nurses across 6 days of study were reduced to 30 unique participants (15 per condition) for analysis. Exclusions: 5 nurses who provided responses under both conditions, 5 nurses who appeared only as message receivers (with no sender-initiated events), and 1 nurse whose 2 within-condition responses were averaged into a single record. T-TPQ: TeamSTEPPS Teamwork Perceptions Questionnaire.



XSL•FO
RenderX

Observed Requests

These 30 participants collectively initiated a total of 87 requests (38 under the control condition and 49 under the intervention condition) across the 6 days of study.

The nonintervention group (control group) made 38 requests, while the intervention group submitted 49 requests. As shown in Table 1, the locations of the requests in the control group (n=38) comprised 13 (34%) requests from 4-bed room, 14 (37%) requests from corridor, 5 (13%) requests from nurse station, 3 (8%) requests from private room, and 3 (8%) requests from day space.

The intervention group (n=49) obtained 22 (45%) requests from the 4-bed room, 12 (24%) requests from corridor, 14 (29%) requests from nurse station, 1 (2%) request from private room, and 0 (0%) requests from day space.

The control group (n=38) included 11 (29%) requests for information sharing, 9 (24%) requests for confirmation to nurse, 8 (21%) requests for care help, 4 (11%) requests for consultation, 4 (11%) requests for inquiry for items, and 2 (5%) requests for confirmation to leader. On the other hand, in the intervention group (n=49), there were 12 (24%) requests for information sharing, 12 (24%) requests for confirmation to nurse, 8 (16%) requests for care help, 11 (22%) requests for consultation, 4 (8%) requests for inquiry for items, and 2 (4%) requests for confirmation to leader.

All observed nurse-initiated requests were classified by the on-site observer into 1 of 6 content categories at the time of each interaction. Care help refers to requests for direct physical assistance with patient care tasks such as repositioning, lifting, or procedural support. Consultation refers to requests for clinical advice or peer consultation regarding patient management decisions. Information sharing refers to the proactive communication of patient-related updates or clinically relevant information to colleagues. Confirmation to nurse refers to requests for verification or acknowledgment from a peer nurse regarding an observation, a procedure, or a care plan. Inquiry for items refers to requests for assistance in locating medical equipment or supplies. Confirmation to leader refers to requests directed to the charge nurse for guidance, approval, or supervisory input.

Table . Number of requests by category per study condition.

Variable	Control (n=38)	Intervention (n=49)
Count of request, n (%)		
Day 1 (control)	8 (21)	—[a]
Day 2 (intervention)	—	21 (43)
Day 3 (control)	18 (47)	—
Day 4 (intervention)	—	12 (24)
Day 5 (control)	12 (32)	—
Day 6 (intervention)	—	16 (33)
Location of request, n (%)		
4-bed room	13 (34)	22 (45)
Corridor	14 (37)	12 (24)
Nurse station	5 (13)	14 (29)
Private room	3 (8)	1 (2)
Day space	3 (8)	0 (0)
Category of request, n (%)		
Information sharing	11 (29)	12 (24)
Confirmation to nurse	9 (24)	12 (24)
Care help	8 (21)	8 (16)
Consultation	4 (11)	11 (22)
Inquiry for items	4 (11)	4 (8)
Confirmation to leader	2 (5)	2 (4)

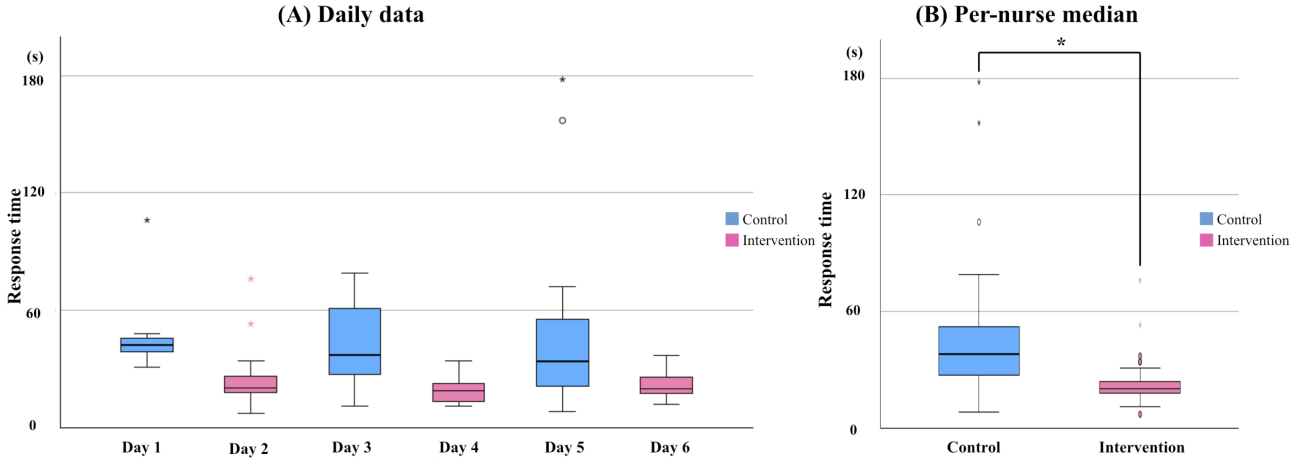
[a]Not applicable.

Primary Analysis: Per-Nurse RT

As the primary analysis, RT observations from the 3-day intervention period and the 3-day nonintervention period were first aggregated to a per-nurse median to account for the within-nurse clustering of repeated observations. The resulting 15 control values and 15 intervention values were then compared using a 2-sided Mann-Whitney *U* test. Daily medians of the unaggregated event-level data are shown in Figure 2 for reference.

XSL•FO
RenderX

Figure 2. Impact of app use on nurse response times: daily intervention and control data. (A) Event-level response times for each of the 6 study days (boxplots: median, IQR, whiskers) and the corresponding daily medians shown in the accompanying table. (B) Per-nurse median response times for the 15 control and 15 intervention nurses, which constitute the primary analysis; the intervention condition significantly reduced response time (Mann-Whitney *U* test=217.5; *P*<.001; rank-biserial *r*=−0.93, large effect). The asterisk symbol (*) in panel (B) denotes a statistically significant difference between the control and intervention conditions (Mann-Whitney *U* test; *P*<.001).



	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
Time (s), median (IQR)	42 (40-44)	20 (18-26)	37 (28-59)	19 (14-22)	34 (21-47)	20 (18-25)

	Control	Intervention	*P* value
Time (s), median (IQR)	38.5 (31-58.5)	19.5 (18.5-23.5)	<.001

○, outlier (1.5–3 box lengths); plot-area *, extreme outlier (>3 box lengths). Symbol colors indicate study condition only.

The median (IQR) of daily RTs is shown in Figure 2 as follows: day 1 (control) was 42 (IQR 40-44) seconds, day 2 (intervention) was 20 (IQR 18-26) seconds, day 3 (control) was 37 (IQR 28-59) seconds, day 4 (intervention) was 19 (IQR 14-22) seconds, day 5 (control) was 34 (IQR 21-47) seconds, and day 6 (intervention) was 20 (IQR 18-25) seconds.

At the per-nurse level, the median RT was 38.5 (IQR 31.0-58.5) seconds in the control condition and 19.5 (IQR 18.5-23.5) seconds in the intervention condition. The difference was statistically significant (Mann-Whitney *U* test=217.5; *P*<.001; rank-biserial *r*=–0.93, large effect).

Sensitivity Analysis: Linear Mixed Effects Model

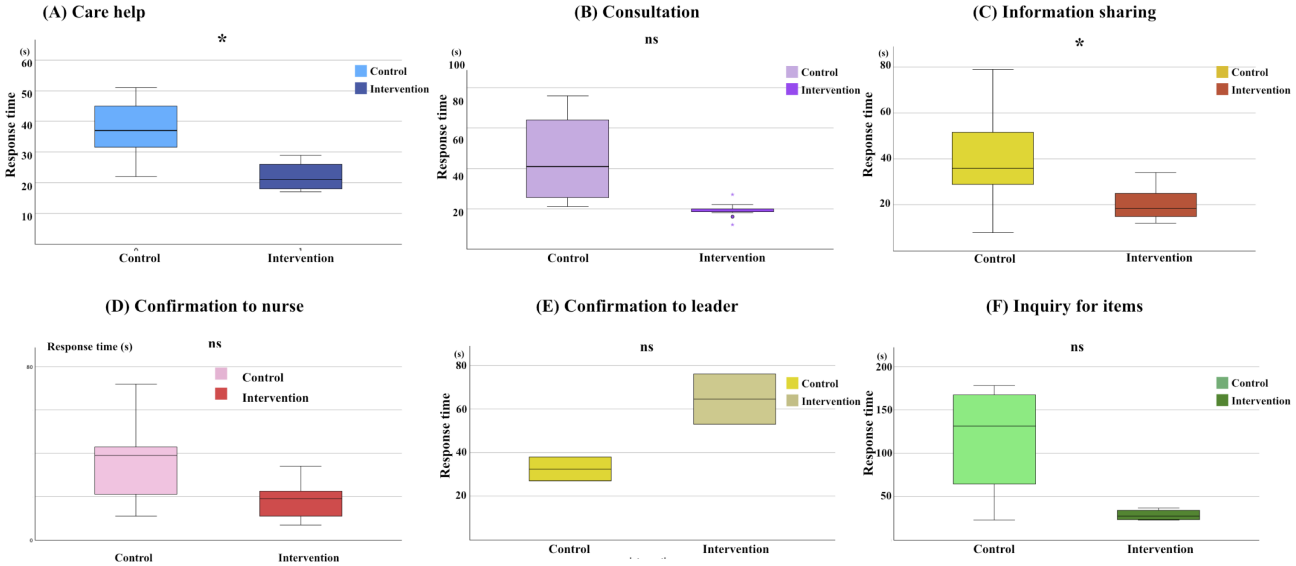
To formally account for the within-nurse clustering of repeated observations, a prespecified sensitivity analysis fitted a linear mixed effects model to all 87 events with log-transformed RT as the outcome, condition as a fixed effect, and a random intercept for nurse (Table S1 in Multimedia Appendix 1). The intervention reduced log-transformed RT by β of −0.67 (95% CI −0.93 to −0.41; *P*<.001), corresponding to a 48.7% reduction in geometric mean RT. The within-nurse intraclass correlation of 0.14 was modest, indicating that the per-nurse aggregation used in the primary analysis provides an interpretable summary, while the mixed effects model formally confirms the result under the residual within-nurse correlation.

Exploratory Analysis: RT by Request Category

As an exploratory analysis, event-level RTs were compared between conditions within each of the 6 request categories using 2-sided Mann-Whitney *U* tests (Figure 3). To control the family-wise error rate across the 6 categories, Bonferroni correction was applied (adjusted α=.05/6≈.0083). Two categories remained significant after correction: “Care help” decreased from 37 (control) to 21 seconds (intervention; *P*=.003) and “Information sharing” decreased from 36 to 18.5 seconds (*P*=.006). Two further categories showed nominal reductions of similar magnitude but did not survive Bonferroni correction: “Consultation” (41-20 s; *P*=.01) and “Confirmation to nurse” (39-19 s; *P*=.02). “Inquiry for items” showed a numerically large reduction from 131.5 to 27.5 seconds that did not reach statistical significance (*P*=.25). For “Confirmation to leader,” median values were 32.5 seconds (control; n=2) and 64.5 seconds (intervention; n=2); with only 2 observations per condition, this comparison was underpowered (*P*=.25) and should be interpreted with caution.

XSL•FO
RenderX

Figure 3. Impact of app use for category of request. Boxplots show the median response time (IQR values) for each of the 6 request categories under the control and intervention conditions, with per-category Mann-Whitney *U* test values and *P* values listed in the accompanying tables. Bonferroni correction was applied across the 6 categories (adjusted α=.05/6≈.0083). Care help (*P*=.003) and information sharing (*P*=.006) remained significant after correction; consultation (*P*=.01) and confirmation to nurse (*P*=.02) showed nominal reductions that did not survive correction; Inquiry for items (*P*=.25) and confirmation to leader (*P*=.25) did not differ between conditions. ns: not significant.



Care help			Consultation			Information sharing		
Intervention, median (IQR), s	Control, median (IQR), s	*P* value	Intervention, median (IQR), s	Control, median (IQR), s	*P* value	Intervention, median (IQR), s	Control, median (IQR), s	*P* value
21 (18–25)	37 (31.8–43.5)	.003	20 (18.5–20)	41 (27.8–58)	.01	18.5 (15–24.5)	36 (29–51.5)	.006

Confirmation to nurse			Confirmation to leader			Inquiry for items		
Intervention, median (IQR), s	Control, median (IQR), s	*P* value	Intervention values, s	Control values, s	*P* value	Intervention, median (IQR), s	Control, median (IQR), s	*P* value
19 (11–22.2)	39 (21–43)	.02	53, 76 (median 64.5)	27, 38 (median 32.5)	.25	27.5 (23.8–32.5)	131.5 (85.2–162.2)	.25

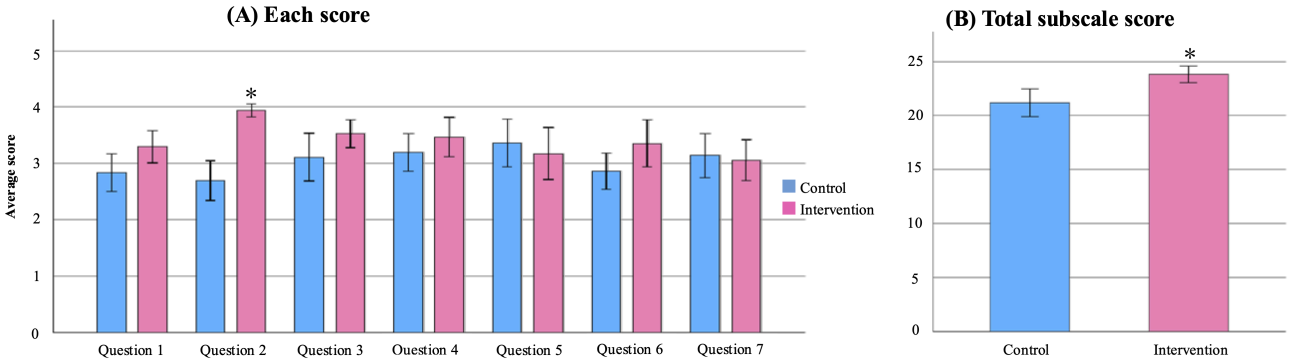
*** Significant after Bonferroni correction across the 6 per-category Mann-Whitney *U* tests (α=.05/6≈.0083).**

Effect on Perceived Situation Monitoring (T-TPQ)

T-TPQ results were analyzed using the Mann-Whitney *U* test on 30 unique participants (n=15 per condition) identified from sender activity logs. For Situation Monitoring, the Mann-Whitney *U* test was applied to check for significant differences between individual T-TPQ items, with a Bonferroni-adjusted α of .0071. As shown in Figure 4, 1 item reached statistical significance after correction: question 2, “Staff monitor each other’s performance” increased from 2.70 (control) to 3.93 (intervention); *P*<.001 (significant, large effect). The remaining 6 items did not reach significance: question 1, “Staff effectively anticipate each other’s needs” (control 2.87 vs intervention 3.27; *P*=.13); question 3, “Staff exchange relevant information as it becomes available” (3.20 vs 3.53; *P*=.44); question 4, “Staff continuously scan the environment for important information” (*P*=.29); question 5, “Staff share information regarding potential complications (eg, patient changes, bed availability)” (3.37 vs 3.20; *P*=.68); question 6, “Staff meet to reevaluate patient care goals when aspects of the situation have changed” (2.77 vs 3.33; *P*=.07); and question 7, “Staff correct each other’s mistakes to ensure that procedures are followed properly” (3.17 vs 3.00; *P*=.54). The Situation Monitoring total score improved from a mean of 3.04 (SD 0.39) in the control group to 3.39 (SD 0.19) in the intervention group (Figure 4), and the difference was statistically significant (*P*=.001; Cohen *d*=1.13, large effect).

XSL•FO
RenderX

Figure 4. Impact of the intervention on perceived Situation Monitoring (TeamSTEPPS Teamwork Perceptions Questionnaire domain). (A) Average score for each of the 7 Situation Monitoring items (Question 1-Question 7) under the control and intervention conditions (error bars represent the SE). (B) Total subscale score (sum across Question 1-Question 7). Item-level Mann-Whitney *U* tests were Bonferroni-adjusted across the 7 items (α=.05/7≈.0071); the subscale-total comparison was a single confirmatory test and was, therefore, not Bonferroni-adjusted. The total subscale score improved significantly under the intervention condition (3.04 to 3.39; *P*=.001; Cohen *d*=1.13, large effect). At the item level, Question 2 ("Staff monitor each other's performance") was the only item that survived Bonferroni correction (*P*<.001). Asterisks (*) denote items significant comparisons: in panel (A), the item that survived Bonferroni correction (question 2; *P*<.001; Bonferroni-adjusted α=.05/7≈.0071); in panel (B), the significant confirmatory subscale-total comparison (*P*=.001; not Bonferroni-adjusted, as prespecified).



Situation Monitoring	Intervention, mean (SD)	Control, mean (SD)	*P value*	Cohen *d*	Bonferroni-adjusted[a]
Total subscale score	23.73 (1.33)	21.30 (2.72)	.001	1.13	Significant
1. Staff effectively anticipate each other's needs	3.27 (0.59)	2.87 (0.74)	.13	0.59	Not significant
2. Staff monitor each other's performance	3.93 (0.26)	2.70 (0.80)	<.001	2.08	Significant
3. Staff exchange relevant information as it becomes available	3.53 (0.52)	3.20 (0.94)	.44	0.44	Not significant
4. Staff continuously scan the environment for important information	3.47 (0.74)	3.23 (0.68)	.29	0.33	Not significant
5. Staff share information regarding potential complications (eg, patient changes, bed availability)	3.20 (0.94)	3.37 (0.90)	.68	−0.18	Not significant
6. Staff meet to reevaluate patient care goals when aspects of the situation have changed	3.33 (0.90)	2.77 (0.68)	.07	0.71	Not significant
7. Staff correct each other's mistakes to ensure that procedures are followed properly	3.00 (0.76)	3.17 (0.84)	.54	−0.21	Not significant

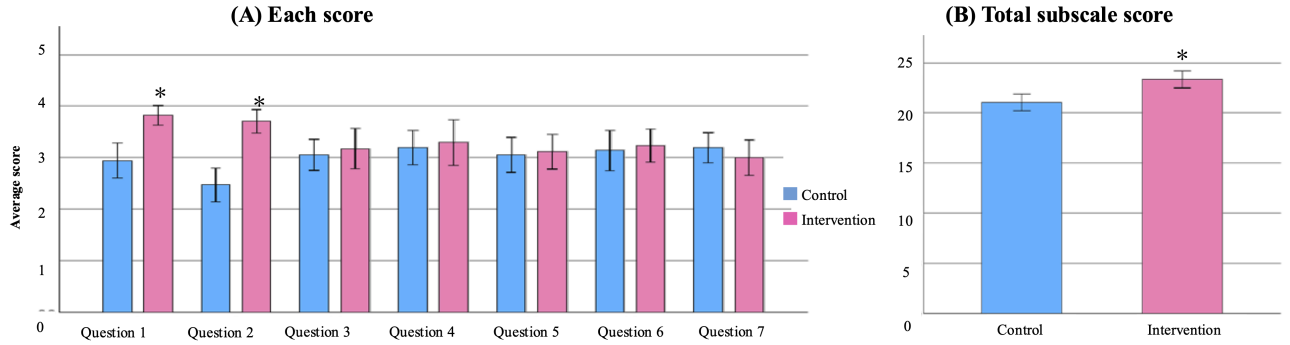
a Bonferroni-adjusted: α=.007 (.05/7); Cohen *d*: 0.2=small, 0.5=medium, and 0.8=large

Effect on Perceived Mutual Support (T-TPQ)

As shown in Figure 5, for Mutual Support, 2 items reached statistical significance after Bonferroni correction: question 1, "Staff assist fellow staff during high workload," increased from 3.00 (control) to 3.80 (intervention), *P*=.003 (significant); question 2, "Staff request assistance from fellow staff when they feel overwhelmed," increased from 2.43 to 3.67, *P*<.001 (significant, large effect). The remaining 5 items did not reach significance: question 3, "Staff caution each other about potentially dangerous situations" (3.07 vs 3.20; *P*=.55); question 4, "Feedback between staff is delivered in a way that promotes positive interactions and future change" (3.30 vs 3.20; *P*=.98); question 5, "Staff advocate for patients even when their opinions conflict with those of a senior member" (2.93 vs 3.13; *P*=.46); question 6, "When staff have a concern about patient safety, they challenge others until they are sure the concern has been heard" (3.10 vs 3.13; *P*=.98); and question 7, "Staff resolve their conflicts, even when the conflicts have become personal" (3.23 vs 3.07; *P*=.49). The Mutual Support total score improved from a mean of 3.01 (SD 0.24) in the control group to 3.31 (SD 0.25) in the intervention group, as shown in Figure 5, and the difference was highly significant (*P*=.003; Cohen *d*=1.24, large effect).

XSL•FO
RenderX

Figure 5. Impact of the intervention on perceived Mutual Support (TeamSTEPPS Teamwork Perceptions Questionnaire domain). (A) Average score for each of the 7 Mutual Support items (Question 1-Question 7) under the control and intervention conditions (error bars represent the SE). (B) Total subscale score (sum across Question 1-Question 7). Item-level Mann-Whitney *U* tests were Bonferroni-adjusted across the 7 items (α=.05/7≈.0071); the subscale-total comparison was a single confirmatory test and was, therefore, not Bonferroni-adjusted. The total subscale score improved significantly under the intervention condition (3.01 to 3.31; *P*=.003; Cohen *d*=1.24, large effect). At the item level, 2 items survived Bonferroni correction: Question 1 ("Staff assist fellow staff during high workload"; *P*=.003) and Question 2 ("Staff request assistance from fellow staff when they feel overwhelmed"; *P*<.001). Asterisks denote items significant comparisons: in panel (A), the items that survived Bonferroni correction (question 1, *P*=.003; question 2, *P*<.001; Bonferroni-adjusted α=.05/7≈.0071); in panel (B), the significant confirmatory subscale-total comparison (*P*=.003).



Mutual Support	Intervention, mean (SD)	Control, mean (SD)	*P value*	Cohen *d*	Bonferroni-adjusted[a]
Total subscale score	23.20 (1.74)	21.07 (1.71)	.003	1.24	Significant
1. Staff assist fellow staff during high workload	3.80 (0.41)	3.00 (0.76)	.003	1.31	Significant
2. Staff request assistance from fellow staff when they feel overwhelmed	3.67 (0.49)	2.43 (0.62)	<.001	2.20	Significant
3. Staff caution each other about potentially dangerous situations	3.20 (0.77)	3.07 (0.59)	.55	0.19	Not significant
4. Feedback between staff is delivered in a way that promotes positive interactions and future change	3.20 (0.94)	3.30 (0.59)	.98	−0.13	Not significant
5. Staff advocate for patients even when their opinions conflict with those of a senior member of the unit	3.13 (0.74)	2.93 (0.70)	.46	0.28	Not significant
6. When staff have a concern about patient safety, they challenge others until they are sure the concern has been heard	3.13 (0.64)	3.10 (0.81)	.98	0.05	Not significant
7. Staff resolve their conflicts, even when the conflicts have become personal	3.07 (0.70)	3.23 (0.68)	.49	−0.24	Not significant

a Bonferroni-adjusted: α=.007 (.05/7); Cohen *d*: 0.2=small, 0.5=medium, and 0.8=large.

Discussion

Effects on Teamwork Perceptions (T-TPQ)

This intervention study was conducted with adequate statistical power across all primary outcomes. Post hoc power analyses confirmed that the achieved sample yielded 90% power for the overall RT outcome (Cohen *d*=1.19), 87% power for Situation Monitoring (Cohen *d*=1.13), and 92% power for Mutual Support (Cohen *d*=1.24), all exceeding the conventional 80% threshold. The findings discussed below are, therefore, considered robust within the scope of this single-site intervention study.

The minimal interaction communication app developed in this study, which uses BLE Mesh network technology, has been shown to be extremely effective in streamlining communication and improving teamwork among nurses. The intervention resulted in a significant reduction in overall RT and a significant improvement in the T-TPQ scores for both Situation Monitoring and Mutual Support. These results strongly suggest that this app can be a practical solution to communication challenges in the nursing field.

The most significant factor in the overall reduction in RT in this study is considered to be the transformation of the communication style itself: from making voice calls to specific individuals to being able to send a brief help signal to the whole team at once, without speaking, when assistance is needed. This way of sending a quick signal to the whole team without speaking streamlined operations through different mechanisms, depending on the nature of the task.

Among the 6 request categories, the most robust improvements—those that survived Bonferroni correction across the 6 per-category comparisons (adjusted α=.0083)—were observed for "Care help" (37-1 s; *P*=.003) and "Information sharing" (36-18.5 s; *P*=.006). For "Care help," assistance can typically be rendered by any available nurse with the requisite physical capacity, and sending the same signal to the whole team at once enabled instant access to the team's collective availability, reducing the time cost of locating a willing helper. For "Information sharing," sending the same signal to the whole team at once allowed clinically relevant updates to reach the team without interrupting the work of a specific recipient: a single signal reaches every team member at once, rather than being relayed from 1 nurse to the other.

Two additional categories—"Consultation" (41-20 s; *P*=.01) and "Confirmation to nurse" (39-19 s; *P*=.02)—showed nominal reductions of comparable magnitude but did not survive Bonferroni correction for multiple comparisons across the 6 categories. These results are best interpreted as hypothesis-generating: they are directionally consistent with the broadcast mechanism described above (ie, requests that can be fulfilled by any qualified peer rather than by a specific individual), but the present feasibility sample lacks the statistical resolution to confirm independent effects in these subcategories. The nonsignificant change in "Inquiry for Items" (131.5-27.5

XSL•FO
RenderX

s; *P*=.25), despite a numerically large median reduction, is likewise underpowered owing to the small per-category sample (n=4 per condition).

On the other hand, the limitations of this model were also revealed by the results for "Confirmation to leader." RT did not improve in this category alone because the request was directed to a specific individual in the team who had an irreplaceable role. In this case, the issue was not the time it took to find the leader but securing the leader's availability to respond, a problem that could not be solved by a simple broadcast. This is related to the busyness and digital stress of leadership positions, as pointed out by Laukka et al [16], and suggests that the technology to be introduced should be optimized according to the nature of the problem to be solved.

Overall, these category-level patterns are consistent with the broadcast mechanism described above.

The T-TPQ served as the assessment tool to measure teamwork perception changes, which showed substantial improvements in both the Situation Monitoring and Mutual Support categories. The app's real-time notification system, which makes each staff member's current situation visible to the rest of the team, likely caused the improvement in Situation Monitoring scores by giving every nurse the same picture of where each colleague is and what each is doing. The Mutual Support score improvement stands as a crucial discovery from this research. The literature shows that measuring Mutual Support in nursing teams becomes difficult because of the difference between staff-provided assistance and staff-received assistance. The app successfully reduced psychological barriers to help requests, according to our findings, which show substantial improvements in "Staff assist fellow staff during high workload" and "Staff request assistance from fellow staff when they feel overwhelmed." The decreased barriers enabled staff members to request help without hesitation, while available staff members provided support voluntarily, thus solving the problem identified in previous studies [17].

On the other hand, no significant difference was observed in items requiring more complex and advanced communication, such as "Staff caution each other about potentially dangerous situations" and "Staff resolve their conflicts." This indicates that while the app contributes to providing an "impetus" for and "streamlining" communication, it does not guarantee the content or quality of the interaction itself. Among individual Situation Monitoring items, question 1 ("Staff effectively anticipate each other's needs") and question 3 ("Staff exchange relevant information as it becomes available") did not reach statistical significance (*P*=.13 and *P*=.44, respectively). Critically, the overall subscale totals for both Situation Monitoring and Mutual Support were significant (*P*=.001 and *P*=.003, respectively), with large effect sizes (Cohen *d*=1.13 and 1.24), supporting the robustness of the primary findings.

Previous research shows that even a simple, content-light signal can convey complex and rich meaning when senders and receivers share an existing working relationship, along with the social, cultural, and emotional context that comes with it [18]. This experiment used these brief team-wide signals—sent without speaking, to the whole team at once—as an example of how a short, content-light signal can still convey rich meaning when the people involved share the clinical context. Because senders and receivers already share the same ward context, a short signal can carry several layers of meaning, and meaningful cooperative behavior becomes possible with very little explicit content. The results show that what matters is not the amount of information carried by the signal, but the depth of interpretation that the professionals on each end of the signal bring to it.

Minimal communication should, therefore, be viewed not as a limitation but as a starting point that draws on the nurses' interpretive skills, builds on their working relationships, and uses their clinical experience to produce a shared understanding.

The result of this study suggests that successful outcome is determined not only by the superiority of the technology itself but also by how it is accepted and used in the field. Livesay et al [19] point out that a major reason for the failure of digital health adoption is the lack of nurse involvement in the design and implementation process, as well as inadequate training, indicating that technology lacking a field perspective can disrupt workflows. Furthermore, Fitzpatrick [20] warns of the overlooked negative aspect of digitalization, where the increase in communication volume can, conversely, contribute to health care provider burnout. Additionally, the adoption rate among staff can be an issue in this type of technology implementation. For example, in a contact tracing study by Curtis et al [21], the participation rate, especially among physicians, was reported to be low despite the use of anonymous tags, possibly due to privacy concerns. The key to the success of these technological innovations lies in the close collaboration between nurses, who have a deep understanding of the nursing field, and researchers and developers in the engineering field. Wang et al [22] emphasize that deep collaboration between nursing professionals and Internet of Things technology developers is essential for the effective implementation of these technologies. It should also be acknowledged that this study evaluated teamwork perceptions using only 2 of the 5 T-TPQ subscales, and that self-report questionnaire data had inherent limitations in capturing the full complexity of team processes. Perceptual scores reflect nurses' subjective appraisals and may be susceptible to response bias or social desirability effects. Objective behavioral measures of teamwork, such as structured observational coding or interaction analysis, would complement the self-report data in future studies. Furthermore, the 3 T-TPQ subscales not assessed in this study (Team Structure, Leadership, and Communication) may also have been influenced by the intervention in ways not captured here, and future research should consider comprehensive multidimensional teamwork assessments. Only when digital health is built on the needs of the clinical floor and keeps the nurses who actually use it firmly in view can it achieve its ultimate goal of improving the quality and safety of health care [19-25].

Limitations

This study has several limitations, including a small sample size of 8 nurses per shift at a single acute care ward; a short study

XSL•FO
RenderX

period of 6 days, which requires caution in generalizing the results; the lack of verification of the Red Call function for emergencies; and the fact that the evaluation was primarily quantitative. For RT data, each nurse could contribute multiple events across multiple days. The primary analysis addressed this by aggregating events to a per-nurse median before testing (n=15 per condition), and a linear mixed effects model was fitted as a sensitivity check; both analyses yielded congruent results. The per-category RT analyses, however, retained the event-level unit of analysis because each nurse contributed too few events per category for per-nurse aggregation within a category to be informative. Per-category *P* values were corrected for multiplicity across the 6 request categories using Bonferroni adjustment (adjusted α=.0083); 2 categories (Care Help and Information Sharing) remained significant after correction, while 2 further categories (Consultation and Confirmation to Nurse) showed nominal reductions that did not survive correction and should be regarded as hypothesis-generating. Additionally, the alternating-day, between-subjects design is susceptible to temporal confounding from factors, such as staffing variation, patient census, and workload differences between study days. Although day-level clustering was examined in a prespecified sensitivity analysis (Table S2 in Multimedia Appendix 1) and did not alter the substantive conclusion, day-level confounders cannot be fully excluded with the present design and require confirmation in a larger study with a greater number of study days.

Conclusions

This study demonstrated that a simple signal communication app using BLE Mesh network technology significantly reduces RTs among nurses and markedly improves the core teamwork elements of Situation Monitoring and Mutual Support. This indicates that the app has immediate and practical value in improving the inefficiency of communication in nursing practice and enhancing the quality of team-based medical care.

Acknowledgments

During the preparation of this work, the authors used NotebookLM to improve the English grammar, spelling, and readability of the manuscript. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published study.

Funding

This research was conducted with funding for education and research provided by the University of Tsukuba.

Data Availability

The datasets generated and/or analyzed during this study are not publicly available due to the subanalysis conducted but are available from the corresponding author upon reasonable request. Researchers interested in accessing the data for future studies may contact YM, the co-responding author, to inquire about data availability.

Authors' Contributions

Conceptualization: YM, MH, KS
Formal analysis: YM
Funding acquisition: KS
Investigation: YM, MH
Methodology: YM
Supervision: MH, KS, YI
Writing—original draft: YM
Writing—review and editing: MH, KS, YI

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables for the linear mixed-effects sensitivity analyses.
[DOCX File, 28 KB - nursing_v9i1e85599_app1.docx]

References

1. Miller ET, Deets C, Miller RV. Nurse call systems: impact on nursing performance. J Nurs Care Qual 1997 Feb;11(3):36-43. [doi: 10.1097/00001786-199702000-00007] [Medline: 9029852]
2. D'Amour D, Ferrada-Videla M, San Martin Rodriguez L, Beaulieu MD. The conceptual basis for interprofessional collaboration: core concepts and theoretical frameworks. J Interprof Care 2005 May;19 Suppl 1:116-131. [doi: 10.1080/13561820500082529] [Medline: 16096150]

XSL•FO
RenderX

3. Institute of Medicine (US) Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century: National Academies Press; 2001. [doi: 10.17226/10027]
4. Keshtkar L, Bennett-Weston A, Khan AS, et al. Impacts of communication type and quality on patient safety incidents: a systematic review. Ann Intern Med 2025 May;178(5):687-700. [doi: 10.7326/ANNALS-24-02904] [Medline: 40228297]
5. Meneses-La-Riva ME, Fernández-Bedoya VH, Suyo-Vega JA, Ocupa-Cabrera HG, Grijalva-Salazar RV, Ocupa-Meneses GDD. Enhancing healthcare efficiency: the relationship between effective communication and teamwork among nurses in Peru. Nurs Rep 2025 Feb 7;15(2):59. [doi: 10.3390/nursrep15020059] [Medline: 39997795]
6. Institute of Medicine (US) Committee on the Robert Wood Johnson Foundation Initiative on the Future of Nursing, at the Institute of Medicine. The Future of Nursing: Leading Change, Advancing Health: National Academies Press; 2011. [doi: 10.17226/12956]
7. Galiano MA, Moreno Fergusson ME, Guerrero WJ, et al. Technological innovation for workload allocation in nursing care management: an integrative review. F1000Res 2023;12:104. [doi: 10.12688/f1000research.125421.1]
8. Hwata C, Gatera O, Ritsch N, et al. Evaluation of the awareness of electromagnetic compatibility and interference for improved medical device management observed in selected Rwanda district hospitals. Medicine (Baltimore) 2024 Dec 27;103(52):e41179. [doi: 10.1097/MD.0000000000041179] [Medline: 39969313]
9. Vazquez Guillamet MC, Rjob A, Zhang J, et al. Leveraging Bluetooth Low-Energy technology to improve contact tracing among healthcare personnel in hospital setting during the coronavirus disease 2019 (COVID-19) pandemic. Infect Control Hosp Epidemiol 2024 Apr;45(4):546-548. [doi: 10.1017/ice.2023.227] [Medline: 37982262]
10. Weller J, Boyd M, Cumin D. Teams, tribes and patient safety: overcoming barriers to effective teamwork in healthcare. Postgrad Med J 2014 Mar;90(1061):149-154. [doi: 10.1136/postgradmedj-2012-131168] [Medline: 24398594]
11. Parker AL, Forsythe LL, Kohlmorgen IK. TeamSTEPPS®: an evidence-based approach to reduce clinical errors threatening safety in outpatient settings: an integrative review. J Healthc Risk Manag 2019 Apr;38(4):19-31. [doi: 10.1002/jhrm.21352] [Medline: 30212604]
12. Unoki T, Matsuishi Y, Tsujimoto T, et al. Translation, reliability and validity of Japanese version the TeamSTEPPS® teamwork perceptions questionnaire. Nurs Open 2020;8(1):115-122. [doi: 10.1002/nop2.609] [Medline: 33318818]
13. Hwang JI, Ahn J. Teamwork and clinical error reporting among nurses in Korean hospitals. Asian Nurs Res (Korean Soc Nurs Sci) 2015 Mar;9(1):14-20. [doi: 10.1016/j.anr.2014.09.002] [Medline: 25829205]
14. Ballangrud R, Husebø SE, Hall-Lord ML. Cross-cultural validation and psychometric testing of the Norwegian version of the TeamSTEPPS® teamwork perceptions questionnaire. BMC Health Serv Res 2017 Dec 2;17(1):799. [doi: 10.1186/s12913-017-2733-y] [Medline: 29197381]
15. Keebler JR, Dietz AS, Lazzara EH, et al. Validation of a teamwork perceptions measure to increase patient safety. BMJ Qual Saf 2014 Sep;23(9):718-726. [doi: 10.1136/bmjqs-2013-001942] [Medline: 24652512]
16. Laukka E, Hammarén M, Pölkki T, Kanste O. Hospital nurse leaders' experiences with digital technologies: a qualitative descriptive study. J Adv Nurs 2023 Jan;79(1):297-308. [doi: 10.1111/jan.15481] [Medline: 36300725]
17. Bae SH, Farasat A, Nikolaev A, et al. Nursing teams: behind the charts. J Nurs Manag 2017 Jul;25(5):354-365. [doi: 10.1111/jonm.12473] [Medline: 28294446]
18. Kaye J. I just clicked to say I love you: rich evaluations of minimal communication. Presented at: CHI '06 Extended Abstracts on Human Factors in Computing Systems; Apr 22-27, 2006. [doi: 10.1145/1125451.1125530]
19. Livesay K, Petersen S, Walter R, Zhao L, Butler-Henderson K, Abdolkhani R. Sociotechnical challenges of digital health in nursing practice during the COVID-19 pandemic: national study. JMIR Nurs 2023 Aug 16;6(1):e46819. [doi: 10.2196/46819] [Medline: 37585256]
20. Fitzpatrick PJ. Improving health literacy using the power of digital communications to achieve better health outcomes for patients and practitioners. Front Digit Health 2023;5:1264780. [doi: 10.3389/fdgth.2023.1264780] [Medline: 38046643]
21. Curtis SJ, Rathnayaka A, Wu F, et al. Feasibility of Bluetooth Low Energy wearable tags to quantify healthcare worker proximity networks and patient close contact: a pilot study. Infect Dis Health 2022 May;27(2):66-70. [doi: 10.1016/j.idh.2021.10.004] [Medline: 34810151]
22. Wang B, Shi X, Han X, Xiao G. The digital transformation of nursing practice: an analysis of advanced IoT technologies and smart nursing systems. Front Med (Lausanne) 2024;11:1471527. [doi: 10.3389/fmed.2024.1471527] [Medline: 39678028]
23. Park J, Park J. Identifying the knowledge structure and trends of nursing informatics: a text network analysis. Comput Inform Nurs 2023 Jan 1;41(1):8-17. [doi: 10.1097/CIN.0000000000000919] [Medline: 35703232]
24. Nashwan AJ, Cabrega JA, Othman MI, et al. The evolving role of nursing informatics in the era of artificial intelligence. Int Nurs Rev 2025 Mar;72(1):e13084. [doi: 10.1111/inr.13084] [Medline: 39794874]
25. Stevens ER, Alfaro Arias V, Luu S, Lawrence K, Groom L. Technology integration to support nurses in an "inpatient room of the future": qualitative analysis. J Med Internet Res 2025 Jun 16;27:e68689. [doi: 10.2196/68689] [Medline: 40522717]

Abbreviations

BLE: Bluetooth Low Energy
RT: response time

XSL•FO
RenderX

T-TPQ: TeamSTEPPS Teamwork Perceptions Questionnaire

Edited by J Sarvestan; submitted 10.Oct.2025; peer-reviewed by T Davidson; revised version received 22.May.2026; accepted 25.May.2026; published 05.Aug.2026.

Please cite as:
Matsuishi Y, Hirokawa M, Inoue Y, Suzuki K
Facilitating Teamwork With Minimal Interaction via a Mesh-Based Communication Device in a Clinical Nursing Setting: Intervention Study
JMIR Nursing 2026;9:e85599
URL: https://nursing.jmir.org/2026/1/e85599
doi:10.2196/85599

© Yujiro Matsuishi, Masakazu Hirokawa, Yoshiaki Inoue, Kenji Suzuki. Originally published in JMIR Nursing (https://nursing.jmir.org), 5.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Impact of Electronic Medical Record–Enabled Versus Paper-Based Systems on the Quality of Nursing Handover: Comparative Case-Study

Lisa Browning[1,2*], BN, GradDipAdvClinNurs; Urooj Raza Khan[1,3*], BSc, MSc, Prof Doc of IT; Sandra Leggat[1,4*], BSc, MHSc, MBA, PhD; Nicholas Monypenny[5*], BN, GradCertInfSys; James H Boyd[1,4*], BSc, PhD

[1]Department of Public Health, La Trobe University, CS 1 Building, 1 Kingsbury Drive, Bundoora, Australia

[2]Nursing and Midwifery Professional Reports, Office of the Chief Nursing and Midwifery Officer, Eastern Health, Box Hill, Australia

[3]Australasian Institute of Digital Health, South Melbourne, Australia

[4]Eastern Health Institute, Eastern Health, Box Hill, Australia

[5]E-Health Program, Infrastructure and Digital, Eastern Health, Box Hill, Australia

[*]all authors contributed equally

Corresponding Author:
Lisa Browning, BN, GradDipAdvClinNurs
Department of Public Health, La Trobe University, CS 1 Building, 1 Kingsbury Drive, Bundoora, Australia

Abstract

Background: Ineffective clinical handover has the potential to compromise patient safety and quality of care. Standardizing the handover process is a widely adopted improvement strategy intended to reduce failures of information transfer. By enabling real-time access to patient information, electronic medical records (EMRs) could address communication issues inherent to nursing handover.

Objective: This case study sought to compare the quality of nursing handover occurring at EMR-enabled sites with that occurring at paper-based sites, within a single Australian public health service.

Methods: A comparative case study design was used, using quantitative data collected from observational audits of 60 handovers and posthandover surveys conducted in EMR-enabled and paper-based ward environments. Handover quality was measured through compliance with the organization's Clinical Handover Standard and staff-reported perceptions, enabling comparison between cohorts.

Results: Compared with paper-based wards, EMR-enabled wards demonstrated more consistent communication of clinical alerts and risks and fewer interruptions, whereas paper-based wards showed higher rates of bedside handover and patient engagement.

Conclusions: EMR implementation alone does not ensure high-quality nursing handover. EMR interface design and functionality may act as a barrier to bedside handover and patient engagement, and contribute to continued reliance on paper-based artifacts. Targeted EMR design and nursing informatics-led optimization are required to better support nursing handover as a complex and cognitively challenging communication process.

(JMIR Nursing 2026;9:e85909) doi:10.2196/85909

KEYWORDS

nursing handover; electronic medical record; electronic health record; health information systems; patient engagement

Introduction

Overview

Nurses play a vital role in determining the clinical outcomes of their patients, particularly in relation to ensuring patient safety and continuity of care [1]. Extensive research demonstrates that deficiencies in clinical communication processes represent a significant risk to patient safety [2] and are a common source of sentinel events [3]. When communication is incomplete, unclear, or inaccurate, patient safety may be compromised [1]. Clinical handover is a critical and complex health care communication process, serving as the primary mechanism for transferring responsibility and accountability between shifts [4]. Effective and high-quality nursing handover is a core patient safety mechanism, facilitating the timely, accurate, and efficient exchange of clinical information [5]. Consistent with the National Safety and Quality Health Service (NSQHS) Standards, effective handover requires information to be clear, comprehensive, and precise to support safe, appropriate, and continuous care [6].

XSL•FO
RenderX

Conversely, inadequate or poorly executed clinical handovers can adversely affect patient outcomes [5]. Ineffective communication during handover has been associated with inaccurate clinical assessments [7], delayed diagnoses [6], inappropriate treatments, and medication errors [5]. The impacts of these failures range from minor disruptions to care to severe and potentially catastrophic outcomes. As health care systems increasingly adopt electronic medical records (EMRs) to support clinical communication, understanding how these systems influence the quality and safety of nursing handover is essential.

Background

Efforts to structure and standardize nursing handover processes are often recognized as key improvement strategies for improving the quality and safety of clinical communication [8]. Standardized communication protocols and structured handover tools such as ISOBAR (Identification Situation Observations Background Assessment/Action Recommendation) [9] are "intended to prevent failures of information transfer" [10] by ensuring that information shared is relevant, concise, and focused [5]. Handover mnemonics like ISOBAR reduce the cognitive workload associated with performing handover [11,12] and were designed to overcome differences in communication styles, a factor contributing to communication errors [13]. By streamlining the handover process, these tools also enhance situational awareness within team-based environments, supporting the development of a shared mental model of a patient's clinical status [13,14]. Standardizing handover processes is therefore a frequent recommendation aimed at improving handover quality and patient outcomes.

In Australia, NSQHS Standards provide a quality assurance mechanism and "a nationally consistent statement about the standard of care consumers can expect from their health service organizations" [15]. NSQHS Standard 6, "Communicating for Safety includes the criterion for communication at clinical handover," encompassing structured clinical handover, defining minimum information content of handover, key principles of clinical handover, and engaging patients and caregivers in clinical handover processes [15].

Increasingly, bedside handover has been adopted as standard practice in Australian hospitals [1] and is promoted as a strategy to enhance patient engagement, transparency, and patient-centered care [16]. Best practice bedside handover requires nurses to communicate accurate, comprehensive, and contemporaneous clinical information in real time [5], often in the presence of the patient, their families, and caregivers. While this model of handover offers important benefits [1,16], it also increases the complexity of information exchange [17], requiring nurses to synthesize and communicate large volumes of clinical data efficiently [18] while maintaining patient engagement [16] and situational awareness [11,13]. These demands place heightened emphasis on the availability, accuracy, and usability of clinical information at the point of care, and on the systems through which that information is accessed and communicated during handover.

In this context, EMRs have been introduced internationally as a means of supporting complex clinical communication processes, including nursing handover. EMRs are clinical information systems that facilitate real-time access to patient data and provide structured system-level processes intended to support patient safety, quality improvement, and compliance with expected standards of practice [19,20]. Key patient safety mechanisms include electronic prescribing and medication administration, the use of alerts, checklists, and predictive tools and embedded evidence-based practices, supporting decision making at the point of care [19,21]. Beyond strengthening compliance, EMRs also enable continuous quality improvement by generating routinely collected data that can be used to monitor performance, identify variations in practice, and inform iterative improvement cycles [19].

With real-time access to up-to-date patient information, EMRs have been proposed as a means to address the communication challenges inherent to clinical nursing handover [9]. EMRs have the potential to support nursing handover through the use of structured EMR-based handover tools, thus aligning with the intent of NSQHS Standard 6, by supporting timely, accurate, and standardized communication of critical patient information [5].

The introduction of an EMR in itself does not guarantee improvements in clinical communication or patient safety. The literature remains equivocal regarding the extent to which EMRs improve efficiencies in health care, with some studies noting potential adverse effects on nursing handover, including increased work and decreased efficiencies [22,23]. Poorly planned or ineffective implementations of an EMR may disrupt established nursing workflows and may fail to provide the required cognitive support nurses require during handover [22,24]. Information fragmentation within EMRs and the need to navigate multiple screens or data sources may further increase nurses' cognitive workload during handover, introducing new risks for both patients and clinicians [25,26]. Understanding both the intended benefits and unintended consequences of EMR adoption is essential to assessing their value in practice [27,28].

While EMRs provide the technological infrastructure to support structured and standardized handover, the effective translation of EMR functionality into safe and high-quality nursing practice is underpinned by nursing informatics capability. Nursing informatics describes the knowledge and skills required by registered nurses to integrate nursing science with information and computer sciences to manage, communicate, and apply clinical data, information, and knowledge in practice [29].

The presence of digital technology alone is insufficient to improve clinical outcomes. Contemporary nursing practice requires nurses to navigate an increasingly complex digital environment, meaningfully engage with digital systems, adapt workflows, and effectively use information within the realities of clinical practice [30,31], including during time-critical activities such as nursing handover. Nursing Informatics, therefore, provides an interpretive and operational lens through which the impact of EMR-facilitated handover can be understood. Within this context, nurses play a critical role in shaping documentation practices, information use, and workflow integration [32,33] in ways that either support or undermine the

XSL•FO
RenderX

safety and quality objectives of a structured, best-practice nursing handover, occurring at the patient's bedside.

While handover standardization alone has the potential to improve the effectiveness and efficiency of nursing handover and contribute to safer patient care [15], high-quality research linking the implementation of EMR-facilitated structured nursing handover to impacts on patient-related outcomes is still lacking. A knowledge gap therefore exists, as does an opportunity for further exploration of EMR-mediated solutions that may serve to support, standardize, and therefore enhance nursing handover.

Despite the widespread implementation of EMRs and increasing adoption of bedside handover, there remains limited empirical evidence examining how EMR-facilitated structured nursing handover translates into practice, particularly when compared with handover occurring in non-EMR environments within the same health service. Understanding how differences in information systems, workflow integration, and nursing informatics capability influence the quality of handover is essential to determining whether EMRs achieve their intended safety and quality objectives.

This gap is particularly salient for health services implementing EMRs in a staged or hybrid manner, where variations in digital maturity create a natural opportunity to examine the impact of EMR-enabled handover in context. It is within this organizational and digital landscape that the present case study was situated.

Context

Eastern Health is a large public health service situated in Victoria, Australia, serving a wide catchment across Melbourne's east. The organization spans 6 local government areas, covering 2816 square kilometers and comprising 21 health service locations. Eastern Health's 3 largest acute hospital sites are Box Hill Hospital, Maroondah Hospital, and Angliss Hospital.

A full EMR was implemented at Box Hill Hospital in October 2017. At Angliss Hospital, Electronic Medication Prescribing and Administration was implemented in November 2016; however, inpatient clinical documentation has remained predominantly paper-based. At the time of the study (conducted March to May 2023), implementation of the full EMR across inpatient wards at Maroondah Hospital (including Electronic Medication Prescribing and Administration) and completion of the full EMR deployment at Angliss were planned for late 2025 and early 2026, respectively. As a result, Eastern Health operated within a hybrid digital environment with varying levels of maturity across sites.

The introduction of full EMR at Box Hill Hospital fundamentally altered the interface through which nurses accessed, synthesized, and communicated clinical information during handover. In contrast, nurses at Maroondah and Angliss continued to rely primarily on paper-based documentation supported by other limited digital systems. These contrasting documentation environments, operating within the same organization, provided a structured basis for systematic comparison of EMR-enabled and predominantly paper-based nursing handover within a shared organizational context.

The sampling frame comprised nursing handovers conducted on all acute general medical and acute general surgical wards across Box Hill, Maroondah, and Angliss Hospitals. Each site has a 24-hour emergency department and intensive care unit, providing comparable patient cohorts within their respective acute medical and acute surgical wards. The selection of these was intended to promote homogeneity between the EMR-enabled and non-EMR sites, thereby minimizing the impact of variability in patient acuity on nursing expertise and handover practices. A total of 12 wards were identified as suitable for inclusion in the study. The target population comprised clinical nursing staff responsible for conducting routine shift-to-shift handovers.

Study Purpose

The purpose of this case study was to examine and compare the quality of nursing handover conducted in fully EMR-enabled wards at Box Hill Hospital with handover occurring in predominantly paper-based documentation environments at Maroondah and Angliss Hospitals. Handover quality was examined within routine clinical practice and assessed using two complementary measures: (1) the degree of compliance with the organization's Clinical Handover Standard, and (2) clinical nursing staff's self-reported perceptions of handover quality.

This purpose aligns with the study's focus on understanding how differences in documentation environments and information access may shape the conduct and perceived quality of bedside nursing handover within a shared organizational context.

Aims

The primary aim of this study was to compare the quality of nursing handover in fully EMR-enabled wards and predominantly paper-based wards within acute inpatient settings.

Specifically, the study sought to compare nursing handover practices in fully EMR-enabled wards with those occurring in predominantly paper-based environments, by examining whether EMR presence was associated with:

- Higher-quality nursing handover, as measured by compliance with the organization's Clinical Handover Standard (Multimedia Appendix 1).
- Differences in duration of nursing handover.
- Differences in the level of patient engagement in the handover process.
- Differences in staff-reported perceptions of handover quality.

Methods

Design

To address the study purpose and aims, nursing handover practices were examined across EMR-enabled and predominantly paper-based wards within the same health service. This study used a comparative quantitative case-study design using 2 complementary quantitative data sources.

XSL•FO
RenderX

Case study research is "a distinctive form of empirical enquiry" [34] that investigates a contemporary phenomenon within its real-life context [35]. This approach is well-suited to health care settings, where complex clinical processes can be examined using observational methods alongside other data sources, allowing meaningful characteristics of real-world practice to be retained [34], while generating opportunities for more comprehensive explanations [36].

As described within the context section of this case study, the staged implementation of the EMR across Eastern Health provided a naturalistic opportunity to apply a comparative case study design, enabling examination of nursing handover practices in wards supported by a fully implemented EMR alongside those operating in predominantly paper-based environments. Comparative evaluations of health information systems remain relatively uncommon, with much of the existing literature relying on descriptive or correlational approaches [27]. As such, this design offered the potential to generate contextually grounded insights relevant to both local system optimization and the broader evidence base on formative evaluations of digital health systems.

This case study used 2 quantitative data collection instruments. First, structured observational audits of routine shift-to-shift nursing handover were conducted to assess handover quality based on observed compliance with key principles of Eastern Health's Clinical Handover Standard. Second, posthandover surveys were administered to capture clinical nursing staff's perceptions of the quality of handover received, with responses quantified using Likert-scales, responses to closed yes/no questions, and counts of affirmative responses. For both data sources, comparisons were undertaken between handover occurring in the EMR-enabled wards and those in predominantly paper-based wards. This multi-source quantitative approach enabled direct observation of the phenomenon being studied alongside quantified staff perceptions of handover quality [34].

Consistent with guidance on health information system evaluation, the use of multiple data sources allowed complementary insights to be integrated, supporting a more comprehensive understanding of nursing handover within differing documentation environments [27]. EMR-supported handover was examined from a nursing informatics perspective, focusing on nurses' use of available documentation systems and the clinical information exchanged during routine handover practice.

The data collected and analyzed in this case study form part of a broader program of research examining nursing handover processes within EMR-enabled clinical environments. Subsequent analyses will explore nurses' interaction with, and use of, EMR functionality during handover in greater depth.

Participants

The study was conducted across acute general medical and acute general surgical inpatient wards within 3 hospital sites of a single public health service in Victoria, Australia. Participants in the observational audits and posthandover surveys were nurses conducting routine shift-to-shift nursing handovers on the 12 selected wards during the data collection period. The distribution of wards by site, documentation environment, and the number of handovers audited is presented in Table 1.

Table . Selected wards by site and number of handovers audited.

Site	Medical wards	Surgical wards	Medical record status	Total number of wards	Number of handovers audited
Box Hill	3	3	Full EMR[a]	6	30
Maroondah	2	2	Paper-based	4	20
Angliss	1	1	Paper-based	2	10
Total	6	6	—[b]	12	60

[a]EMR: electronic medical record.

[b]Not applicable.

For this comparative case study, the case was defined as the selected wards within the 3 hospitals of a single public health service. The unit of analysis was the individual shift-to-shift nursing handover. Data were collected between March and May 2023, across morning and afternoon shifts, encompassing both paper-based and EMR-enabled documentation environments.

Sampling

A convenience sampling approach was used, whereby nurses giving handover on the selected wards during scheduled data collection periods were invited to participate. To promote variation in observed practice, handovers were sampled sequentially across different days, shifts, and times within each ward to capture variation in routine practice. Five handovers per ward were targeted to ensure adequate representation of routine handover practices within each clinical context, while remaining feasible given the resource-intensive nature of direct observational auditing across multiple sites. Five nursing handovers were sampled per ward, resulting in a total of 60 observed handovers across the 12 wards.

Inclusion and Exclusion Criteria

Eligibility for participation was defined to ensure consistency across wards and sites.

Inclusion Criteria

Participants were eligible if they were registered nurses with current nursing registration, used by Eastern Health, and scheduled to participate in shift-to-shift nursing handover during the data collection period. All participants were required to provide informed consent.

XSL•FO
RenderX

Exclusion Criteria

Undergraduate nurses, assistants in nursing, agency nurses, and staff who declined to participate or did not consent to the collection of basic demographic information were excluded from the study.

Ethical Considerations

Ethics approval was obtained from Eastern Health's Human Research Ethics Committee, HREC number LR22-042-84698. Reciprocal ethics approval was obtained from La Trobe University. Written informed consent was obtained from all nurse participants involved in the observational audits and posthandover surveys, including nurses giving and receiving handovers. Patients whose handovers were observed were provided with written information about the observational audit in advance. Patients were informed that no personal identifiers or health information would be collected or recorded as part of the audit. Verbal consent to proceed with handover observation was obtained from patients immediately prior to handover commencement. Participants received no financial or other compensation for their participation.

Data Collection

Overview

A total of 60 observational audits and corresponding posthandover surveys were conducted between March and May 2023. Observational audits occurred during the morning-to-afternoon (AM to PM) shift-to-shift nursing handover. All observational audits were undertaken by a single auditor to ensure consistency. Both the observational audit tool and the posthandover survey were completed using paper-based instruments at the point of data collection. Immediately following each observed handover, the nurse receiving the handover completed a posthandover survey. All paper-based data were subsequently entered into REDCap by the same auditor later on the day of collection, following completion of data collection activities for that day.

Each nurse was observed delivering handover once; however, nurses may have received handover and completed the posthandover survey on more than one occasion, consistent with the sampling approach and routine practice, resulting in up to 120 participants.

For each participating ward, 5 observational audits and posthandover surveys were completed over a period of 1 to 2 days. The observational audits comprised 30 handovers in fully EMR-enabled wards and 30 handovers conducted and observed in predominantly paper-based documentation environments. As this study represented an observational comparative case study and an exploratory examination of nursing handover practices, formal sample size or power calculations were not undertaken.

Observational Audit

The observational audit was designed to quantify the quality of nursing handover by measuring compliance with the organization's Clinical Handover Standard (Multimedia Appendix 1). This Standard addresses 11 key principles underpinning best practice clinical handover and is aligned with the NSQHS Standard 6: Communicating for Safety. These principles are summarized in Table 2.

XSL•FO
RenderX

Table . Principles for best practice handover. Extracted from the Eastern Health Clinical Handover Standard, November 2022, available in Multimedia Appendix 1.

Number	Principle
1	Are undertaken using a consistent and structured process, "ISOBAR,"[a] which includes a minimum dataset.
2	Recognize that clinical handover requires effective communication between clinicians, including an opportunity to clarify information.
3	Maintain the patient's right to confidentiality and privacy; sharing of information is based upon the relevance and impact to care and outcomes
4	Promote inclusion of patients and, where appropriate, their caregivers, while acknowledging that sensitivity is required where other patients or visitors may overhear patient information.
5	Consider the need to engage interpreters if patients have communication difficulties, such as cultural and linguistic diverse backgrounds or hearing impairments.
6	Occur prior to, or at the time when clinician/s transfer care and accountability, and acknowledge the transfer of accountability for some or all of the patient's care.
7	Ensure adequate preparation prior to undertaking a clinical handover, making certain that the process is efficient and that all relevant information is transferred.
8	The clinician providing direct care leads the clinical handover, and where possible, handover occurs as face-to-face communication.
9	Respect the importance of handover, with minimal interruptions or distractions.
10	Within mental health settings, a team handover occurs prior to individual handover to ensure all staff are aware of the current milieu and risk issues present within the unit.
11	Documentation of the handover is recorded within the progress notes, or equivalent, in accordance with the Clinical Documentation Standard.

[a]ISOBAR: Identification Situation Observations Background Assessment/Action Recommendation.

Health service organizations are required to implement governance frameworks, systems, and processes that support effective clinical communication during handover. Locally contextualized clinical handover standards provide a mechanism through which NSQHS requirements are operationalized at the organizational level, thereby supporting patient safety and quality of care [15].

Audit Tool Development

The observational audit tool was developed to assess whether handovers complied with established principles of effective clinical communication. Tool development occurred in partnership with Eastern Health's Communicating for Safety Clinical Risk Governance Committee (CFS CRGC) to ensure content validity and alignment with established clinical governance standards. Audit questions were aligned with those used in previous organization-wide Clinical Nursing and Midwifery Handover Audits, enabling comparison with existing organizational data and reinforcing alignment with established quality and safety processes.

In determining the final audit tool, several decisions were made:

- Two principles of the Clinical Handover Standard (Principles 10 and 11) were considered outside of the scope of this study and were therefore excluded.
- The location of handover was identified as a key variable of interest in the context of a newly introduced digital interface. Specified within the Standard is that "where possible, clinical handover will occur face to face and in the patient's presence (bedside handover)." This element was therefore included as a distinct principle for auditing.
- Handover efficiency was measured through handover duration and whether handover commenced and concluded on time, with all required staff present.
- Respect for the handover process was quantified by recording the number of interruptions occurring during each handover episode.

Ten principles were therefore included in the final audit framework. A total of 28 audit questions were used to assess compliance with these principles. Each question required a yes or no response, with the exception of interruptions and handover duration, which were recorded as counts and time (minutes), respectively. The full list of audit questions and their alignment with each principle is presented in Table 3.

XSL•FO
RenderX

Table . Observational audit questions in relation to each principle.

Principle	Questions	Response	Interpretation and data analysis
Handover occurs at the bedside	• Did handover occur at the bedside?	Y/N	% compliance with the principle per ward/cohort/overall.
Handover follows ISOBAR[a]	• Identify: Was the person/team involved in the handover introduced to the patient? • Identify: Was a 3-point ID check completed? • Situation: Was the current clinical situation of the patient (ie, stable, improving, deteriorating, and discharge plan) discussed? • Observation: Were recent clinical observations discussed? • Background: Was the relevant patient background/history, evaluation, and management to date shared? • Assessment: Were relevant clinical assessments, any concerns from assessments, or outstanding assessments/actions discussed? • Assessment: Were clinical risks (ie, falls, pressure injuries, delirium, malnutrition, behavior of concern [BOC]) discussed? • Assessment: Were clinical alerts (ie, safety concerns such as aggression/harms to others, infection control) discussed? • Request/recommendation: Were actions required and recommendations for ongoing care discussed, including responsibility for actions and agreed transfer of responsibility?	Y/N	% compliance for each question and % compliance with the principle per ward/cohort/overall.
Communication is effective	• Did the nurse leading the handover provide an opportunity to ask questions? • Did the nurse receiving the handover ask any questions to clarify the information received? • Did the nurse receiving the handover ask any questions to seek information that was not offered?	Y/N	Principle deemed met if any 1 question answered in the affirmative and % compliance with the principle per ward/cohort/overall.
Patient's right to confidentiality is maintained	• Was all the information provided at the bedside clinically relevant, avoiding information pertaining to sensitive social or clinical conditions?	Y/N	—[b]

XSL•FO
RenderX

Principle	Questions	Response	Interpretation and data analysis
Inclusion of the patient, family, or caregiver is promoted	• Was the patient greeted during the handover process? • Was the handover process explained to the patient? • Was the patient and caregiver/family member present included in the handover? • Was the patient given an opportunity to ask questions or discuss goals or raise concerns?	Y/N	% compliance for each question and % compliance with the principle per ward/cohort/overall.
Interpreters are used where needed	• Is the patient's first language or native language a language other than English? • Is there a language barrier present that would prevent the patient from being included in the handover? • Was an interpreter used during handover? • Was the future planning or booking, or did an interpreter discuss?	Y/N	% compliance with the principle per ward/cohort/overall. Logic used: If N recorded for Q1 or Q2=N/A If Y recorded for Q1 & Q2 (interpreter deemed needed) AND Y recorded for either Q3 OR Q4 (interpreter is used or planned)=Y/N.
Transfer of care and accountability is acknowledged	• Was there an acknowledgment of a transfer of care or accountability? *SELECT ALL THAT APPLY: • Discussion of actions required and responsibility for actions occurred • Review of treatment and care plan, clinical risks, management, and recommendations occurred • Review of information and clarification sought by the recipient • Acknowledgment and agreement from the recipient of the plan of care	Y/N. Principle deemed met if any 1 question answered in the affirmative	% compliance with the principle per ward/cohort/overall.
The direct care clinician leads handover and, where possible, delivered face to face	• Did the clinician providing direct care lead the handover? • Was the handover delivered face to face?	Y/N	% compliance for each question and % compliance with the principle per ward/cohort/overall.
Respect for process – minimized interruptions and distractions	Was the handover interrupted?	Number of interruptions recorded	Number of interruptions recorded per ward/cohort/overall.
Handover is efficient	Time taken to deliver the handover	Duration of handover in minutes (rounded)	Duration of handover per ward/cohort/overall.
Handover is efficient	Handover started and finished on time, with all required staff present?	Y/N	% compliance for the question.

[a]ISOBAR: Identification Situation Observations Background Assessment/Action Recommendation.

[b]Not applicable.

Posthandover Survey

The posthandover survey was designed to complement the objective data captured in the observational audits by capturing the immediate assessment of handover quality from the perspective of the handover recipient. Survey items were developed to quantify staff perceptions of satisfaction, quality, efficacy, and efficiency associated with each handover episode. This approach is consistent with the literature, which emphasizes the importance of accurate and comprehensive communication during clinical handover to support patient safety [6].

The survey was developed collaboratively with the organization's CFS CRGC to ensure alignment with clinical

XSL•FO
RenderX

handover principles and organizational standards, and comprised 11 items: 5 Likert-scale questions, 5 closed-ended (yes/no) questions, and 1 free text response. Likert-scale items were scored on a 5-point scale, with a composite score (out of 25) calculated for questions assessing satisfaction, quality, and perceived efficacy. Closed-ended items were analyzed using frequency counts and percentages. Free-text responses were reviewed descriptively; however, due to limited volume and specificity, they were not subjected to formal qualitative analysis. The full list of posthandover survey questions, response formats, and corresponding analytical approaches is provided in Table 4.

While formal psychometric analyses (eg, internal consistency) were not feasible due to the exploratory nature of the study, sample size, methodological triangulation with the observational audits supported interpretive validity.

XSL•FO
RenderX

Table . Posthandover survey questions.

Question	Construct	Response	Interpretation and data analysis
Are you satisfied with the handover you just received?	Satisfaction	Likert scale 1 - 5: 1=very dissatisfied 2=dissatisfied 3=neither satisfied nor dissatisfied 4=satisfied 5=very satisfied	Score (out of 5) for each question per ward/cohort/overall Composite score (out of 25) for questions 1 - 5 per ward/cohort/overall
Rate the overall quality of the handover received	Quality	Likert scale 1 - 5: 1=very poor 2=poor 3=average 4=good 5=excellent	Composite score (out of 25) for questions 1 - 5 per ward/cohort/overall
Did the handover follow the ISOBAR format?	Quality	Likert scale 1 - 5: 1=strongly disagree 2=disagree 3=neither agree nor disagree 4=agree 5=strongly agree	Composite score (out of 25) for questions 1 - 5 per ward/cohort/overall
Rate how accurate you believe the handover was?	Efficacy	Likert scale 1 - 5: 1=very poor 2=poor 3=average 4=good 5=excellent	Composite score (out of 25) for questions 1 - 5 per ward/cohort/overall
Rate how comprehensive you believe the handover was?	Efficacy	Likert scale 1 - 5: 1=very poor 2=poor 3=average 4=good 5=excellent	Composite score (out of 25) for questions 1 - 5 per ward/cohort/overall
Was the required content effectively communicated?	Efficacy	Y/N	% occurrence (count of yes) per ward/cohort/overall
Were you given the opportunity to clarify information?	Efficacy	Y/N	% occurrence (count of yes) per ward/cohort/overall
Was there any specific information you felt was missing?	Efficacy	Y/N	% occurrence (count of no) per ward/cohort/overall
If Y to Q8: In a few short words, briefly state what information you believe was missing	Efficacy	Free text	To be themed where possible
Was the duration of handover appropriate?	Efficiency	Y/ N - too short N - too long	% occurrence (count of yes) per ward/cohort/overall
Was the inclusion of the patient, family, or caregiver promoted?	Quality	Y/N	% occurrence (count of yes) per ward/cohort/overall

It should be noted that staff perceptions of patient engagement were higher than what was observed in audits, suggesting possible overestimation due to social desirability bias or limited understanding of effective patient inclusion during handover. This discrepancy highlights the importance of combining objective and self-reported measures when evaluating handover quality.

XSL•FO
RenderX

Data Analysis

Overview

Data were extracted from REDCap into Microsoft Excel by the research team and deidentified prior to analysis. All data handling and analysis complied with institutional governance, ethical, and privacy requirements.

Observational Audits

For each audit question, the frequency of affirmative responses was calculated. Compliance with each handover principle was determined by aggregating question-level responses and summarized at the ward level, by cohort (EMR-enabled vs paper-based), and for the overall sample. Descriptive statistics, including mean and median values where appropriate, were generated to enable comparison between cohorts. The detailed analytic approach, itemized by audit question, is presented in Table 3.

Nonparametric tests (Kruskal-Wallis rank-sum test) were used for intergroup comparisons due to nonnormal distribution of data. Handover duration and frequency of interruptions were analyzed descriptively.

Posthandover Surveys

Posthandover survey data were analyzed descriptively. Likert-scale responses were summarized using mean and median values at the ward level, by cohort, and overall. For closed-ended (yes/no) questions, frequencies and percentages of affirmative responses were calculated across the same groupings.

A composite score was derived from the 5 Likert-scale items assessing satisfaction, quality, and perceived efficacy to provide an overall measure of staff perceptions of handover quality. Comparative analysis was undertaken to examine alignment between staff-reported perceptions and observed handover quality as measured through compliance with the Clinical Handover Standard.

The posthandover survey also included a single free-text question inviting respondents to identify any information they believed was missing from handover. Responses were limited and lacked sufficient depth for qualitative analysis and were therefore not subjected to formal qualitative analysis or reported in this paper.

Given the exploratory nature of the study and the relatively small sample size, formal effect size estimation was not undertaken, and results were interpreted primarily descriptively.

Validity and Reliability

Overview

The observational audit and posthandover survey instruments were developed in collaboration with the organization's CFS CRGC, supporting content validity and ensuring alignment with established clinical governance standards. Audit items were mapped directly to the organization's Clinical Handover Standard and aligned with the NSQHS Standard 6: Communicating for Safety, ensuring construct relevance and consistency with best-practice frameworks.

To enhance the validity of the findings, objective measures of handover quality derived from observational audit compliance with the Clinical Handover Standard were compared with self-reported staff perceptions of the handover quality captured through the posthandover survey, providing methodological triangulation. Use of a single trained auditor and standardized data collection procedures further supported reliability and methodological rigor.

Observational Audits

Design of the observational audits was purposefully aligned with previous organization-wide Clinical Nursing and Midwifery Handover Audits. This alignment supported consistency in measurement and enabled comparison with existing audit data, strengthening confidence in the reliability of the observational audit approach.

Posthandover Survey

As the posthandover survey captured individual staff perceptions, inherent limitations relating to subjectivity and response bias are acknowledged. While the survey instrument had not been previously administered within the organization, its development was informed by established handover constructs and aligned with the observational audit framework, supporting interpretive validity.

Results

Observational Audits

Overview

Observational audit data were first analyzed to compare overall compliance with the Clinical Handover Standard between paper-based and EMR-enabled wards. Compliance with individual principles of the Standard was then examined to identify areas of variation between the 2 cohorts.

Overall Compliance With the Clinical Handover Standard

Overview

Overall compliance with the Clinical Handover Standard was assessed across 10 key principles. Eight of these principles were evaluated as either met or not met. Table 5 presents percentage compliance for each of these 8 principles, along with overall compliance by cohort, expressed using mean and median values. This composite measure, derived from compliance across the audited principles, indicated broadly comparable overall handover quality between EMR-enabled and paper-based wards.

XSL•FO
RenderX

Table . Overall compliance with the standard.

Principle	Paper-based[a]		EMR[b]-enabled[c]	
	Mean	Median	Mean	Median
Handover occurs at the bedside (%)	83	80	63	60
Handover follows ISOBAR[d] (%)	80	78	83	78
Communication is effective (%)	100	100	100	100
Patient's right to confidentiality is maintained (%)	100	100	100	100
Inclusion of the patient, family, or caregiver is promoted (%)	70	68	43	37
Interpreters are employed where needed	—[e]	—	—	—
Transfer of care and accountability is acknowledged (%)	100	100	100	100
The direct care clinician leads handover and, where possible, delivered face to face (%)	100	100	100	100
Overall compliance (first 8 principles) (%)	90	100	84	100
Handover is efficient (handover duration) (minutes)	3.87	4	4.13	3.5

[a]Respect for process: 7 interruptions.

[b]EMR: electronic medical record.

[c]Respect for process: 7 interruptions.

[d]ISOBAR: Identification Situation Observations Background Assessment/Action Recommendation

[e]Not applicable. The principle, "Interpreters are employed as needed,"demonstrated insufficient data for comparative analysis, with only 2 instances of interpreter need identified across the sample. This has been reported descriptively and excluded from composite compliance to avoid disproportionate influence on results.

Paper-based wards demonstrated slightly higher mean overall compliance with the Clinical Handover Standard (90%) than EMR-enabled wards (84%), although median compliance was equivalent across cohorts. This pattern suggests broadly comparable aggregate performance, with observed differences driven by variation across specific handover principles, rather than consistent divergence across all domains.

The 2 remaining principles, "Respect for Process" and "Handover is Efficient," were assessed using alternative metrics, quantified by the number of interruptions that occurred, and the duration of handover, respectively.

Interruptions During Handover

Across the 60 handovers observed, a total of 8 interruptions were recorded. Seven interruptions occurred in 2 paper-based wards, while only one interruption was observed in an EMR-enabled ward.

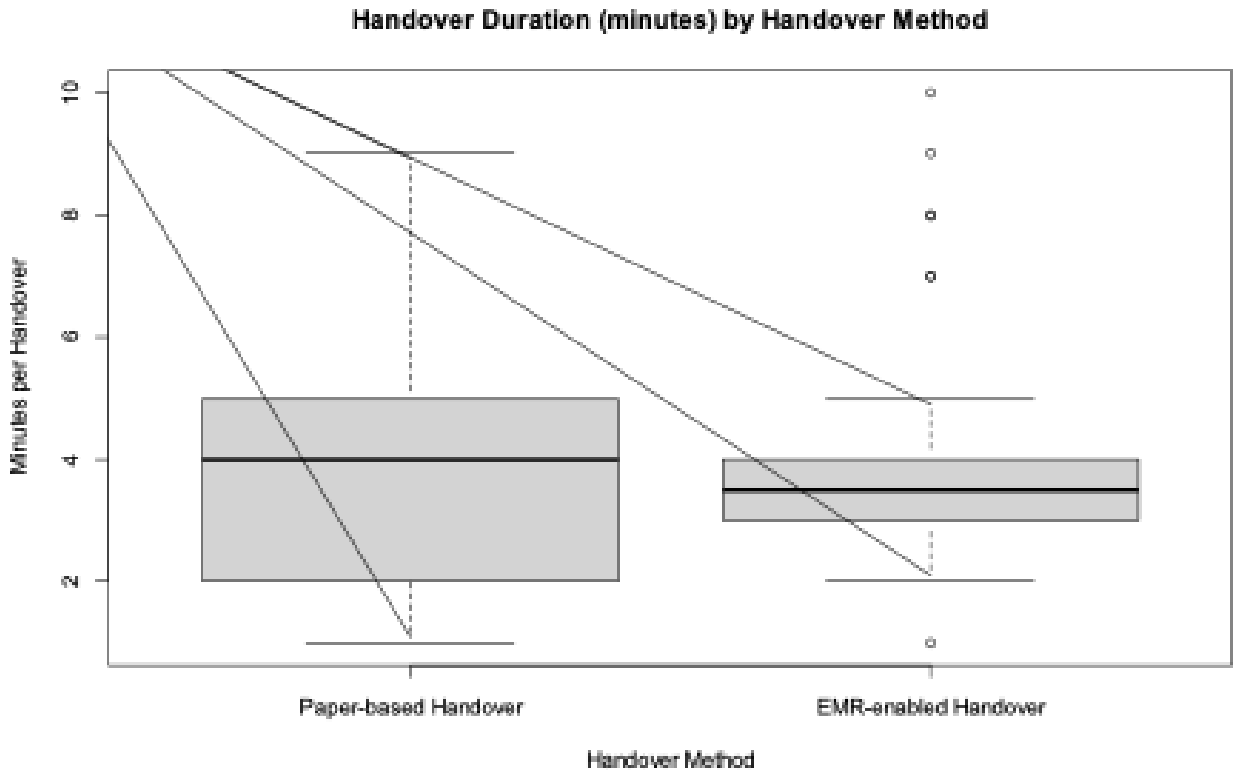
Handover Duration

Median handover duration was slightly longer in paper-based wards (4 min) compared with EMR-enabled wards (3.5 min). Initial analysis was considered using ANOVA; however, the data failed tests of normality (Shapiro–Wilk test for normality). Consequently, nonparametric methods, including the Kruskal–Wallis rank-sum test, were used. This analysis indicated no statistically significant differences between the paper-based and EMR-enabled wards, likely influenced by the small sample size.

As illustrated in Figure 1, the median handover duration was slightly longer for paper-based handover (4 min) compared with EMR-enabled handover (3.5 min), suggesting a marginally quicker handover in EMR settings. However, the IQR was wider for paper-based handovers (2.75 vs 1), indicating greater variability within the middle 50% of values, suggesting more consistent handover duration in EMR-enabled settings.

XSL•FO
RenderX

Figure 1. Handover duration by handover method.



Compliance Against Individual Principles of the Standard

Overview

Four principles demonstrated 100% compliance across all 12 wards:

- Communication is effective,
- Patient's right to confidentiality is maintained,
- Transfer of care and accountability occur,
- Direct care clinician leads handover.

Analysis of the remaining principles revealed several areas of divergence between cohorts, as outlined below.

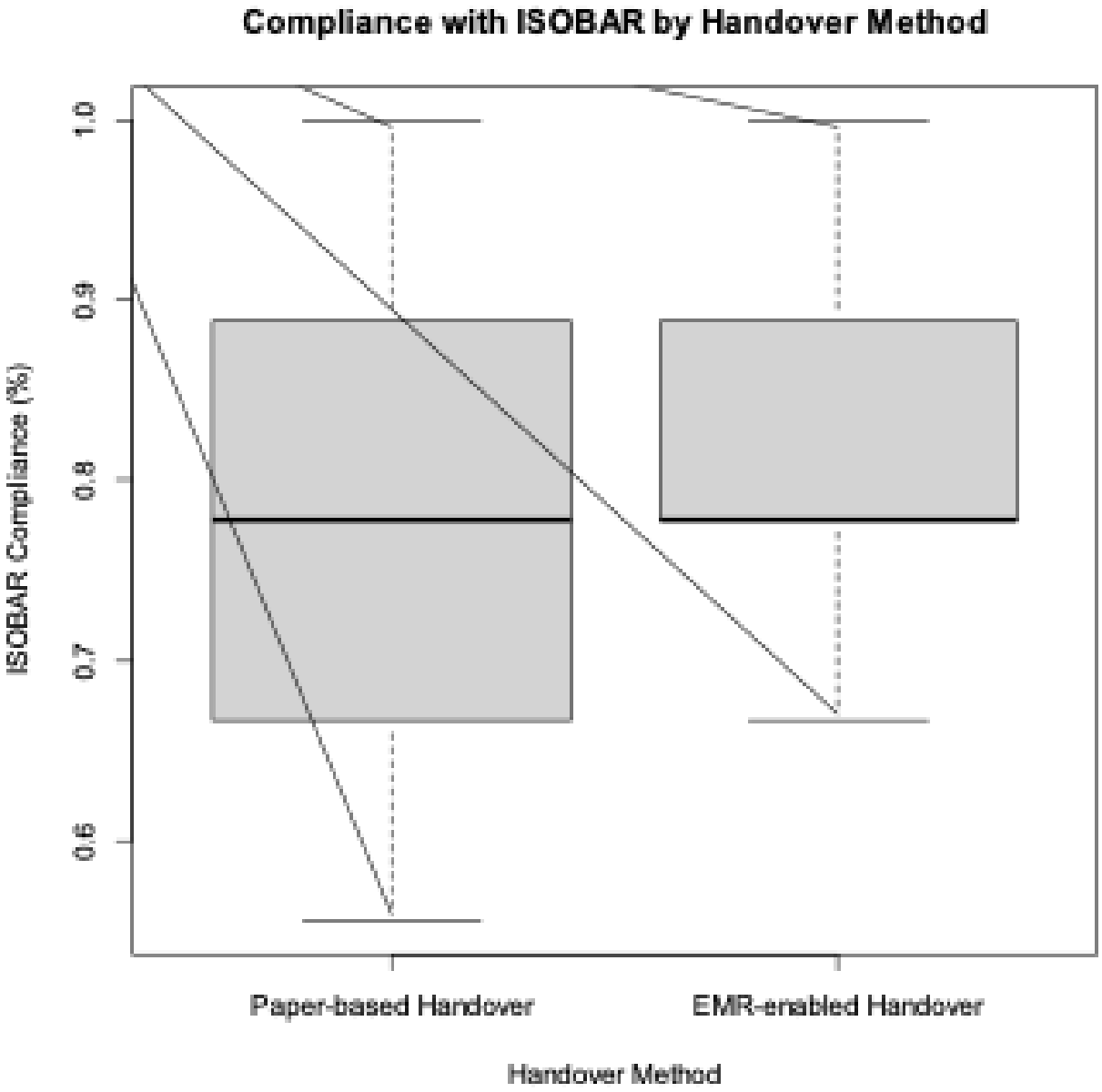
Compliance With ISOBAR

Compliance with all 9 ISOBAR elements was reviewed for both cohorts. Median compliance scores were equivalent between cohorts (0.778); however, EMR-enabled wards demonstrated a slightly higher mean compliance score (0.833) compared with paper-based wards (0.804). Paper-based wards exhibited greater variability, reflected by a wider IQR (0.194 vs 0.111). Both cohorts achieved the maximum possible compliance score of 1.0. These findings are illustrated in Figure 2.

Further examination of individual ISOBAR elements identified differences in the discussion of clinical risks and clinical alerts (Figure 3).

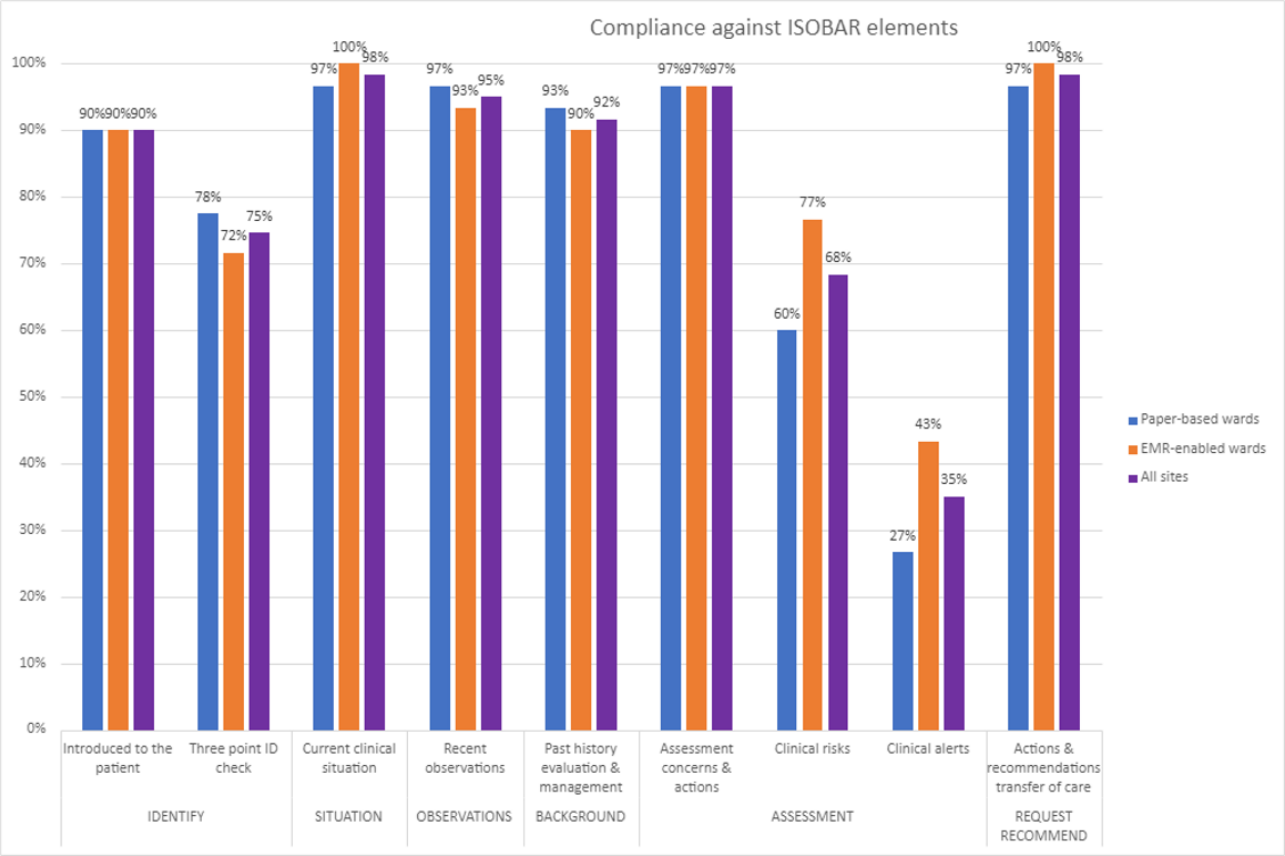
XSL•FO
RenderX

Figure 2. Compliance with ISOBAR by handover method. EMR: electronic medical record; ISOBAR: Identification, Situation, Observations, Background, Assessment/Action, Recommendation.



XSL•FO
RenderX

Figure 3. Compliance against ISOBAR elements. EMR: electronic medical record; ISOBAR: Identification, Situation, Observations, Background, Assessment/Action, Recommendation.



Clinical Risks and Clinical Alerts Were Discussed

Handover of clinical risks and clinical alerts occurred more frequently in EMR-enabled wards. Clinical risks were discussed in 77% of EMR-enabled handovers compared with 60% in paper-based wards, while clinical alerts were discussed in 43% of EMR-enabled handovers compared with only 27% in the paper-based wards.

Bedside Handover

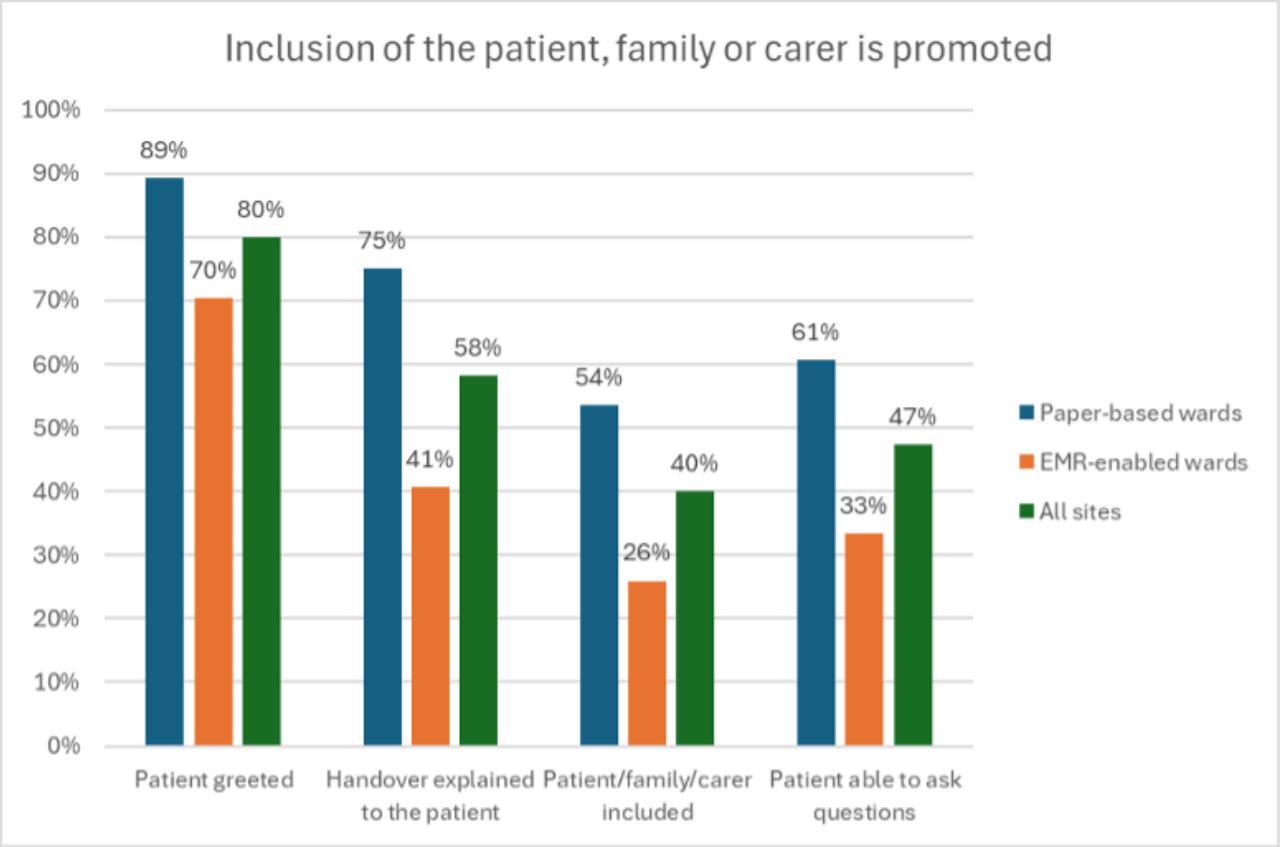
Handover occurred at the bedside more frequently in paper-based wards, with a mean compliance of 83%, compared with 63% in EMR-enabled wards.

Inclusion of the Patient, Family, or Caregiver

Promotion of patient, family, or caregiver inclusion occurred more frequently in paper-based wards. Compliance with all 4 patient inclusion indicators was observed in 70% of handovers in paper-based wards, compared with 43% in the EMR-enabled wards (Figure 4).

XSL•FO
RenderX

Figure 4. Inclusion of the patient, family, or caregiver is promoted. EMR: electronic medical record.



Comparison With Organization-Wide Handover Audits

Observational audit findings were compared with results from Eastern Health's Clinical Handover Nursing and Midwifery Audits conducted in 2022‐2023. Areas identified for improvement in the organization-wide audits aligned closely with findings in this case study, particularly in relation to bedside handover and patient and caregiver involvement.

Notably, overall performance relating to patient and caregiver inclusion was lower in this case study than in the organization-wide audits, suggesting potential variability in practice interpretation and reinforcing the need for further exploration of what constitutes effective patient inclusion and engagement in handover. Overall, these observational findings suggest that EMR availability was associated with variation in how handover was enacted across specific domains, rather than a uniform improvement or decline in overall handover quality.

Posthandover Surveys

Overview

Analysis of the posthandover survey data commenced with the calculation of a composite score derived from the 5 Likert-scale questions assessing staff satisfaction, perceived handover quality, adherence to ISOBAR, accuracy, and comprehensiveness of handover. A maximum composite score of 25 was possible. Results for individual items and composite scores by cohort are presented in Tables 6 and 7.

XSL•FO
RenderX

Table . Posthandover survey results (Likert scale questions). Overall score across all survey questions was 92% (median 97%) for paper-based and 91% (median 97%) for electronic medical record (EMR)-enabled handover.

Question	Paper-based mean (score)	Paper-based mean (%)	Paper-based median (score)	Paper-based median (%)	EMR-enabled mean (score)	EMR-enabled mean (%)	EMR-enabled median (score)	EMR-enabled median (%)
1. Are you satisfied with the handover you just received?	4.73	95	5	100	4.67	93	5	100
2. Rate the overall quality of the handover received	4.60	92	5	100	4.43	89	5	100
3. Did the handover follow the ISOBAR[a] format?	4.40	88	5	100	4.53	91	5	100
4. Rate how accurate you believe the handover was?	4.80	96	5	100	4.70	94	5	100
5. Rate how comprehensive you believe the handover was?	4.60	92	5	100	4.47	89	5	100
Composite score (Q1-Q5)	23.1	93	23	92	22.8	91	23	92

[a]ISOBAR: Identification Situation Observations Background Assessment/Action Recommendation.

Table . Posthandover survey results (Closed questions). Overall score across all survey questions was 92% (median 97%) for paper-based and 91% (median 97%) for electronic medical record (EMR)-enabled handover.

Closed questions (Y/N)	Paper-based mean (%)	Paper-based median (%)	EMR-enabled mean (%)	EMR-enabled median (%)
6. Was the required content effectively communicated?	100	100	100	100
7. Were you given the opportunity to clarify information?	100	100	97	100
8. Was there any specific information you felt was missing? (count of no)	93	100	90	100
10. Was the duration of handover appropriate?	93	100	100	100
11. Was the inclusion of the patient, family, or caregiver promoted?	70	70	70	70

Composite scores demonstrated minimal difference between cohorts, indicating similar levels of staff satisfaction and perceived handover quality. Across 10 of the 11 survey questions, responses were overwhelmingly positive, with no mean scores falling below 88% in either cohort.

Staff perception of the promotion of patient, family, or caregiver inclusion was the only survey item demonstrating a notable reduction in perceived performance.

Staff Perception - Promotion of Patient Inclusion

While observed patient engagement during the observational audits was limited, staff-reported perceptions indicated higher levels of patient involvement during handover. Promotion of patient inclusion was recorded as equivalent across cohorts, with both paper-based and EMR-enabled wards reporting 70% affirmative responses to this survey item.

XSL•FO
RenderX

Discussion

Principal Findings

This case study examined the impact of EMR implementation on both the process and perceived quality of nursing handover. Particular attention was given to how digital information systems shape nurses' access to, synthesis of, and communication about patient information. In this study, handover quality was operationalized using 2 complementary process-level measures: observed compliance with the organization's Clinical Handover Standard, and nurses' self-reported perceptions of handover quality following shift-to-shift handover exchange. Together, these measures provide insight into how EMR-mediated workflows influence handover behaviors, rather than downstream patient outcomes. Overall, EMR implementation did not uniformly improve handover quality; rather, its effects were heterogeneous across process domains, partially supporting the study objectives.

Quantitatively, overall compliance with the Clinical Handover Standard did not differ meaningfully between EMR-enabled and paper-based wards, nor did overall staff perceptions of handover quality. However, analysis of individual observational indicators revealed variation across specific domains. These included bedside location of handover, promotion of patient engagement, communication of clinical risks and alerts, handover duration, and frequency of interruptions. These findings suggest that while aggregate measures of handover quality appeared comparable, EMR implementation influenced *how* handover was enacted across specific components of the process.

Observational Audits

EMR-Facilitated Handover Creates a Barrier to Performing Bedside Handover

Observational audit data demonstrated a lower frequency of bedside handover in EMR-enabled wards compared with paper-based wards.

Bedside handover encourages direct and real-time engagement with the patient [1]. Active patient participation in the process can provide immediate information on the patient's condition [1]. Moreover, exchanging clinical information in the patient's presence can assist in identifying patient concerns, promoting continuity, quality of care, and improved patient satisfaction [1]. Implementing bedside handover is therefore key to promoting and supporting patient engagement and participation in care.

In this study, the presence of the EMR appeared to act as a barrier to nurses performing bedside handover. This finding points to challenges associated with EMR implementation that may, in turn, reduce the level of patient engagement during handover. The following factors may help explain this finding.

Limitations of the Physical Environment

In the EMR-enabled wards in this case study, nurses access the EMR on mobile workstations via computers mounted onto purpose-built trolleys, commonly referred to as Workstations on Wheels (WOWs). While ergonomic and adequately mobile, these WOWs are still relatively large. Often observed during the observational audit was one nurse from the AM shift giving handover to 2 nurses on the PM shift. Frequently, both nurses receiving handover would each be interacting with their own WOW. At times, the nurse giving handover would also be accessing the EMR on their own WOW, though this appeared to occur less frequently, with the nurse giving handover most commonly referring to their own printed handover sheet, inclusive of hand-written notes.

Despite the EMR-enabled wards being located in a relatively new and modern facility with spacious patient rooms, the available physical space at the bedside remained insufficient for 3 nurses and 2 WOWs. The devices are bulky and difficult to maneuver. As a result, staff appeared reluctant to take the WOWs fully into patient rooms and were frequently observed remaining in or just outside the doorway. The lead researcher was also made aware that in some patient rooms within the EMR-enabled cohort, Wi-Fi connectivity was reduced, resulting in poor or limited EMR functionality at the bedside.

Social Barriers to Using Technology at the Bedside

In addition to these identified physical limitations, a more nuanced barrier is a social one, where the presence of the EMR at the patient's bedside may be perceived as incongruous with a therapeutic environment. A positive and workable nurse-patient relationship develops when trust is established and maintained through an open exchange of information, supporting the delivery of interpersonal care [37,38]. In inpatient hospital settings, the nurse-patient relationship develops at the patient's bedside. Use of active listening techniques during nurse-patient interactions, such as making and maintaining appropriate eye contact, facing the patient, and minimizing distractions [37,38], further contributes to this therapeutic connection. Introducing digital technology into the therapeutic bedside space may be perceived by nurses as intrusive. It may disrupt active listening practices and interfere with the development of the therapeutic relationship. The patient may therefore perceive that the information being shared is not valued.

Duffy et al [38] similarly note that EMR documentation at the point of care can distract nurses' attention from the patient. This may reduce communication quality, appropriate eye contact, and overall meaningful interaction [38]. Misto [39] explored the impact of electronic documentation in the patient's presence on the nurse-patient relationship and noted that nursing staff perceptions of the EMR's influence on the therapeutic relationship were mixed [39]. Positive impacts included enhanced access to data at the patient's bedside and the ability to document contemporaneously and more efficiently, resulting in more time to spend with the patient [39]. Negative impacts centered on the recurring premise that using the EMR at the bedside is a distraction, resulting in both 'missed opportunities to connect with patients,' and an increased likelihood that subtle, nonverbal cues could be missed [39]. This was attributed to nurses being preoccupied with interacting with the EMR rather than being fully attentive to the patient [39]. Nurses in that study

XSL•FO
RenderX

also reported that patients felt less valued when this occurred [39].

Gaudet [40] also explored the changes in nurse-patient interactions associated with the introduction of EMR documentation at the bedside, noting that this may create "an automatic, machine-like, task centered bedside environment" [40].

It is therefore possible that nurses feel that the insertion of a large and bulky, potentially distracting piece of digital technology into the space at the patient's bedside may negatively impact the therapeutic relationship, serving only to make handover more impersonal rather than enhance it.

Cognitive Barriers to Performing EMR-Facilitated Handover at the Bedside

A recent review article exploring the impact of an EMR implementation on nursing handover determined that existing EMR interfaces are not currently meeting the needs of nurses during handover [22]. Navigating the EMR has been shown to impact cognitive workload due to the range of tasks required of users [26]. Combined with the way information is presented in the EMR, this results in fragmented or siloed data that must be skimmed through, sorted, and deciphered [18,24,25]. In this study, nurses were frequently required to navigate multiple EMR screens to locate alerts, risks, and clinical updates, reflecting well-recognized information fragmentation within EMR systems [25]. In this context, fragmentation shifts cognitive effort away from clinical reasoning and toward system navigation [25,41].

Handover is already a cognitively challenging task, requiring the nurse to be able to swiftly analyze and synthesize patient data across multiple information sources [24,42,43] and engage in complex knowledge-sharing processes [18]. The expectation that this demanding task also occurs at the bedside, involves the patient, and is mediated through an EMR that is often difficult to navigate, may further increase this burden. This raises an important question: Does this combination unintentionally work against the goal of improving handover quality and patient engagement? Is the presence of the patient themselves now an additional barrier to achieving the desired level of handover quality? Gaudet [40] explores this conundrum where 'the complexity of the patients, the computerized prioritisations of tasks and the interruptions, limited the nurses' ability to nurture and become familiar with their patients, thus creating "an environment that does not put the patients' concerns at the center of care" [40].

From an informatics perspective, EMR-mediated handover requires nurses to simultaneously navigate, prioritize, and synthesize information. At the same time, they must maintain effective interpersonal communication with both colleagues and patients. Together, these demands increase cognitive workload during an already complex task. Given that a fully optimized EMR solution for nursing has not yet been realized [22], nurses may be reluctant to bring EMR-facilitated handover to the bedside. Doing so may be perceived as increasing complexity and potentially compromising handover quality.

Patient Engagement Was Better Promoted in Paper-Based Environments

Where a reduced frequency of bedside handover was observed in the EMR-enabled cohort, so too were lower levels of patient engagement and decreased promotion of patient, family, and caregiver inclusion in handover. Although the compliance gap for initially greeting the patient was smaller between the cohorts, meaningful patient inclusion cannot occur when handover takes place away from the bedside. This was demonstrated by the widening performance gaps for the remaining patient engagement indicators, "Including the Patient in Handover" and "Providing the Patient an Opportunity to Ask Questions."

Reduced frequency of bedside handover in the EMR-enabled wards was therefore identified as a stand-alone barrier to promoting patient engagement as measured by observational indicators of inclusion and the opportunity to ask questions.

Other Known Barriers to Bedside Handover and Patient Engagement

Despite the known benefits of bedside handover in improving patient engagement and supporting patient-centered care, studies continue to elucidate reluctance from nurses to perform handover at the bedside. Independent of EMR introduction, commonly cited barriers to bedside handover include concerns about patient privacy and confidentiality [44-46]. Nurses also report worries about being asked difficult questions or appearing unprofessional in front of patients [44,45]. Additional concerns relate to perceived prolongation of handover [47,48] and disruption to communication flow [46,47].

A pre-existing resistance to bedside handover for these myriad reasons must therefore be considered as a contributing factor to the gap in bedside handover implementation across both cohorts in this study. Where an existing baseline gap in practice may already be evident, the introduction of the EMR, and all the additional potential barriers it brings, has possibly further compounded this issue in the EMR-enabled wards, making bedside handover even less achievable.

Handover of Critical Information Occurred More Readily in EMR-Enabled Wards

In the hospital environment, EMRs are accessed by clinicians throughout the course of a patient's treatment, replacing bulky paper-based medical records. In this case-study context, an EMR was introduced and mobile WOWs provided, providing up-to-date clinical information at the point of care and during handover. An EMR facilitates real-time sharing of patient data and has the potential to improve efficiency and care coordination [49]. The results of the observational audit support this, with improved handover of critical information such as clinical alerts and clinical risks evident in the EMR-enabled wards compared with the paper-based cohort.

This also aligns with marginally better performance observed for overall ISOBAR compliance in the EMR-enabled wards, where more consistent practice in using the mnemonic may have contributed to the improved handover of risks and alerts.

XSL•FO
RenderX

Reliance on Other Tools to Support Handover

In paper-based environments in this case study, nurses predominantly rely on printed handover sheets generated by bed management software known as Patient Flow Manager (PFM). PFM is not linked to the organization's EMR and clinical information stored within it, and it does not form part of the medical record.

To populate these handover sheets, clinical information is transcribed from the scanned patient history or EMR and manually entered into PFM. Nursing staff add and update clinical information in PFM throughout their shift, including the removal of any outdated information. It is not routinely used by medical or allied health staff in this way. Unfortunately, a reliance on duplication and manual data entry processes has the potential to result in transcription errors, decreasing the accuracy of the information contained on these handover sheets. Conversely, the EMR is considered the source of truth for clinical patient information, inclusive of clinical alerts and clinical risks, which have been configured to provide a visual alert to clinicians.

In the EMR-enabled wards, nurses also use PFM and are provided with PFM-generated printouts at the commencement of their shifts. In these wards, however, the EMR is always accessible during handover and is frequently used to check information, most notably by the handover recipients. This is consistent with recent studies, which predominantly described EMR use during handover as verifying or double-checking information [24,25,41] and keeping up to date with the latest developments [18].

In paper-based environments, nurses rely heavily on printed handover sheets. The information on these sheets may be outdated, incomplete, or inaccurate. Where critical changes in the patient's condition occur, it is imperative that nurses are immediately aware of changes to any clinical alerts and clinical risks. This will ensure that the most appropriate care, interventions, and safety precautions can be implemented to promote patient and staff safety.

Persistent reliance on PFM-generated paper handover sheets was observed in both cohorts, including the EMR-enabled wards. In this study, this pattern signaled a socio-technical misalignment between the EMR design and nursing handover workflows. Despite the EMR being the authoritative source of clinical information, nurses continued to depend on nonintegrated artifacts to support their cognitive load, information sequencing, and real-time note-taking during handover. This behavior suggests that existing EMR interfaces did not adequately support nurses' information needs during handover, reinforcing findings that poor task–technology fit drives workarounds and partial adoption of digital systems [22].

Handover Was More Efficient in the EMR-Enabled Wards

Observational audit data demonstrated reduced median handover duration in EMR-enabled wards compared with paper-based wards. Performance was also more consistent in the EMR-enabled environments, with more variability of duration noted in the paper-based wards.

Across a number of studies, accessing and navigating the EMR has been described as time-consuming [24,41], cumbersome [41], arduous [24], and not necessarily able to provide pertinent information quickly [43].

Despite these potentially negative connotations, studies reporting specific impacts of EMR implementation on handover duration offered mixed results. Hertzum and Simonsen [50] sought to understand the impact of a trial EMR implementation, finding that there was no significant difference in handover duration postintervention [50]. Correspondingly, Alghenami's [51] 2012 study exploring the role of EMRs in structuring handover communications found that nurses did not identify the EMR as a factor that would impact handover duration [51]. There is growing evidence, however, that use of specific EMR-structured handover tools or EMR-generated printouts can assist in making handover quicker and more efficient [52-54]. It is important to note, nonetheless, that positive outcomes are highly contingent on whether these tools are adequately aligned with the nature of the clinical work [55] and offer sufficient functionality to support the task [56]. Conversely, where an EMR interface does not adequately suit the task at hand, it fails the end user, which may result in incomplete adoption of the new technology and a continued reliance on paper-based tools [22]. Ultimately, when a gap in functionality or design is not readily addressed, or poor user experience persists, this has the potential to lead to user frustration and even boycott [56].

At Eastern Health, in the EMR-enabled wards, nurses were observed navigating various parts of the EMR during handover, yet they did not use any specific EMR-structured handover tools or EMR-generated printouts. Further to this, nurses across both cohorts were almost always observed consulting or making handwritten notes on paper handover sheets printed from PFM. Nurses in the EMR-enabled cohort demonstrated willingness to adopt EMR use during handover. However, the continued reliance on PFM-generated printouts suggests a gap in EMR design or functionality. While the PFM-generated printouts clearly provide support to nurses during handover, they remain problematic as they do not draw clinical information directly from the EMR. Ideally, an EMR-generated printout would replace those generated via PFM and thus provide the most up-to-date clinical information. Ultimately, these printouts must be purposefully designed by and for nurses in order that they fully support the task being performed.

Handover Occurring in EMR-Enabled Wards Experienced Fewer Interruptions

The paper-based cohort experienced more interruptions during handover, noting this may well be an anomaly given the very small sample size. Current research on the impact of interruptions during EMR use appears limited to those that occur via the EMR itself, such as automated alerts, reminders, and other clinical decision support mechanisms [57-59], or broader workflow interruptions associated with EMR implementations in general [60-62]. This case study sought to measure external interruptions caused by patients, phones, alarms, or other staff in the ward environment. The reason for fewer interruptions occurring in the EMR-enabled cohort is, therefore, entirely speculative.

XSL•FO
RenderX

A very well-researched field that could offer some useful comparisons is the frequency and impact of interruptions on medication administration workflows, which have the potential to result in medication errors. As a known high-risk activity, efforts to prevent and reduce interruptions that may lead to errors during medication rounds have also been thoroughly explored. Frequently used strategies to reduce and manage the risk of interruptions include the use of standardized checklists and protocols such as the 5 rights of medication administration, quiet zones for medication preparation, diversion strategies, and visual alerts such as wearable sashes and vests asking that others "Do Not Disturb" and "Do Not Interrupt" [63-65]. Nurses are therefore highly aware of this increased risk to patient safety, noting also that these types of interruptions are significant contributing factors to increasing mental workloads in hospital work environments [66]. One possible explanation is that engagement with the EMR may provide a visible cue to other staff that nurses are occupied with handover, potentially reducing external interruptions in a manner similar to visual "Do No Disturb" signals used during medication administration.

These findings should therefore be interpreted as indicative process trends rather than definitive performance improvements, and further investigation with larger samples or longitudinal designs may clarify their significance.

Staff Perception of Patient Engagement

Observational audit findings relating to patient engagement during nursing handover did not correlate with staff's self-reported perceptions of their own efforts to promote patient participation.

While bedside handover alone is insufficient to ensure meaningful patient engagement in the process [45], it remains a fundamental prerequisite in order for patient participation to occur. In this study, observational audits identified a compliance gap with bedside handover across both cohorts, with this gap more significant in the EMR-enabled wards. On this basis, lower observed rates of bedside handover in the EMR-enabled cohort would be expected to correspond with similarly reduced rates of self-reported promotion of patient engagement during handover. And yet the staff survey found that this was occurring in equal proportions across both cohorts, around 70% of the time.

The large mixed methodology study of Street et al [45] examining patient participation in handover may assist in explaining this, observing that whilst nurses do perceive the value in patient participation, 'the best way to effectively help patients be more involved in bedside nursing handover is poorly understood' [45]. The findings of this case study support this notion, suggesting that nurses, who may not understand how to effectively engage patients, may have limited insight into whether their handover practices genuinely facilitate patient engagement and participation.

This discrepancy may reflect an overestimation of patient engagement in staff self-reported data, potentially influenced by social desirability bias, whereby participants provide responses and report practices that are perceived as more acceptable [67]. In this instance, the self-reported practice aligned better with organizational expectations and the requirements of the Clinical Handover Standard, despite the decreased incidence of bedside handover necessary to facilitate the desired engagement.

Collectively, these findings highlight the value of using multiple measurement approaches when evaluating handover quality, as reliance on self-report alone may overestimate adherence to patient engagement principles.

Staff Satisfaction With Handover Quality

Despite the reduced occurrence of bedside handover, potential data fragmentation, workflow changes, and reliance on workarounds (including paper-based and nonintegrated artifacts) in EMR-enabled wards, staff satisfaction with overall handover quality remained high across both cohorts. Clinicians may report satisfaction with systems they have learned to cope with, even when those systems impose additional cognitive or workflow burdens [56]. These findings likely reflect nurses' ability to adapt to existing system constraints rather than optimal alignment between EMR design and nursing handover workflows.

Taken together, these findings suggest that staff satisfaction alone is an insufficient indicator of EMR effectiveness, reinforcing the need to examine how digital systems shape cognitive workload, workflow integration, and communication practices during nursing handover.

The contrast between consistently high staff satisfaction and observed variability in handover behaviors suggests adaptation and normalization of system constraints, reinforcing the importance of examining both objective and perceptual measures when evaluating digital clinical systems. Integrating observational and perceptual data strengthened interpretive validity and enabled identification of important differences between enacted and perceived practice that would not have been apparent using a single measurement approach.

Limitations

Several limitations should be considered when interpreting the findings of this study.

First, the relatively small sample size of 60 handovers limited the ability to detect statistically significant differences across all measures, particularly given the complexity and variability inherent in clinical handover practices. While sampling 5 handovers per ward enabled exposure to routine practice within each setting, the number of observations per cohort was insufficient to support more granular analyses or make more robust comparisons. As a result, findings should be interpreted as indicative of patterns of use and areas of variation rather than definitive estimates of practice prevalence. In addition, this study focused on process-level indicators of nursing handover quality, including observed compliance with the organization's Clinical Handover Standard and nurses' self-reported perceptions of handover quality. These measures provide important insights into how handover is conducted in different environments but do not directly assess patient outcomes or safety events. As such, findings should be interpreted as reflecting differences in handover processes only. Furthermore,

XSL•FO
RenderX

as an exploratory case study conducted within a single health service, the findings have limited generalizability beyond similar organizational and hybrid EMR contexts. However, the design provides rich, practice-based insight into how nursing handover is shaped within real-world digital environments.

Second, the use of a single auditor introduces the potential for investigator bias. To mitigate this risk, the observational audit tool was designed to minimize subjectivity, with items structured wherever possible to limit reliance on auditor interpretation. Consistent use of a single trained auditor also supported standardization of data collection across sites.

Third, a degree of Hawthorne effect may have occurred, as participants were aware that they were being observed. This is particularly relevant for highly visible elements of the Clinical Handover Standard, such as conducting handover at the bedside and performing patient identification checks. While this may have influenced overall adherence to the Clinical Handover Standard, it is unlikely to have affected the comparative analysis between the 2 cohorts, as any observation effect would have been equally probable in both groups.

Finally, the study focused exclusively on shift-to-shift nursing handover within inpatient wards and did not include handovers occurring in other settings such as interdepartmental transfers and other transitions of care. Future research should consider observing and evaluating handovers across a broader range of clinical contexts, with particular attention to how EMRs are configured and used to facilitate and support nursing handover under varying workflow, environmental, and cognitive demands.

Conclusions

This case study demonstrates that EMR implementation has the potential to produce both beneficial and unintended effects on the quality of shift-to-shift nursing handovers. While EMR-enabled wards showed positive impacts in accessing and communicating up-to-date clinical information, particularly in relation to clinical alerts and risks, these benefits did not translate uniformly across all dimensions of handover quality.

Findings indicate that a pre-existing reluctance to perform bedside handover may be exacerbated in EMR-enabled environments, with negative implications for patient inclusion and engagement. Limitations related to the physical environment, EMR interface and design, social factors, and cognitive factors appear to impede nurses' ability to consistently promote patient participation during handover. These findings suggest that nurses may not have a clear or comprehensive understanding of the actions necessary to actively promote patient engagement, nor how to effectively assess these efforts, a challenge that appears more pronounced in digitally mediated handover contexts. This suggests a gap in informatics-enabled support for nursing practice, particularly in translating digital handover processes into patient-centered behaviors.

Positive impacts of EMR implementation included improved efficiency and reduced interruptions during handover, indicating a degree of user acceptance and adaptation. However, continued reliance on paper-based PFM-generated handover forms highlights persistent gaps in EMR functionality or design, and the need for workarounds to support cognitive and workflow demands during handover. While this study did not examine patient or safety outcomes directly, it provides important process-level evidence to inform the optimization of nursing handover within digitally enabled environments.

Collectively, these findings indicate that the presence of an EMR alone is insufficient to ensure high-quality nursing handover. Rather, the impact of EMR implementation on handover quality is mediated by nursing informatics capability and the degree to which digital systems are designed and aligned with nurses' cognitive processes and clinical workflows. Where EMR interfaces fragment information or fail to provide integrated, task-specific handover views, nurses compensate through workarounds that may limit the potential benefits of digital systems. Optimizing nursing handover in digitally enabled environments requires EMR designs that explicitly support handover as a complex, cognitively demanding, and safety-critical nursing activity, rather than assuming that quality improvements will occur through digitization alone.

Future Research

Future research should examine in greater depth how nurses interact with EMR systems during handover, including patterns of use, information navigation, and reliance on workarounds. Detailed analysis of user interaction with the EMR would assist in determining the degree to which the current technology suits the task being undertaken and which current digital tools are suited to the cognitive and workflow demands of nursing handover.

Qualitative exploration of nurses' perceived barriers and facilitators to EMR use during handover would further inform understanding of how EMR interface design, functionality, workflow integration, and device availability may influence handover practices. Such insights could support the identification of design flaws that contribute to the ongoing reliance on paper-based handover forms and other nonintegrated artifacts in EMR-enabled environments.

Future studies may benefit from the evaluation of discrete components of EMR implementation, including device footprint, software interface design, workflow integration, and overall user experience, to better understand each element's relative impact on the quality and effectiveness of nursing handover.

Comparative analyses across wards or sites with differing levels of digital maturity, including electronic prescribing and medication administration, may offer further clarification as to how specific EMR functionalities influence nursing handover processes and outcomes.

Acknowledgments

We are grateful for the clinical governance expertise, support, and input from members of the Communicating for Safety Clinical Risk Governance Committee in the development of the audit tools used in this case study.

XSL•FO
RenderX

Funding

No funding was received for this research.

Data Availability

The original contributions presented in this study are included in the article and Multimedia Appendix 1. Further inquiries can be directed to the corresponding author.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Eastern Health clinical handover standard.
[DOCX File, 374 KB - nursing_v9i1e85909_app1.docx]

References

1. Slade D, Murray KA, Pun JKH, Eggins S. Nurses' perceptions of mandatory bedside clinical handovers: an Australian hospital study. J Nurs Manag 2019 Jan;27(1):161-171. [doi: 10.1111/jonm.12661] [Medline: 30198619]
2. Chapman YL. Nurse satisfaction with information technology enhanced bedside handoff. Medsurg Nurs 2016;25(5) [FREE Full text]
3. Dunsford J. Structured communication: improving patient safety with SBAR. Nurs Womens Health 2009 Oct;13(5):384-390. [doi: 10.1111/j.1751-486X.2009.01456.x] [Medline: 19821914]
4. Jorm CM, White S, Kaneen T. Clinical handover: critical communications. Med J Aust 2009 Jun 1;190(S11):S108-S109. [doi: 10.5694/j.1326-5377.2009.tb02613.x] [Medline: 19485856]
5. ACo S, QiH C. Safety and Quality Improvement Guide Standard 6: Clinical Handover: ACSQHC Sydney; 2012.
6. Patterson ES, Wears RL. Patient handoffs: standardized and reliable measurement tools remain elusive. Jt Comm J Qual Patient Saf 2010 Feb;36(2):52-61. [doi: 10.1016/s1553-7250(10)36011-9] [Medline: 20180437]
7. Johnson M, Sanchez P, Zheng C. Reducing patient clinical management errors using structured content and electronic nursing handover. J Nurs Care Qual 2016;31(3):245-253. [doi: 10.1097/NCQ.0000000000000167]
8. Ong MS, Coiera E. A systematic review of failures in handoff communication during intrahospital transfers. Jt Comm J Qual Patient Saf 2011 Jun;37(6):274-284. [doi: 10.1016/s1553-7250(11)37035-3] [Medline: 21706987]
9. Yee KC, Wong MC, Turner P. "HAND ME AN ISOBAR": a pilot study of an evidence-based approach to improving shift-to-shift clinical handover. Med J Aust 2009 Jun 1;190(S11):S121-S124. [doi: 10.5694/j.1326-5377.2009.tb02617.x] [Medline: 19485859]
10. Sujan M, Spurgeon P, Cooke M. The role of dynamic trade-offs in creating safety—a qualitative study of handover across care boundaries in emergency care. Reliab Eng Syst Saf 2015 Sep;141:54-62. [doi: 10.1016/j.ress.2015.03.006]
11. Clark JR, Stanton NA, Revell KMA. Identified handover tools and techniques in high-risk domains: using distributed situation awareness theory to inform current practices. Saf Sci 2019 Oct;118:915-924. [doi: 10.1016/j.ssci.2019.06.033]
12. Milesky JL, Baptiste DL, Shelton BK. An observational study of patient handover communications among nurses on an oncology critical care unit. Contemp Nurse 2018 Feb;54(1):77-87. [doi: 10.1080/10376178.2017.1416306] [Medline: 29235419]
13. Haig KM, Sutton S, Whittington J. SBAR: a shared mental model for improving communication between clinicians. Jt Comm J Qual Patient Saf 2006 Mar;32(3):167-175. [doi: 10.1016/s1553-7250(06)32022-3] [Medline: 16617948]
14. Stanton NA, Salmon PM, Walker GH, Salas E, Hancock PA. State-of-science: situation awareness in individuals, teams and systems. Ergonomics 2017 Apr 3;60(4):449-466. [doi: 10.1080/00140139.2017.1278796]
15. National Safety and Quality Health Service Standards, 2nd edition: Australian Commission on Safety and Quality in Health Care; 2021.
16. Chaboyer W, McMurray A, Wallis M. Bedside nursing handover: a case study. Int J Nurs Pract 2010 Feb;16(1):27-34. [doi: 10.1111/j.1440-172X.2009.01809.x] [Medline: 20158545]
17. Forde MF, Coffey A, Hegarty J. Bedside handover at the change of nursing shift: a mixed-methods study. J Clin Nurs 2020 Oct;29(19-20):3731-3742. [doi: 10.1111/jocn.15403] [Medline: 32644255]
18. Ihlebæk HM. Lost in translation - silent reporting and electronic patient records in nursing handovers: An ethnographic study. Int J Nurs Stud 2020 Sep;109:103636. [doi: 10.1016/j.ijnurstu.2020.103636] [Medline: 32593879]
19. Jimma BL, Enyew DB. Barriers to the acceptance of electronic medical records from the perspective of physicians and nurses: a scoping review. Inform Med Unlocked 2022;31:100991. [doi: 10.1016/j.imu.2022.100991]
20. Silow-Carroll S, Edwards JN, Rodin D. Using electronic health records to improve quality and efficiency: the experiences of leading hospitals. Issue Brief (Commonw Fund) 2012 Jul;17(1):1-40 [FREE Full text] [Medline: 22826903]
21. Gatiti P, Ndirangu E, Mwangi J, Mwanzu A, Ramadhani T. Enhancing healthcare quality in hospitals through electronic health records: a systematic review. J Health Inform Dev Ctries 2021;15(2):1 [FREE Full text]

XSL•FO
RenderX

22. Browning L, Raza-Khan U, Leggat S, Boyd JH. The impact of electronic medical record implementation on the process and outcomes of nursing handover: a rapid evidence assessment. J Nurs Manag 2025 Jan;2025(1). [doi: 10.1155/jonm/5585723] [Medline: 5585723]
23. Boonstra A, Versluis A, Vos JFJ. Implementing electronic health records in hospitals: a systematic literature review. BMC Health Serv Res 2014 Sep 4;14(1):370. [doi: 10.1186/1472-6963-14-370] [Medline: 25190184]
24. Staggers N, Clark L, Blaz JW, Kapsandoy S. Why patient summaries in electronic health records do not provide the cognitive support necessary for nurses' handoffs on medical and surgical units: insights from interviews and observations. Health Informatics J 2011 Sep;17(3):209-223. [doi: 10.1177/1460458211405809] [Medline: 21937463]
25. Lindroth HL, Pinevich Y, Barwise AK, et al. Information and data visualization needs among direct care nurses in the intensive care unit. Appl Clin Inform 2022 Oct;13(5):1207-1213. [doi: 10.1055/s-0042-1758735] [Medline: 36577501]
26. Wisner K, Lyndon A, Chesla CA. The electronic health record's impact on nurses' cognitive work: an integrative review. Int J Nurs Stud 2019 Jun;94:74-84. [doi: 10.1016/j.ijnurstu.2019.03.003] [Medline: 30939418]
27. Van Der Meijden MJ, Tange HJ, Troost J, Hasman A. Determinants of success of inpatient clinical information systems: a literature review. J Am Med Inform Assoc 2003;10(3):235-243. [doi: 10.1197/jamia.M1094] [Medline: 12626373]
28. Goodhue DL. Understanding user evaluations of information systems. Manage Sci 1995 Dec;41(12):1827-1844. [doi: 10.1287/mnsc.41.12.1827]
29. Graves JR, Corcoran S. The study of nursing informatics. Image (IN) 1989 Dec;21(4):227-231 [FREE Full text] [doi: 10.1111/j.1547-5069.1989.tb00148.x]
30. Staggers N, Gassert CA, Curran C. A Delphi study to determine informatics competencies for nurses at four levels of practice. Nurs Res 2002;51(6):383-390. [doi: 10.1097/00006199-200211000-00006] [Medline: 12464758]
31. Woods L, Cummings E, Dobroff N, et al. Intended use of the national nursing and midwifery digital health capability framework. In: Healthier Lives, Digitally Enabled: IOS Press; 2021:106-112. [doi: 10.3233/SHTI210018]
32. Darvish A, Bahramnezhad F, Keyhanian S, Navidhamidi M. The role of nursing informatics on promoting quality of health care and the need for appropriate education. Glob J Health Sci 2014 Jun 25;6(6):11-18. [doi: 10.5539/gjhs.v6n6p11] [Medline: 25363114]
33. Mather C, Cummings E, Gale F. Nurses as stakeholders in the adoption of mobile technology in Australian health care environments: interview study. JMIR Nurs 2019;2(1):e14279. [doi: 10.2196/14279] [Medline: 34345771]
34. Yin RK. Case Study Research: Design and Methods: SAGE Publications; 2009, Vol. 5.
35. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing among Five Approaches: Sage publications; 2016.
36. Morgan SJ, Pullon SRH, Macdonald LM, McKinlay EM, Gray BV. Case study observational research: a framework for conducting case study research where observation data are the focus. Qual Health Res 2017 Jun;27(7):1060-1068. [doi: 10.1177/1049732316649160] [Medline: 27217290]
37. Arnold EC, Boggs KU. Interpersonal Relationships E-Book: Professional Communication Skills for Nurses: Elsevier Health Sciences; 2019.
38. Duffy WJ, Kharasch MS, Du H. Point of care documentation impact on the nurse-patient interaction. Nurs Adm Q 2010;34(1):E1-E10. [doi: 10.1097/NAQ.0b013e3181c95ec4] [Medline: 20023554]
39. Misto K, Padula C, Bryand E, Nadeau K. Nurses' perception of the impact of electronic documentation on the nurse-patient relationship. J Nurs Care Qual 2019;34(2):163-168. [doi: 10.1097/NCQ.0000000000000339] [Medline: 29975218]
40. Gaudet CA. Electronic documentation and nurse-patient interaction. ANS Adv Nurs Sci 2016;39(1):3-14. [doi: 10.1097/ANS.0000000000000098] [Medline: 26539694]
41. Ghosh K, Dohan MS, Curl E, Goodwin M, Morrell P, Guidroz P. Information tools for care coordination in patient handover: is an electronic medical record enough to support nurses? Health Care Manage Rev 2022;47(2):100-108. [doi: 10.1097/HMR.0000000000000296]
42. Staggers N, Clark L, Blaz JW, Kapsandoy S. Nurses' information management and use of electronic tools during acute care handoffs. West J Nurs Res 2012 Mar;34(2):153-173. [doi: 10.1177/0193945911407089] [Medline: 21540355]
43. Galatzan BJ, Carrington JM. Exploring the state of the science of the nursing hand-off communication. CIN 2018;36(10):484-493. [doi: 10.1097/CIN.0000000000000461]
44. Gregory S, Tan D, Tilrico M, Edwardson N, Gamm L. Bedside shift reports: what does the evidence say? J Nurs Adm 2014 Oct;44(10):541-545. [doi: 10.1097/NNA.0000000000000115] [Medline: 25208269]
45. Street M, Dempster J, Berry D, et al. Enhancing active patient participation in nursing handover: a mixed methods study. J Clin Nurs 2022 Apr;31(7-8):1016-1029. [doi: 10.1111/jocn.15961] [Medline: 34268829]
46. Tobiano G, Whitty JA, Bucknall T, Chaboyer W. Nurses' perceived barriers to bedside handover and their implication for clinical practice. Worldviews Evid Based Nurs 2017 Oct;14(5):343-349. [doi: 10.1111/wvn.12241] [Medline: 28759162]
47. Drach-Zahavy A, Goldblatt H, Maizel A. Between standardisation and resilience: nurses' emergent risk management strategies during handovers. J Clin Nurs 2015 Feb;24(3-4):592-601. [doi: 10.1111/jocn.12725] [Medline: 25363742]
48. Malfait S, Eeckloo K, Van Biesen W, Van Hecke A. Barriers and facilitators for the use of NURSING bedside handovers: implications for evidence-based practice. Worldviews Evid Based Nurs 2019 Aug;16(4):289-298. [doi: 10.1111/wvn.12386] [Medline: 31309690]

XSL•FO
RenderX

49. Wills MJ, El-Gayar OF, Sarnikar S. Beyond meaningful use: a model for evaluating electronic health record success. 2011 Presented at: 2011 44th Hawaii International Conference on System Sciences. [doi: 10.1109/HICSS.2011.81]
50. Hertzum M, Simonsen J. Positive effects of electronic patient records on three clinical activities. Int J Med Inform 2008 Dec;77(12):809-817. [doi: 10.1016/j.ijmedinf.2008.03.006] [Medline: 18457987]
51. Alghenaimi S. The Role of Electronic Health Records in Structuring Nursing Handoff Communication and Maintaining Situation Awareness: University of Missouri-Columbia; 2012.
52. Nickel N, Amin D, Shakeel F, Germain A, Machry J. Handoff standardization in the neonatal intensive care unit with an EMR-based handoff tool. J Perinatol 2021 Mar;41(3):634-640. [doi: 10.1038/s41372-020-0740-z] [Medline: 32690861]
53. Panda S. Nursing shift handoff process: using an electronic health record tool to improve quality. CJON 2020 Oct 1;24(5):583-585. [doi: 10.1188/20.CJON.583-585]
54. Vawdrey DK, Stein DM, Fred MR, Bostwick SB, Stetson PD. Implementation of a computerized patient handoff application. AMIA Annu Symp Proc 2013;2013:1395-1400. [doi: 10.12691/ajphr-4-4-5] [Medline: 24551415]
55. Bergey MR, Goldsack JC, Robinson EJ. Invisible work and changing roles: Health information technology implementation and reorganization of work practices for the inpatient nursing team. Soc Sci Med 2019 Aug;235:112387. [doi: 10.1016/j.socscimed.2019.112387]
56. Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. BMC Med Inform Decis Mak 2006 Jan 9;6(1):3. [doi: 10.1186/1472-6947-6-3] [Medline: 16401336]
57. Oluoch T, Santas X, Kwaro D, et al. The effect of electronic medical record-based clinical decision support on HIV care in resource-constrained settings: a systematic review. Int J Med Inform 2012 Oct;81(10):e83-e92. [doi: 10.1016/j.ijmedinf.2012.07.010] [Medline: 22921485]
58. Powers EM, Shiffman RN, Melnick ER, Hickner A, Sharifi M. Efficacy and unintended consequences of hard-stop alerts in electronic health record systems: a systematic review. J Am Med Inform Assoc 2018 Nov 1;25(11):1556-1566. [doi: 10.1093/jamia/ocy112] [Medline: 30239810]
59. Todd B, Shinthia N, Nierenberg L, Mansour L, Miller M, Otero R. Impact of electronic medical record alerts on emergency physician workflow and medical management. J Emerg Med 2021 Mar;60(3):390-395. [doi: 10.1016/j.jemermed.2020.10.017] [Medline: 33298357]
60. Gephart S, Carrington JM, Finley B. A systematic review of nurses' experiences with unintended consequences when using the electronic health record. Nurs Adm Q 2015;39(4):345-356. [doi: 10.1097/NAQ.0000000000000119]
61. Lin A, Harris M, Zalis M. Initial observations of electronic medical record usage during CT and MRI interpretation: frequency of use and impact on workflow. AJR Am J Roentgenol 2010 Jul;195(1):188-193. [doi: 10.2214/AJR.09.2946] [Medline: 20566815]
62. Makoul G, Curry RH, Tang PC. The use of electronic medical records: communication patterns in outpatient encounters. J Am Med Inform Assoc 2001;8(6):610-615. [doi: 10.1136/jamia.2001.0080610] [Medline: 11687567]
63. Raban MZ, Westbrook JI. Are interventions to reduce interruptions and errors during medication administration effective?: a systematic review. BMJ Qual Saf 2014 May;23(5):414-421. [doi: 10.1136/bmjqs-2013-002118]
64. Relihan E, O'Brien V, O'Hara S, Silke B. The impact of a set of interventions to reduce interruptions and distractions to nurses during medication administration. Qual Saf Health Care 2010 Oct;19(5):e52. [doi: 10.1136/qshc.2009.036871] [Medline: 20511596]
65. Yoder M, Schadewald D, Dietrich K. The effect of a safe zone on nurse interruptions, distractions, and medication administration errors. J Infus Nurs 2015;38(2):140-151. [doi: 10.1097/NAN.0000000000000095] [Medline: 25723837]
66. Kim JH, Parameshwara N, Guo W, Pasupathy KS. The impact of interrupting nurses on mental workload in emergency departments. Int J Hum Comput Interact 2019 Feb 7;35(3):206-217. [doi: 10.1080/10447318.2018.1447421]
67. Matthews BA, Baker F, Spillers RL. How true is true? assessing socially desirable response bias. Qual Quant 2003 Aug;37(3):327-335. [doi: 10.1023/A:1024420231961]

Abbreviations

CFS CRGC: Communicating for Safety Clinical Risk Governance Committee
EMR: electronic medical record
ISOBAR: Identification Situation Observations Background Assessment/Action Recommendation
NSQHS: National Safety and Quality Health Service
PFM: Patient Flow Manager
WOW: Workstations on Wheels

XSL•FO
RenderX

Edited by G Strudwick; submitted 14.Oct.2025; peer-reviewed by BJ Galatzan, S Manurung; revised version received 09.Mar.2026; accepted 22.Mar.2026; published 12.May.2026.

Please cite as:
Browning L, Raza Khan U, Leggat S, Monypenny N, Boyd JH
Exploring the Impact of Electronic Medical Record–Enabled Versus Paper-Based Systems on the Quality of Nursing Handover: Comparative Case-Study
JMIR Nursing 2026;9:e85909
URL: https://nursing.jmir.org/2026/1/e85909
doi:10.2196/85909

© Lisa Browning, Urooj Raza Khan, Sandra Leggat, Nicholas Monypenny, James H Boyd. Originally published in JMIR Nursing (https://nursing.jmir.org), 12.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of a Digital Decision Support System on Diabetes Management in a Real-World Home Care Setting: Prospective Quasi-Experimental Feasibility and Implementation Study

Katharina Maria Lichtenegger[1], BSc, MSc, PhD; Lara Scholle-Winkler[1], MD; Julia Kopanz[1], BSc, MSc, PhD; Angela Libiseller[1], BSc, MSc; Klaus Donsa[2], DI, PhD; Felix Aberer[1], MD, PhD; Thomas Rudolf Pieber[1,2], MD; Julia Katharina Mader[1], MD

[1]Division of Endocrinology and Diabetology, Medical University of Graz, Auenbruggerplatz 15, Graz, Styria, Austria

[2]Joanneum Research Forschungsgesellschaft mbH, Graz, Styria, Austria

Corresponding Author:
Katharina Maria Lichtenegger, BSc, MSc, PhD
Division of Endocrinology and Diabetology, Medical University of Graz, Auenbruggerplatz 15, Graz, Styria, Austria

Abstract

Background: Older adults with diabetes are vulnerable, facing multimorbidity and challenges in reaching glycemic targets. When insulin therapy is required, limited knowledge of geriatric therapeutic goals often leads to regimens that increase the risk of hypoglycemia. Simplified approaches such as once-daily basal insulin are recommended, yet titration protocols and unclear regimens remain barriers in home care nursing. Digital decision support systems (DDSSs) such as GlucoTab address this gap by offering evidence-based titration guidance integrated into nursing workflows. DDSSs also strengthen nurse autonomy, which leads to improved care processes and better outcomes in home care settings.

Objective: This study assessed the effects of a nurse-guided intervention including a DDSS for diabetes management in older adults with type 2 diabetes receiving home care nursing.

Methods: The prospective quasi-experimental feasibility and implementation study was conducted using GlucoTab for diabetes management as part of a multicomponent intervention at 2 home care nursing sites. Nine adults with type 2 diabetes receiving insulin therapy were enrolled (n=5 women; mean age 77, SD 10 years; mean BMI 27.7, SD 4.7 kg/m^2; mean baseline hemoglobin A_{1c} 59.5, SD 12.8 mmol/mol). After receiving training, nurses implemented the DDSS, which integrates evidence-based diabetes management recommendations for home care and nursing workflows. Diabetes-related hospital admissions and glycemic outcomes were assessed before, during, and after implementation of the DDSS intervention.

Results: Before implementation, 5 diabetes-related emergency admissions led to 4 hospitalizations, compared with none during implementation and 4 after implementation, mostly due to hyperglycemia. The median hospital stay was 9 (IQR 8-15) days (before and after implementation). Diabetes management shifted from predominantly premixed insulin therapy at baseline to basal insulin with or without bolus insulin during implementation. Diabetes-related home care visits decreased from 1.81 to 1.01 per day (before vs during implementation), and general practitioner contacts for insulin adjustments declined (19 before implementation vs 16 after implementation). Mean morning blood glucose decreased from 212 (SD 72) mg/dL to 162 (SD 42) mg/dL, and the proportion of measurements within the target range (70 - 180 mg/dL) increased from 39.0% to 70.8% before vs during implementation. No incidence of hypoglycemia requiring third-party help was observed during implementation. Adherence of nurses to DDSS recommendations exceeded 94%.

Conclusions: This study suggests that integrating a DDSS in a multicomponent intervention into home care nursing yielded longitudinal changes in diabetes outcomes. Emergency visits and hospitalizations were lower during implementation, no severe hypoglycemia was observed, glycemic control improved throughout the observation period, and therapy was simplified in routine care. Nursing visits decreased, suggesting potential improvements in care processes and resource use. High adherence and acceptance by nurses indicate feasibility and usability in home care practice. These findings align with evidence from hospitalized patients and telehealth studies, suggesting the potential for clinical, safety, and efficiency benefits.

Trial Registration: German Clinical Trials Register DRKS00015059; https://drks.de/search/en/trial/DRKS00015059

(JMIR Nursing 2026;9:e89687) doi:10.2196/89687

KEYWORDS

diabetes management; diabetes mellitus; digital decision support system; home care service; implementation science; digital health

XSL•FO
RenderX

Introduction

Background

Worldwide, approximately 422 million people live with diabetes, most residing in low- and middle-income countries. The prevalence of diabetes has been steadily increasing across all income levels over recent decades [1]. The most common form is type 2 diabetes, accounting for more than 90% of all cases worldwide, with prevalence rising further in geriatric populations due to increasing life expectancy. In this context, Europe represents one of the regions with the highest diabetes prevalence worldwide, with a substantial and growing proportion of older adults affected, placing increasing pressure on health care systems and home-based care services [2].

Complexity of Diabetes Management

Diabetes is a chronic, metabolic disorder characterized by elevated blood glucose levels and can be accompanied by a range of complications, including microvascular and macrovascular damage, as well as a multitude of common comorbidities (eg, hypertension, obesity, cardiovascular disease, and chronic kidney disease) [3]. Older adults in particular face an increased risk of adverse health outcomes due to comorbidities and often fail to achieve age-appropriate glycemic goals [4]. Clinical evidence confirms that effective blood glucose management can reduce diabetes-related complications [5]. The American Diabetes Association recommends simplified treatment regimens for older adults, emphasizing hypoglycemia prevention as a critical factor of individualized glycemic targets. With the aging population, more older adults will live with diabetes, and age-related declines in self-management skills and memory may further complicate care. Basal insulin regimens, often administered via prefilled pens once daily, are especially suitable for older adults due to their ease of use and low hypoglycemia risk. Despite the availability of modern antihyperglycemic agents, some individuals with type 2 diabetes will ultimately require and benefit from insulin therapy due to decreasing endogenous insulin secretion. Evidence from hospitalized patients with type 2 diabetes has shown that a basal-bolus regimen is clearly superior to premixed insulin, resulting in fewer hypoglycemic events at comparable mean blood glucose levels [5,6].

Empowering Nurses in Diabetes Care

Older adults with diabetes in home care nursing frequently do not receive optimal diabetes management due to lack of information, support, and professional cohesion [7,8]. Effective chronic diabetes management demands an interdisciplinary approach in which nurses represent a leading, empowered player. Enhancing nurses' clinical autonomy, confidence, access to up-to-date information, and appropriate training is essential to enable this role [9]. Nurse-led care models have consistently shown to improve the outcomes of people with chronic diseases such as diabetes. Outcomes also include improved quality of care, better chronic disease management, and reduced health care use in primary care settings [10]. Evidence indicates that nurses often spend a great deal of time engaging with individuals, enabling more comprehensive assessment and support [11]. Systematic reviews further show that nurse-led interventions are associated with higher patient satisfaction; reductions in hospitalizations; and, in some cases, even lower mortality [12]. Additionally, care delivered by well-trained nurses can be as effective or even superior to that provided by primary care physicians [13].

Digital Decision Support Systems

Innovative point-of-care digital decision support systems (DDSSs) are increasingly recognized as useful tools to facilitate evidence-based care specifically in people of older age and with impaired health status and more pronounced comorbidity profiles [14,15]. Evidence from randomized controlled trials has demonstrated that technology-supported and digitally enabled care interventions can improve clinical outcomes in chronic diseases, including diabetes [16]. Similarly, telehealth and structured chronic disease management approaches have been shown to improve care processes and reduce health care use [17]. Developments in glycemic care include computerized order sets, blood glucose data networks with visualization, case-finding tools, insulin calculators, electronic treatment protocols, titration algorithms, and alert systems [18,19]. For example, the DDSS GlucoTab supports basal-bolus insulin titration workflows by structuring care processes and offering standardized recommendations in the hospital setting. While such systems are often well established in hospitals [20-22], it remains unclear whether these tools can be effectively used in home care nursing. Although evidence from hospital and telehealth settings supports the effectiveness of such digital interventions, their transferability to home care nursing environments remains underexplored [17]. Each tool must be tested in its new care setting as it cannot be assumed that the systems designed for hospitals will function seamlessly in home care nursing.

To address this gap, DDSSs designed for nursing care are gaining more and more attention. A recent review highlights the importance of integrating nurses into every stage of innovation design and implementation to preserve care-centric perspectives [23]. Current desktop tools and mobile apps for the nursing care of older adults use nursing process frameworks to personalize interventions, showing a trend toward smart, home-based support systems [24]. These findings align with broader themes of nurse empowerment and digitally supported, person-centered care.

Objectives

This study aimed to assess the effects of a multicomponent intervention including a DDSS on diabetes management by assessing diabetes-related glycemic management outcomes and hospital admissions in older adults with type 2 diabetes receiving home care nursing. This study was guided by the following research questions: (1) How does the implementation of a nurse-led DDSS affect glycemic outcomes across pre-, during, and postimplementation phases? (2) How does the DDSS influence diabetes-related hospital admissions in older home-cared adults? Outcomes were examined over 3 phases: 6 months before implementation, 3 months during implementation, and 6 months after implementation of the nurse-guided DDSS for insulin therapy.

XSL•FO
RenderX

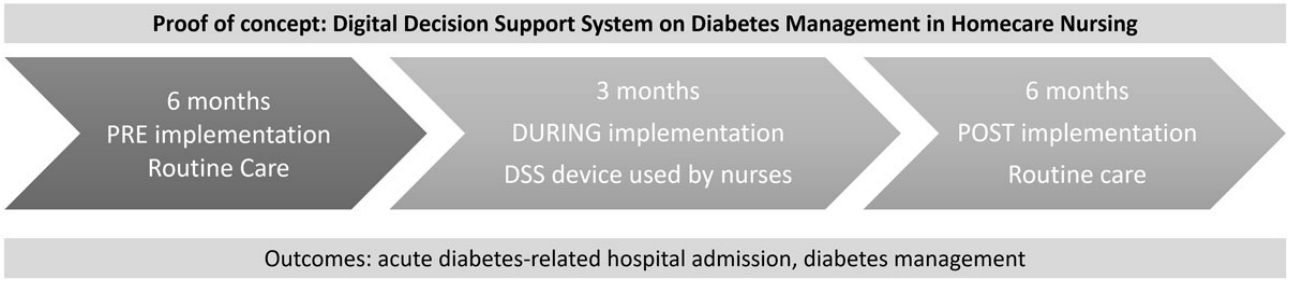
Methods

Overview

This study was conducted and reported as a prospective quasi-experimental feasibility and implementation study in a real-world setting. Reporting followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines where applicable given the nonrandomized observational design. Relevant elements from the CONSORT-AI (Consolidated Standards of Reporting Trials–AI) extension were considered for the description of the DDSS [25]; however, the CONSORT-AI was not fully applicable due to the nonrandomized, nontrial design of this study. Furthermore, GlucoTab (decide Clinical Software GmbH) is a rule-based clinical decision support system and not a machine learning–based AI system.

Following the overall study design, assessments were conducted at 3 predefined time points: before implementation (6 months), during implementation (implementation phase with active DDSS use; 3 months), and after implementation (postimplementation follow-up phase after DDSS discontinuation; 6 months; Figure 1). This study was conducted in Graz, Austria, at 2 home care nursing sites of the Austrian Red Cross, a private, nonprofit organization.

Figure 1. Study design to assess outcomes in the pre-, during, and postimplementation phases of a digital decision support system (DDSS) for diabetes management in home care nursing.



Participants

A total of 9 individuals with type 2 diabetes were enrolled in the study. All 9 participants were assessed before implementation and during implementation. Only 8 participants were assessed after implementation as 1 participant died in the postimplementation phase.

Eligible participants were adults with type 2 diabetes treated with insulin therapy who received home care nursing from the Austrian Red Cross. Recruitment took place at 2 Austrian Red Cross home care sites. Local home care teams identified patients who met the predefined eligibility criteria and informed the study team about potentially eligible individuals. The most important exclusion criteria were medical conditions interfering with study participation or safety, other forms of diabetes (eg, type 1 diabetes), mental incapacity to provide consent, or terminal illness.

Study Intervention

The intervention involved a multicomponent restructuring of care delivery, including structured nursing training, a DDSS, physician-guided initiation of insulin therapy, and standardized insulin titration protocols. The DDSS used (GlucoTab) included decision support for health care professionals regarding automated suggestions for insulin dose, adjustments for basal insulin therapy, and changes in the insulin regimen. The DDSS also included suggestions for blood glucose measurement frequency, workflow support through visualization of open tasks, and support with a documentation interface. In a previous study, the basal insulin algorithm was specifically adapted for geriatric individuals [26]. An evidence-based version developed in line with international guidelines [3] was digitally implemented integrated into the GlucoTab software, with the system and workflows tailored to the routines of home care nursing. Importantly, the system empowers nurses to act directly in the individual's home based on current blood glucose values independently of physician availability, including evenings and weekends, when medical support is often limited.

Before implementation, home care nurses and supervising study physicians participated in a structured group training covering the intervention.

During the implementation phase of the intervention, the study physicians prescribed the initial insulin regimen, and subsequent dosing was managed by home care nurses using GlucoTab. The DDSS allows physicians to either independently determine the initial basal insulin dose for participants (already receiving insulin therapy) or request an algorithm-generated dosing recommendation for participants newly initiated on insulin therapy. In both cases, the final decision on initial insulin therapy remains physician led. GlucoTab automatically flags unsafe situations (eg, low blood glucose values) and ensures physician review when necessary. The DDSS automatically records all generated recommendations and corresponding nursing actions, enabling objective and reproducible calculation of adherence rates [20-22]. The system uses a mobile client-server architecture connected via a secure Long Term Evolution network to a centralized, secure server environment hosted on the premises of Joanneum Research Forschungsgesellschaft mbH (Austria).

During the implementation phase, nurses accessed GlucoTab via standard tablet devices using individual secure log-ins during home care visits. The system does not require dedicated medical hardware and was used directly at the point of care to document blood glucose values, insulin administration, and relevant clinical information. On the basis of the entered data, GlucoTab generated individualized, algorithm-based decision support recommendations for insulin management. Nurses reviewed

XSL•FO
RenderX

these recommendations and integrated them into their clinical decision-making process while retaining responsibility for the final insulin dosing and administration decision.

During the pre- and postimplementation phases, participants received routine care, which encompassed paper-based diabetes management, including diabetes therapy, and the frequency of blood glucose measurements was prescribed by the treating general practitioner. Home care nurses performed blood glucose monitoring and administered diabetes therapy as prescribed. Importantly, the postimplementation phase represents routine care after DDSS discontinuation rather than a genuine return to baseline conditions as training effects and workflow changes established during implementation may persist into the postimplementation phase.

Assessments and Data Analysis

In all phases, clinical and demographic data and diabetes management data (blood glucose values, measurement frequency, insulin dosing and injection, general practitioner contacts, and diabetes-related hospital admissions) were collected from (electronic) health records and paper-based nursing documentation of the home care nursing provider.

Hypoglycemia was defined as capillary blood glucose values below 54 mg/dL and/or hypoglycemic events requiring third-party assistance. Hypoglycemic events were identified from nursing documentation and blood glucose records. Continuous glucose monitoring was not available in this study.

During the implementation phase, additional data regarding diabetes management were recorded directly in the DDSS for nurses. Clinical and demographic data were digitally entered into electronic case report forms (OpenClinica; OpenClinica, LLC) and spreadsheets by a qualified researcher. Quality control was ensured by applying a 4-eye principle to verify completeness and accuracy.

Nurse acceptance was operationalized as adherence to DDSS-generated recommendations and was objectively measured using automatically logged system data capturing nurse responses and actions within the DDSS.

Data analysis was performed using the statistical program SPSS Statistics (version 26; IBM Corp) and Microsoft Excel. Data were first summarized using descriptive statistics. For numerical data, depending on the distribution, means and SDs; medians; minimum, maximum, and range; and quartiles were calculated. Categorical data are presented as absolute and relative frequencies. Due to the quasi-experimental feasibility and implementation design of the study, no inferential statistical analyses and no sample size calculation were conducted.

Demographics

Nine participants receiving home care nursing were included (n=5 women; mean age 77, SD 10 years; mean hemoglobin A_{1c} [HbA_{1c}] 59.5, SD 12.8 mmol/mol; mean BMI 27.7, SD 4.7 kg/m^2 [before implementation]). Before implementation, 4 participants had diabetes-related comorbidities: retinopathy (n=2), neuropathy (n=3), and nephropathy (n=3).

The 9 participants completed all 3 study phases as per the protocol. One participant was hospitalized (unrelated to the study) and withdrew but was included in the analysis as intervention duration was completed. One participant died during the postimplementation phase for reasons unrelated to the study.

Ethical Considerations

This study was approved by the Ethics Committee of the Medical University of Graz (approval numbers EK 30 - 287 ex 17/18 and EK 31 - 522 ex 18/19) and relevant legal authorities. It was conducted in compliance with good clinical practice and the principles of the Declaration of Helsinki.

All eligible patients identified were invited to participate; all provided informed consent and were enrolled in the study. To ensure confidentiality, participants were assigned unique study identification numbers, preventing traceability of personal data. No participant compensation was provided.

Results

Diabetes-Related Hospital Admissions

During the preimplementation phase, 5 diabetes-related emergency admissions were documented in 4 participants, resulting in 4 hospitalizations. In the postimplementation phase (after DDSS discontinuation), 4 diabetes-related emergency admissions occurred in 2 participants, leading to 1 referral to the diabetes outpatient clinic and 2 hospitalizations. No diabetes-related emergency admissions or hospitalizations were observed during the implementation phase (Table 1).

XSL•FO
RenderX

Table . Diabetes management and outcomes before, during, and after implementation of the GlucoTab system.

	Routine care (6 months before implementation; n=9)	Decision support (3 months during implementation; n=9)	Routine care (6 months after implementation; n=8[a])
Insulin therapy, n (%)			
Premixed insulin	5 (55.6)	0 (0)	0 (0)
Premixed insulin+bolus	2 (22.2)	0 (0)	0 (0)
Basal insulin	1 (11.1)	9 (100)	5 (62.5)
Basal insulin+bolus[b]	1 (11.1)	0 (0)	3 (37.5)
Hospital admissions, n	5	0	4
Inpatient stays among admissions	4	0	2
Morning capillary blood glucose (mg/dL), mean (SD)	212 (72)	Month 1: 171 (68); month 2: 150 (46); month 3: 145 (35)	162 (42)
Blood glucose values by range (mg/dL), % (n/N)[c]			
<54	0.1	0.0 (0/992)	0.0
<70	0.2	0.3 (3/992)	0.4
70-180	39.0	77.0 (764/992)	70.8
>180	60.8	22.1 (219/ 992)	28.8
HbA_{1c}[d] (mmol/mol), mean (SD)	59.5 (12.8)	56.5 (12.4)	—[e]

[a]One participant died during the postimplementation phase.

[b]Bolus insulin for main meal or for correctional insulin.

[c]Absolute frequencies (n/N) cannot be provided for before and after implementation because the original dataset is currently unavailable.

[d]HbA_{1c}: hemoglobin A_{1c}.

[e]HbA_{1c} was not available in the postimplementation phase.

Hyperglycemia was the reason for referral in 8 of the 9 diabetes-related emergency admissions; the remaining diabetes-related emergency admission was due to hyperglycemia. The length of inpatient stays ranged from 8 to 15 days, with a median of 9 days.

Nurse-Led Diabetes Processes in Home Care

Table 1 summarizes diabetes management and outcomes across the 3 study phases. In the preimplementation phase, 77.8% (7/9) of the participants received premixed insulin, including 22.2% (2/9) with additional bolus insulin. A total of 22.2% (2/9) of the participants received basal insulin, of whom 1 also received bolus insulin. After study completion, all participants (8/8, 100%) were treated with basal insulin, with 37.5% (3/8) additionally receiving bolus insulin for correcting hyperglycemia. In 62.5% (5/8) of the participants, the suggested basal insulin titrations were implemented by general practitioners; in 37.5% (3/8) of the participants, the recommended corrections were not adopted during the postimplementation phase.

During the implementation phase, therapy adjustments based on DDSS recommendations were made on average 9 times in the 3 months (up to 18 times). The number of diabetes-related home care nursing visits decreased from 1.81 to 1.01 per day when comparing the pre- and during implementation phases, largely due to reduced necessity of blood glucose checks.

In total, 35 diabetes-related contacts with general practitioners were documented, of which 19 (54.3%) occurred in the 6-month preimplementation phase and 16 (45.7%) occurred in the postimplementation phase, indicating a slight reduction after the implementation phase. In the preimplementation phase, 36.8% (7/19) of the general practitioner contacts were due to a planned checkup, 57.9% (11/19) were for therapy adjustment, and 5.3% (1/19) were for HbA_{1c} determination. In the postimplementation phase, 31.3% (5/16) of the general practitioner contacts were due to planned controls, and 68.8% (11/16) were due to therapy adjustment. During the implementation phase, nurses contacted the outpatient diabetes clinic via telephone as suggested by the DDSS for receiving medical advice, but no quantitative data were systematically recorded in this phase.

Glycemic Outcomes

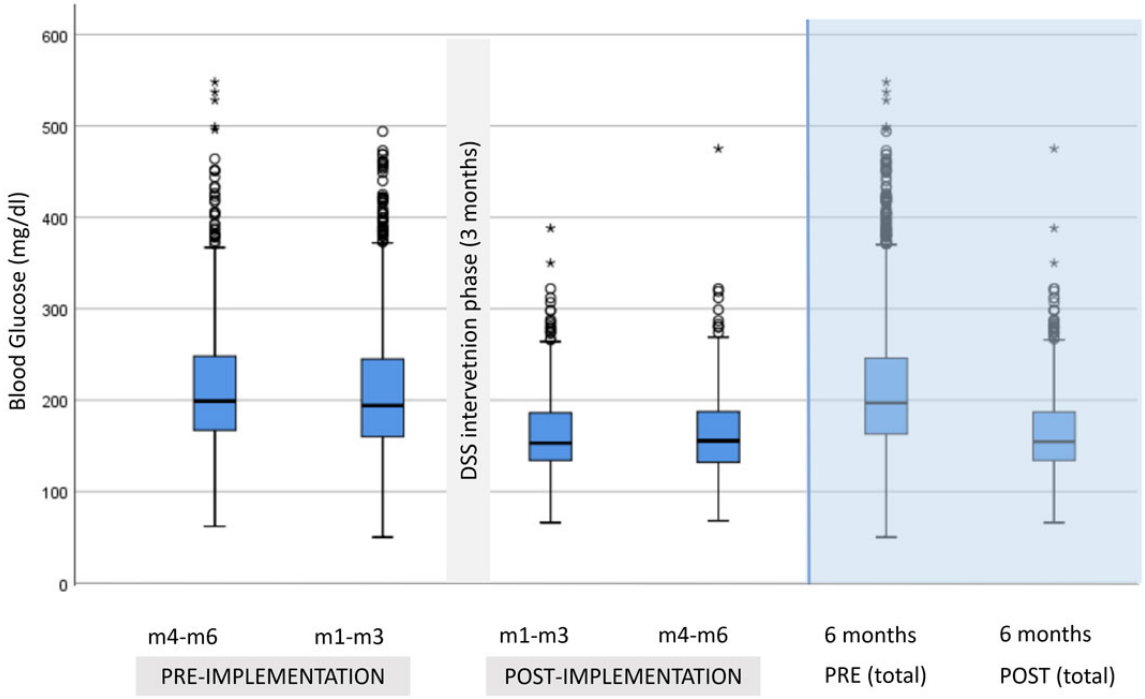
Mean morning blood glucose values decreased from 212 (SD 72) mg/dL before implementation to 162 (SD 42) mg/dL after implementation. No hypoglycemic events requiring third-party assistance were observed during any study phase. Biochemical hypoglycemia (<54 mg/dL) was rare, with 0.1% of glucose measurements below this threshold in the preimplementation phase and no values below 54 mg/dL during implementation or after implementation. The proportion of blood glucose measurements within the prespecified target range (70 - 180 mg/dL) was 39.0% before implementation, 77% (764/992) during implementation, and 70.8% after implementation,

XSL•FO
RenderX

whereas measurements above 180 mg/dL were 60.8%, 22.1% (219/992), and 28.8%, respectively (n/N values were only available for the during-implementation period).

The distribution of blood glucose values over time showed a clear trend toward improved glycemic control (Figure 2, right panel). Higher mean values and wider variability were observed in the preimplementation phase. In contrast, a marked reduction in median blood glucose levels and variability was found in the implementation phase. This improvement was sustained during the following 3 months, with the 6-month postimplementation phase showing improved glycemic control compared to the 6-month preimplementation phase.

Figure 2. Box plots for blood glucose values in the pre- and postimplementation phases (boxes represent the IQR with the median, whiskers indicate 1.5 times the IQR, dots indicate outliers [1.5 - 3 times the IQR], and stars indicate extreme outliers [>3 times the IQR]). The figure shows longitudinal glucose measurements over time, with repeated measurements from the same participants reflecting real-world glucose monitoring. DSS: decision support system.



User Acceptance

In 88.9% (8/9) of cases, physicians chose to set the initial basal insulin dose independently without using the DDSS-generated recommendation (all participants already received insulin therapy before GlucoTab implementation). Adherence to the DDSS recommendations by nurses was high: 94.7% (552/583) for blood glucose measurement frequency, 95.8% (295/308) for basal insulin dose adjustment, 99.7% (764/766) for basal insulin injections, 97.9% (142/145) for bolus insulin injections, and 100% (231/231) for titration time points.

Discussion

Principal Findings

This prospective quasi-experimental feasibility and implementation study suggests that integration of a DDSS into home care nursing corresponded with improved diabetes management outcomes over time. During the implementation of the DDSS, emergency department visits and hospitalizations were reduced, highlighting improved patient safety and favorable clinical outcomes. Glycemic control improved over time, while therapy appeared simplified and nursing visits decreased, suggesting potential improvements in efficiency and resource use.

Diabetes management in home care or long-term care facilities is often challenged by the complexity of insulin regimens used in older adults, the high variability of daily blood glucose levels, and the frequent need for medical decisions. Many older adults with diabetes receive once- or twice-daily basal or premixed insulin, which still requires regular dose adjustments [8,27]. However, in clinical practice, contacts with general practitioners occur infrequently, dose adaptations are often delayed, and nurses are frequently forced to act autonomously without formal medical orders. As a consequence, clinical uncertainties can lead to avoidable emergency department transfers or hospital admissions [8]. DDSSs directly address these challenges by providing structured, algorithm-guided titration that supports timely and safe decisions at the point of care.

Our study findings extend the evidence base for algorithm-guided diabetes management using GlucoTab as initially demonstrated in hospitalized patients [20-22], where the decision support system guided insulin dosing, achieving effective glycemic control with reduced hypoglycemia and appropriate adherence to dose suggestions. A recent comprehensive review emphasizes that technological solutions, including connected insulin pens, continuous glucose monitoring, automated insulin delivery systems, and other digital tools, play an increasingly important role in the care of older adults with diabetes. Through reduction of blood glucose

XSL•FO
RenderX

variability and hypoglycemia risk, as well as by improving quality of life and reducing hospital admissions, such technologies support safe and effective long-term diabetes management in older, multimorbid populations [15]. This broad evidence for diabetes technology in older adults reinforces the relevance and potential impact of GlucoTab in home care nursing.

Beyond home care, randomized controlled trials have shown that nurse-assisted insulin titration can safely improve glycemic control. For example, decision-supported telephone coaching by diabetes specialty nurses enabled individuals to self-titrate basal insulin safely, resulting in greater HbA_{1c} reductions without increasing hypoglycemia compared to usual care [28]. This previous study underscores that combining decision support with nurse engagement meaningfully improves glycemic outcomes in noninstitutionalized individuals, supporting our findings from this quasi-experimental feasibility and implementation study in home care nursing.

The high acceptance levels that we observed among nurses, with adherence rates exceeding 94%, align with nursing-focused digital health literature. These findings are in line with evidence that nurse-led care models improve chronic disease outcomes and strengthen care quality in primary care settings [10]. Recent studies published in *JMIR Nursing* and *JMIR Diabetes* underscore the importance of implementing user-friendly digital tools in fostering adherence and enabling nurse-led care innovations [29,30]. Usability is a key concept in the evaluation of digital health technologies and can be defined as the extent to which a system enables users to achieve specified goals with effectiveness, efficiency, and satisfaction. Importantly, usability is influenced by the users, tasks, and context of use, which is particularly relevant in home care settings where clinical decision-making occurs in complex real-world environments [31]. Established usability evaluation methods include inspection-based approaches and user-centered evaluation techniques that are used to identify usability problems and barriers to system use during development and implementation [32]. However, formal usability testing was not conducted in the present study. Recent digital health research further emphasizes that successful implementation of health technologies requires the combined evaluation of usability, user experience, and implementation outcomes as these factors jointly influence adoption, safety, and long-term sustainability [33]. This supports the notion that the DDSS not only augments clinical outcomes but also empowers nursing autonomy and confidence, allowing nurses to integrate the DDSS seamlessly into daily routines and care activities and spend more time on direct patient care rather than administrative tasks or frequent coordination with general practitioners.

The observed reduction in home nursing visits and the simplification of therapy processes through GlucoTab suggest considerable potential for cost savings and efficiency gains in home care nursing services. At the same time, the DDSS strengthens nurses' autonomy and supports safe, evidence-based decision-making without continuous physician oversight. This is in line with international guidelines advocating for structured nurse-led insulin titration under standardized protocols (eg, the Irish national insulin titration guideline for nurses) [34]. These effects align with broader health system goals of optimizing resource use without compromising quality, a particularly relevant consideration in international settings with varying home care constraints. Similar findings have been reported in other studies. A systematic review demonstrated that telehealth interventions for communication between individuals with type 2 diabetes and health care professionals significantly improved glycemic control, with 80% of included studies also showing potential for cost reduction [35]. Another study reported a reduction in overall health care costs of approximately US $24,000 over 7 months through telehealth-based diabetes management, highlighting substantial savings for the health care system [36].

Recent advances in AI and digital health research further contextualize these findings. Emerging AI-enabled decision support prototypes for basal insulin titration have shown promising usability and safety in hospital and early primary care evaluations, indicating that nurse-in-the-loop models could further enhance precision and individualization of insulin therapy in home care settings. Integrating such tools may help reduce practice variation, strengthen guideline adherence, and potentially open pathways for decision support in broader chronic disease management beyond insulin therapy [24,37-39].

These findings suggest the potential of digital interventions such as GlucoTab not only to improve clinical outcomes but also to enhance efficiency and reduce costs in home care nursing services by minimizing the need for frequent coordination with general practitioners, enabling nurses to work more autonomously with decision support, and reducing dependence on continuous medical orders.

Limitations and Future Directions

The relatively small sample size and observational design may limit the generalizability of our findings. The observational pretest-during-posttest design without a control group limits causal inference. All outcomes should be interpreted as descriptive and hypothesis generating rather than causal evidence of effectiveness. A key limitation is that the postimplementation phase does not reflect a return to baseline as participants and care processes may have been influenced by DDSS implementation, structured training, and persistent workflow adaptations. Pretest-posttest comparisons should thus not be interpreted as changes between fully independent states.

The variation in the lengths of the study phases (6-3-6 months) and the limited external validity due to participation of only two study sites in a single region also constrain the transferability to other settings. Long-term sustainability of the observed improvements, therefore, requires confirmation through larger, ideally randomized studies conducted across diverse health care systems.

Future research should focus on (1) evaluating cost-effectiveness through detailed economic analyses; (2) testing scalability and adaptation of GlucoTab in varied international home care models; and (3) exploring the effects of DDSSs on broader populations, including non–insulin-dependent individuals with type 2 diabetes or other chronic diseases and further

XSL•FO
RenderX

patient-centered clinical end points (eg, mortality and disability). Furthermore, glycemic control might be more intensively supervised by using (blinded) continuous glucose monitoring systems in future studies to obtain a broader picture of the efficacy of such a DDSS.

Relevance to Clinical Practice for the Wider Global Clinical Community

This study provides in-depth insights into the quality of diabetes management for people receiving home care nursing, a population that is often underrepresented in research. It generates evidence to guide policymakers and nursing care managers in designing strategies that ensure high-quality diabetes care in home care settings by strengthening the role and competencies of nursing staff. Furthermore, it presents an exemplary digital decision support solution that empowers nurses, enhances the quality of care, and promotes interdisciplinary collaboration with general practitioners, thereby offering a transferable model for other health care systems.

Conclusions

In the long term, DDSSs have the potential not only to reduce costs in the home care setting but also to enhance the evidence-based quality and safety of care for people with type 2 diabetes. This quasi-experimental feasibility and implementation study provides the first evidence that a nurse-led DDSS such as GlucoTab can safely and effectively improve glycemic control in older adults receiving home care nursing, reduce hospitalizations, simplify therapy, and enhance nurse autonomy. Given its feasibility, this approach could empower nurse-led diabetes management, optimize resource use, and improve outcomes for people with diabetes in diverse health care settings. Collectively, these findings highlight the potential of digitally supported, nurse-driven interventions as scalable, patient-centered solutions for managing complex diabetes care in community and long-term care environments.

Acknowledgments

The authors thank all the collaborating health care professionals from the Austrian Red Cross Graz for their cooperation and effort. All authors declared that they had insufficient funding to support open-access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this article. Parts of the results on the implementation phase were published in *Frontiers in Clinical Diabetes and Healthcare* [40]. The authors used the generative AI tool ChatGPT (version 4) by OpenAI to revise the manuscript for publication, which was further reviewed and revised by the authors and study group.

Funding

Parts of the published results were funded by Zukunftsfonds Steiermark (project ABT08-183051/2016-22).

Conflicts of Interest

JKM, TRP, and KD are founders of decide Clinical Software GmbH. JKM is a member of advisory boards for Abbott Diabetes Care, Becton Dickinson, Biomea Fusion, Dexcom, Eli Lilly and Company, Embecta, Insulet, Medtronic, mylife Diabetes Care AG, Novo Nordisk A/S, PharmaSens, Roche Diabetes Care, Sanofi-Aventis, Tandem, and Viatris and has received speaker honoraria from A. Menarini Diagnostics, Abbott Diabetes Care, Buzud, Dexcom, Eli Lilly and Company, MedTrust, MSD, Novo Nordisk A/S, Roche Diabetes Care, Sanofi, Sinocare, Viatris, and Ypsomed. She is a shareholder of decide Clinical Software GmbH and elyte diagnostics GmbH and serves as chief medical officer of elyte diagnostics GmbH. FA has received speaker honoraria from Sanofi-Aventis, Novo Nordisk A/S, Amgen, and Eli Lilly and Company and was a member of advisory boards for Bayer, Novo Nordisk A/S, and Eli Lilly and Company. TRP is an advisory board member for Novo Nordisk A/S; consultant for Roche Diabetes Care, Novo Nordisk A/S, Eli Lilly and Company, Infineon Technologies, and Carnegie Investment Bank; and shareholder of decide Clinical Software GmbH and is on the speaker's bureau of Novo Nordisk A/S and AstraZeneca. All other authors declare no other conflicts of interest.

References

1. Diabetes. World Health Organization. URL: https://www.who.int/health-topics/diabetes#tab=tab_1 [accessed 2026-07-06]
2. IDF Diabetes Atlas, 11th edition: International Diabetes Federation; 2025.
3. American Diabetes Association Professional Practice Committee for Diabetes. 13. Older adults: Standards of Care in Diabetes-2026. Diabetes Care 2026 Jan 1;49(Suppl 1):S277-S296. [doi: 10.2337/dc26-S013] [Medline: 41358888]
4. Alam W, Syamala S, Al Hamad H, et al. Improving monitoring of diabetic complications in home care patients. BMJ Open Qual 2017;6(2):e000053. [doi: 10.1136/bmjoq-2017-000053] [Medline: 29450271]
5. American Diabetes Association Professional Practice Committee for Diabetes. 16. Diabetes care in the hospital: Standards of Care in Diabetes-2026. Diabetes Care 2026 Jan 1;49(Supplement_1):S339-S355. [doi: 10.2337/dc26-S016] [Medline: 41358892]
6. Bellido V, Suarez L, Rodriguez MG, et al. Comparison of basal-bolus and premixed insulin regimens in hospitalized patients with type 2 diabetes. Diabetes Care 2015 Dec;38(12):2211-2216. [doi: 10.2337/dc15-0160] [Medline: 26459273]

XSL•FO
RenderX

7. Hermann M, Heimro LS, Haugstvedt A, Hernar I, Sigurdardottir AK, Graue M. Hypoglycaemia in older home-dwelling people with diabetes- a scoping review. BMC Geriatr 2021 Jan 7;21(1):20. [doi: 10.1186/s12877-020-01961-6] [Medline: 33413148]
8. Heimro LS, Hermann M, Davies TT, Haugstvedt A, Haltbakk J, Graue M. Documented diabetes care among older people receiving home care services: a cross-sectional study. BMC Endocr Disord 2021 Mar 10;21(1):46. [doi: 10.1186/s12902-021-00713-w] [Medline: 33691687]
9. Dufour E, Bolduc J, Leclerc-Loiselle J, et al. Examining nursing processes in primary care settings using the Chronic Care Model: an umbrella review. BMC Prim Care 2023 Sep 4;24(1):176. [doi: 10.1186/s12875-023-02089-3] [Medline: 37661248]
10. O'Connor PJ, Sperl-Hillen JM, Rush WA, et al. Impact of electronic health record clinical decision support on diabetes care: a randomized trial. Ann Fam Med 2011;9(1):12-21. [doi: 10.1370/afm.1196] [Medline: 21242556]
11. Uğur E, Demir H, Akbal E. Postgraduate education needs of nurses' who are caregivers for patients with diabetes. Pak J Med Sci 2015;31(3):637-642. [doi: 10.12669/pjms.313.6732] [Medline: 26150859]
12. Martínez-González NA, Djalali S, Tandjung R, et al. Substitution of physicians by nurses in primary care: a systematic review and meta-analysis. BMC Health Serv Res 2014 May 12;14:214. [doi: 10.1186/1472-6963-14-214] [Medline: 24884763]
13. Laurant M, van der Biezen M, Wijers N, Watananirun K, Kontopantelis E, van Vught AJ. Nurses as substitutes for doctors in primary care. Cochrane Database Syst Rev 2018 Jul 16;7(7):CD001271. [doi: 10.1002/14651858.CD001271.pub3] [Medline: 30011347]
14. Idrees T, Castro-Revoredo IA, Migdal AL, Moreno EM, Umpierrez GE. Update on the management of diabetes in long-term care facilities. BMJ Open Diabetes Res Care 2022 Jul;10(4):e002705. [doi: 10.1136/bmjdrc-2021-002705] [Medline: 35858714]
15. Maltese G, McAuley SA, Trawley S, Sinclair AJ. Ageing well with diabetes: the role of technology. Diabetologia 2024 Oct;67(10):2085-2102. [doi: 10.1007/s00125-024-06240-2] [Medline: 39138689]
16. Or CK, Liu K, So MK, et al. Improving self-care in patients with coexisting type 2 diabetes and hypertension by technological surrogate nursing: randomized controlled trial. J Med Internet Res 2020 Mar 27;22(3):e16769. [doi: 10.2196/16769] [Medline: 32217498]
17. Holbrook A, Thabane L, Keshavjee K, et al. Individualized electronic decision support and reminders to improve diabetes care in the community: COMPETE II randomized trial. CMAJ 2009 Jul 7;181(1-2):37-44. [doi: 10.1503/cmaj.081272] [Medline: 19581618]
18. Pichardo-Lowden AR, Haidet P, Umpierrez GE, et al. Clinical decision support for glycemic management reduces hospital length of stay. Diabetes Care 2022 Nov 1;45(11):2526-2534. [doi: 10.2337/dc21-0829] [Medline: 36084251]
19. Pichardo-Lowden A, Umpierrez G, Lehman EB, et al. Clinical decision support to improve management of diabetes and dysglycemia in the hospital: a path to optimizing practice and outcomes. BMJ Open Diabetes Res Care 2021 Jan;9(1):e001557. [doi: 10.1136/bmjdrc-2020-001557] [Medline: 33462075]
20. Aberer F, Lichtenegger KM, Smajic E, et al. GlucoTab-guided insulin therapy using insulin glargine U300 enables glycaemic control with low risk of hypoglycaemia in hospitalized patients with type 2 diabetes. Diabetes Obes Metab 2019 Mar;21(3):584-591. [doi: 10.1111/dom.13559] [Medline: 30328252]
21. Mader JK, Neubauer KM, Schaupp L, et al. Efficacy, usability and sequence of operations of a workflow-integrated algorithm for basal-bolus insulin therapy in hospitalized type 2 diabetes patients. Diabetes Obes Metab 2014 Feb;16(2):137-146. [doi: 10.1111/dom.12186] [Medline: 23910952]
22. Neubauer KM, Mader JK, Höll B, et al. Standardized glycemic management with a computerized workflow and decision support system for hospitalized patients with type 2 diabetes on different wards. Diabetes Technol Ther 2015 Oct;17(10):685-692. [doi: 10.1089/dia.2015.0027] [Medline: 26355756]
23. Groeneveld S, Bin Noon G, den Ouden ME, et al. The cooperation between nurses and a new digital colleague "AI-driven lifestyle monitoring" in long-term care for older adults: viewpoint. JMIR Nurs 2024 May 23;7:e56474. [doi: 10.2196/56474] [Medline: 38781012]
24. Guo R, Xiao S, Yang F, et al. Development of a knowledge base for an integrated older adult care model (SMART system) based on an intervention mapping framework: mixed methods study. JMIR Nurs 2025 Aug 14;8:e59276. [doi: 10.2196/59276] [Medline: 40811830]
25. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020 Sep;26(9):1364-1374. [doi: 10.1038/s41591-020-1034-x] [Medline: 32908283]
26. Libiseller A, Kopanz J, Lichtenegger KM, et al. Study protocol for assessing the user acceptance, safety and efficacy of a tablet-based workflow and decision support system with incorporated basal insulin algorithm for glycaemic management in participants with type 2 diabetes receiving home health care: a single-centre, open-label, uncontrolled proof-of-concept study. Contemp Clin Trials Commun 2020 Jul 17;19:100620. [doi: 10.1016/j.conctc.2020.100620] [Medline: 32775762]
27. Langerman C, Forbes A, Robert G. A qualitative study of the experiences of insulin use by older people with type 2 diabetes mellitus. BMC Prim Care 2024 May 22;25(1):180. [doi: 10.1186/s12875-024-02318-3] [Medline: 38778253]

XSL•FO
RenderX

28. Hu X, Deng H, Zhang Y, et al. Efficacy and safety of a decision support intervention for basal insulin self-titration assisted by the nurse in outpatients with T2DM: a randomized controlled trial. Diabetes Metab Syndr Obes 2021 Mar 22;14:1315-1327. [doi: 10.2147/DMSO.S297913] [Medline: 33790599]
29. Alsahli M, Abd-Alrazaq A, Fathy DM, et al. Effectiveness of patients' education and telenursing follow-ups on self-care practices of patients with diabetes mellitus: cross-sectional and quasi-experimental study. JMIR Nurs 2025 Mar 21;8:e67339. [doi: 10.2196/67339] [Medline: 40116715]
30. Noonan SA, Gauld AL, Constantino MI, et al. A nurse-led telemonitoring approach in diabetes during the COVID-19 pandemic: prospective cohort study. JMIR Diabetes 2025 Aug 8;10:e68214. [doi: 10.2196/68214] [Medline: 40779770]
31. Carroll C, Marsden P, Soden P, Naylor E, New J, Dornan T. Involving users in the design and usability evaluation of a clinical decision support system. Comput Methods Programs Biomed 2002 Aug;69(2):123-135. [doi: 10.1016/s0169-2607(02)00036-6] [Medline: 12100792]
32. Kalun Or C, Chan AH. Inspection methods for usability evaluation. In: Stephanidis C, Salvendy G, editors. User Experience Methods and Tools in Human-Computer Interaction, 1st edition: CRC Press; 2024:170-192. [doi: 10.1201/9781003495161-7]
33. Bayor AA, Li J, Yang IA, Varnfield M. Designing clinical decision support systems (CDSS)-a user-centered lens of the design characteristics, challenges, and implications: systematic review. J Med Internet Res 2025 Jun 20;27:e63733. [doi: 10.2196/63733] [Medline: 40540451]
34. National Insulin Titration Guideline for Nurses Working With People With Diabetes Who Require Subcutaneous Insulin Injections, Version 2.: National Clinical Programmes for Diabetes and the Irish Diabetes Nurse Specialists Association; 2023 URL: https://web.archive.org/web/20231008140658/https://www.hse.ie/eng/about/who/cspd/ncps/diabetes/resources/national-insulin-titration-guideline-for-nurses.pdf [accessed 2026-09-18]
35. Eberle C, Stichling S. Clinical improvements by telemedicine interventions managing type 1 and type 2 diabetes: systematic meta-review. J Med Internet Res 2021 Feb 19;23(2):e23244. [doi: 10.2196/23244] [Medline: 33605889]
36. Snoswell CL, Taylor ML, Comans TA, Smith AC, Gray LC, Caffery LJ. Determining if telehealth can reduce health system costs: scoping review. J Med Internet Res 2020 Oct 19;22(10):e17298. [doi: 10.2196/17298] [Medline: 33074157]
37. Thomsen CH, Hangaard S, Kronborg T, Vestergaard P, Hejlesen O, Jensen MH. Time for using machine learning for dose guidance in titration of people with type 2 diabetes? A systematic review of basal insulin dose guidance. J Diabetes Sci Technol 2024 Sep;18(5):1185-1197. [doi: 10.1177/19322968221145964] [Medline: 36562599]
38. Thomsen CH, Kronborg T, Hangaard S, Vestergaard P, Jensen MH. Developing an AI-based clinical decision support system for basal insulin titration in type 2 diabetes in primary care: a mixed-methods evaluation using heuristic analysis, user feedback, and eye tracking. Int J Med Inform 2025 Mar;195:105783. [doi: 10.1016/j.ijmedinf.2024.105783] [Medline: 39787659]
39. Persson V, Lovén Wickman U. Artificial intelligence as a tool for self-care in patients with type 1 and type 2 diabetes-an integrative literature review. Healthcare (Basel) 2025 Apr 21;13(8):950. [doi: 10.3390/healthcare13080950] [Medline: 40281899]
40. Kopanz J, Mader JK, Donsa K, et al. Digital algorithm-guided insulin therapy in home healthcare for elderly persons with type 2 diabetes: a proof-of-concept study. Front Clin Diabetes Healthc 2022;3:986672. [doi: 10.3389/fcdhc.2022.986672] [Medline: 36992745]

Abbreviations

CONSORT-AI: Consolidated Standards of Reporting Trials–AI
DDSS: digital decision support system
HbA_{1c}: hemoglobin A_{1c}
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by L Donelle; submitted 22.Dec.2025; peer-reviewed by C Or, K Seibert; revised version received 07.Jul.2026; accepted 20.Jul.2026; published 24.Sep.2026.

Please cite as:
Lichtenegger KM, Scholle-Winkler L, Kopanz J, Libiseller A, Donsa K, Aberer F, Pieber TR, Mader JK
Effects of a Digital Decision Support System on Diabetes Management in a Real-World Home Care Setting: Prospective Quasi-Experimental Feasibility and Implementation Study
JMIR Nursing 2026;9:e89687
URL: https://nursing.jmir.org/2026/1/e89687
doi:10.2196/89687

XSL•FO
RenderX

© Katharina Maria Lichtenegger, Lara Scholle-Winkler, Julia Kopanz, Angela Libiseller, Klaus Donsa, Felix Aberer, Thomas Rudolf Pieber, Julia Katharina Mader. Originally published in JMIR Nursing (https://nursing.jmir.org), 24.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Nursing Process Data for Health Care Cost Prediction Using Machine Learning: Longitudinal Study

María Consuelo Company-Sancho[1,2*], RN, MSc; Víctor M González-Chordá[3,4*], RN, PhD; María Isabel Orts-Cortés[3,5*], RN, PhD

[1]Health Promotion Service, Directorate General for Public Health, Canary Islands Health Service, Las Palmas de Gran Canaria, Canary Islands, Spain

[2]Nursing and Healthcare Research Unit (Investén-isciii), Research Network on Chronicity, Primary Care and Health Promotion (RICAPPS), Carlos III Health Institute, Madrid, Spain

[3]Nursing and Healthcare Research Unit (Investén-isciii), CIBER of Frailty and Healthy Ageing (CIBERFES), Carlos III Health Institute, Madrid, Spain

[4]Nursing Research Group GIENF-241, Nursing Department, Universitat Jaume I, Castellón de la Plana, Castellón, Spain

[5]Department of Nursing (BALMIS), Institute for Health and Biomedical Research (ISABIAL, Group 23), University of Alicante, Alicante, Spain

[*]all authors contributed equally

Corresponding Author:
Víctor M González-Chordá, RN, PhD
Nursing and Healthcare Research Unit (Investén-isciii), CIBER of Frailty and Healthy Ageing (CIBERFES), Carlos III Health Institute, Madrid, Spain

Abstract

Background: Machine learning (ML) has been demonstrated to enhance health care cost prediction by handling high-dimensional data and identifying complex patterns. However, current risk-adjustment models rarely incorporate structured nursing information derived from the nursing process. This information captures care needs and human responses to health problems.

Objective: This study aimed to evaluate the impact of integrating nursing process data into ML-based predictive models of individual health care costs, including cost component analyses, compared with models based solely on sociodemographic, clinical, and morbidity-related variables.

Methods: A retrospective observational study was conducted using a population-based cohort of 1,691,075 individuals aged 15 years or younger who were registered with the Canary Islands Health Service. Predictors were derived from data available up to 2017 and included sociodemographic and clinical variables, Adjusted Morbidity Groups, health care use, and structured nursing records (Functional Health Patterns [FHP], North American Nursing Diagnosis Association [NANDA], Nursing Outcomes Classification [NOC], and Nursing Interventions Classification [NIC]). Predictive models were developed using feedforward neural networks and extreme gradient boosting; predictions were combined using an ensemble approach. An autoencoder was applied as a dimensionality-reduction technique for the nursing variables. Model performance with and without nursing variables was compared on total cost and individual cost components, and the coefficient of determination (R^2) was used on the test set.

Results: Including the nursing methodology yielded small numerical increases in predictive performance. With respect to total cost, the ensemble model improved the R^2 from 0.5023 to 0.5058 when the nursing variables were added. Although directionally consistent, these gains were limited in magnitude. In the component-level analyses, performance gains were observed in hospital care (R^2=0.2396) and pharmaceutical costs (R^2=0.6631). Reducing 789 nursing variables to 16 latent dimensions using an autoencoder substantially simplified the feature space, with predictive performance remaining broadly comparable but without a substantial additional gain.

Conclusions: Integrating structured information from the nursing process is associated with small incremental improvements in ML-based predictive models and complements commonly used sociodemographic, clinical, and morbidity variables. The systematic incorporation of nursing data into predictive tools may contribute to more accurate health care cost prediction and support more holistic, person-centered approaches.

(JMIR Nursing 2026;9:e93638) doi:10.2196/93638

KEYWORDS

standardized nursing terminology; nursing process; nursing diagnosis; machine learning; predictive modeling; health care costs; risk adjustment

XSL•FO
RenderX

Introduction

Efficient allocation of resources in health care systems requires accurate and reliable estimates of future expenditure, particularly in a context characterized by budgetary constraints and increasing care complexity. Inaccurate cost estimation may lead to inefficient decision-making and compromise the financial sustainability of health services [1-3]. In this context, machine learning (ML) has gained relevance because it provides tools capable of identifying complex patterns, handling missing values, and improving predictive accuracy compared with traditional statistical methods [4-9].

Within the field of cost prediction, risk-adjustment systems such as Adjusted Clinical Groups (ACGs) [10], Clinical Risk Groups (CRGs) [11], and Adjusted Morbidity Groups (AMGs) [12] play key roles in population risk stratification by anticipating care needs and budgetary requirements. These models have been widely used to estimate costs associated with different patient profiles, and their performance is typically assessed using metrics such as the coefficient of determination (R^2) [13]. In Spain, the AMG risk-adjustment system explains approximately 31% of the variability in total health care costs, suggesting substantial room for improvement [14]. In response to these limitations, ML techniques have increasingly been applied to risk-adjustment frameworks because of their greater capacity to identify complex patterns in large datasets. The use of neural networks, boosting models, and penalized regression approaches has been shown to improve predictive accuracy [13,15]. Methods such as neural networks and ridge regression have shown good performance in cost estimation using historical data [6].

Moreover, the types of data incorporated into risk-adjustment models have evolved. Electronic health records, which capture sociodemographic data and multiple dimensions of health status, are a rich source of information. Supervised learning models are particularly well suited for analyzing these data, as they can detect latent interactions and manage nonnormally distributed variables, such as health care costs. Leveraging these data allows more accurate estimation of health care resource use and identification of relevant patterns of association [16].

In this context, Bertsimas et al [17] reported that the accuracy of cost prediction models depends largely on the quality, completeness, and standardization of the data used and suggested that incorporating new variables that more faithfully reflect actual care delivery may enhance predictive performance. However, most predictive models, including those based on variables commonly used in ACG, CRG, and AMG risk-adjustment systems, rely primarily on sociodemographic and clinical variables such as gender, age, comorbidities, procedures, and pharmaceutical expenditures [11]. The inclusion of social determinants has been shown to significantly improve the predictive capacity of these models. For example, a recent study reported that incorporating such factors increased the R^2 from 0.327 to 0.388, reducing estimation errors by 3.5 million dollars per 10,000 individuals added [18]. Similarly, factors such as low income, educational level, food insecurity, or housing instability have been significantly associated with high-cost patient profiles [19], although these variables are rarely included in cost prediction models. Harnessing the richness of electronic health record data, together with big data analytics in health care, offers opportunities to improve outcomes and reduce costs [20].

Despite these advances, most predictive models continue to exclude information derived from the nursing process, even though it has the potential to capture essential aspects of health status and service use [2,6]. The incorporation of standardized nursing terminologies such as the North American Nursing Diagnosis Association (NANDA), the Nursing Outcomes Classification (NOC), and the Nursing Interventions Classification (NIC) could substantially enhance predictive models by capturing the impact of nursing interventions. These terminologies, which have been progressively integrated into electronic health records, enable systematic characterization of nursing practice, facilitate assessment of its impact on health outcomes, and support the development of robust predictive models [21]. Understanding and quantifying nursing care in a systematic and effective manner may contribute to optimizing resource allocation and improving health system sustainability [22]. In this context, Functional Health Patterns (FHP) described by Gordon provide a nursing assessment framework that offers a holistic view of patient health, encompassing physical, emotional, and social dimensions. Although this framework may add value in terms of efficiency and resource management, its impact on health care cost prediction has rarely been explored [23]. Nevertheless, a previous study revealed that more than 21% of the variability in total health care costs could be explained by a model including gender, age, and variables derived from the nursing process (NANDA, NOC, NIC, and FHP) [24]. This study adds a comparison between nursing process data and the clinical, sociodemographic, and morbidity-related variables commonly used in cost prediction. It also extends the analysis to partial cost components and explores dimensionality reduction.

Recent literature has highlighted that incorporating additional layers of information, such as social determinants or measures of well-being, significantly improves the performance of health care cost prediction models and helps correct biases affecting vulnerable populations [18,19]. In parallel, the inclusion of structured nursing data has been shown to provide valuable insights into the impact of nursing practice on health outcomes and patient safety [21]. However, to date, no studies have systematically integrated information derived from the nursing process into ML-based models specifically designed for health care cost prediction.

Against this background, leveraging the potential of nursing big data and integrating it with advanced ML techniques represents an opportunity to improve risk-adjustment systems, enhance population stratification, and support more efficient and equitable resource allocation. Accordingly, the aim of this study was to evaluate the impact of integrating nursing process data into ML-based predictive models of individual health care costs, including cost component analyses, compared with models based solely on sociodemographic, clinical, and morbidity-related variables.

XSL•FO
RenderX

Methods

Study Design and Setting

The present study forms part of a longitudinal research program based on a single population cohort that aims to understand the determinants of health care costs from a comprehensive perspective. In the first phase, individual health care costs were estimated, and their variability was analyzed according to AMG using multiple linear regression [14]. In the second phase, ML-based predictive models were developed to explain the variability in total health care costs, incorporating information derived from the nursing process [24].

The present (third) phase extends this analytical framework by focusing specifically on the contribution of nursing process variables (NANDA, NOC, NIC, and FHP) to health care cost prediction. To this end, a retrospective, observational, and analytical study was conducted using data from multiple health care information systems. The study was carried out in the Autonomous Community of the Canary Islands (Spain) using data from 2017 and 2018. The predictor variables were derived from data available up to 2017, whereas health care costs were measured for 2018, thus ensuring appropriate temporal ordering for predictive modeling.

Ethical Considerations

In accordance with the Declaration of Helsinki and Spanish legislation, this study was approved by the Clinical Research Ethics Committee of the Dr. Negrín University Hospital of Gran Canaria (date: January 29, 2021; code: 2021-037-1). All applicable ethical and regulatory principles for studies using secondary health data were observed.

Study Population

The study population included all individuals aged 15 years or older who were registered in the Canary Islands Health Service health care card database as of December 31, 2017 (N=1,691,075). The entire target population was included. Individuals covered by mutual insurance schemes were excluded; these correspond to public-sector employees whose health care and social care are managed outside the Canary Islands Health Service.

Variables

Predictor variables were derived from data available up to 2017 and included sociodemographic, clinical, morbidity-related, and nursing process information. The sociodemographic variables comprised age, gender, type of affiliation with the social security system, and the percentage of pharmaceutical copayment. Clinical variables included a Barthel Index score greater than 60, a Pfeiffer test score of 5 or lower, 2 or more hospital admissions in the previous 12 months, and the number of prescribed active pharmaceutical ingredients. Morbidity-related variables were obtained from the AMG system and included the complexity weight and morbidity group.

Nursing process variables were also incorporated and consisted of 44 variables derived from the assessment of the 11 FHP (with 4 possible outcomes per pattern), as well as 211 NANDA nursing diagnoses, 388 NOC outcomes, and 558 NIC interventions. The large number of variables reflects the conceptual structure of nursing data, which captures complementary dimensions of human responses, outcomes, and interventions; therefore, regularization and hyperparameter tuning were applied to mitigate the dimensional burden.

The primary outcome was total health care cost per individual in 2018. This was calculated as the sum of the costs associated with primary care and specialist consultations, emergency department visits, hospital admissions, major ambulatory surgery in public and contracted centers, and pharmaceutical dispensing (Multimedia Appendix 1). In addition, a cost component analysis was conducted for different care settings (primary care, hospital care, and contracted care) and service types (hospital admissions, outpatient consultations, and emergency department visits).

Data Sources

Data were obtained from the following information systems: the Canary Islands Health Service health care card database (sociodemographic variables), primary care electronic health records (medical and nursing consultations and nursing process variables), hospital electronic health records (hospital consultations and emergency department visits), the Minimum Basic Dataset of the Hospital Activity Registry (hospital admissions and surgical procedures), the Hospital Contracted Care Information System (activity in contracted centers), and the electronic prescribing system of the Canary Islands Health Service (pharmaceutical dispensing). All databases were linked using a unique, encrypted clinical identifier for each individual.

Data Preprocessing and Analytical Strategy

Prior to analysis, numerical variables (age and AMG complexity weight) were standardized. Cost variables were transformed using natural logarithms to approximate normal distributions. Nursing diagnoses (NANDA), outcomes (NOC), and interventions (NIC) with a prevalence lower than 1 per 10,000 individuals were excluded, as they provided limited information, had a high proportion of zero or missing values, and unnecessarily increased model dimensionality. This threshold was applied pragmatically, as no comparable reference studies were identified, with the aim of filtering infrequent variables and thereby reducing sparsity and computational burden.

After the data were cleaned, a descriptive analysis of the study population was performed. As the entire population was included, no statistical inference techniques were applied. The analytical focus was exclusively on the development and evaluation of predictive models.

Two ML algorithms capable of handling large-scale, high-dimensional health care data were evaluated using complementary modeling approaches. First, a feedforward neural network was implemented [25], which consisted of interconnected layers designed to capture nonlinear relationships and latent patterns. Second, extreme gradient boosting (XGBoost) [26], a gradient-boosted decision tree algorithm recognized for its efficiency, accuracy, and ability to handle missing values and skewed distributions, was applied. Predictions from both models were combined using an ensemble approach based on their averaged outputs to improve robustness [27].

XSL•FO
RenderX

The dataset was randomly split into training (80%), validation (10%), and test (10%) subsets. During model training, the mean squared error (MSE) was used as the optimization metric, while overall performance was assessed using the coefficient of determination (R^2). Because health care cost data were highly skewed, the outcome was log-transformed before modeling, and both model fitting and performance metrics, including R^2, were calculated on that log-transformed scale. R^2 was used as the primary metric to compare relative model performance across the different predictor sets. Given the very large cohort size, a single training, validation, or test split was considered sufficient for initial model development and evaluation, as both the validation and test subsets each included more than 160,000 individuals. For the feedforward neural network, the final architecture consisted of 4 hidden dense layers (128, 64, 32, and 16 units) with rectified linear unit activation and L1 regularization (l1=0.0000625; l2=0). The model was trained using the root mean square propagation optimizer with a learning rate of 0.0005, a batch size of 64, and a maximum of 60 epochs. Early stopping was applied based on the validation R^2 (monitor=val_r_square, mode=max, and patience=5), and the best model weights were restored. For XGBoost, the final model was trained with a learning rate of 0.0960, a maximum tree depth of 10,500 estimators, a subsample of 0.9413, full column sampling per tree (colsample_bytree=1.0), L1 regularization (reg_alpha=10.0), and L2 regularization (reg_lambda=0.001). Hyperparameter optimization was conducted on the validation set, and the final model performance was evaluated exclusively on the test set. A fixed random seed of 42 was used to ensure reproducibility.

The analytical strategy involved 3 stages. First, baseline models excluding the nursing process variables and extended models including these variables were developed and compared for total health care cost prediction. Second, cost component analyses were conducted at both the care setting level (primary care, hospital care, and contracted care) and the service level (hospital admissions, outpatient consultations, and emergency department visits), and model performance was compared with and without the nursing process variables. Third, an autoencoder was applied to reduce the dimensionality of the nursing process variables, as the feature block was high-dimensional and predominantly binary, making a more flexible nonlinear approach preferable to the conventional principal component analysis. This neural network architecture compresses high-dimensional input data into a lower-dimensional latent representation and reconstructs it, thereby enabling the identification of unsupervised latent patterns [28]. Autoencoders have previously been applied in hospital cost prediction studies [29]. All analyses were performed using Python version 3.10.8, with the *pandas*, *NumPy*, *scikit-learn*, *TensorFlow*, and *XGBoost* libraries.

Results

Characteristics of the Study Population

The study included 1,691,075 individuals aged over 15 years, including 863,071 (51.0%) women; the mean age was 46.6 (SD 17.9) years, and 17.4% (n=294,077) were older than 65 years. The mean AMG index weight was 5.82 (SD 5.89) points, which was higher in women (6.25, SD 6.06) than in men (5.10, SD 5.62) and increased with age (Spearman ρ=0.47). A total of 216,630 (12.8) women and 175,536 (10.4) men were pensioners. The mean number of prescribed medications was 4.35 (SD 5.10), with higher values observed in women.

Overall, 22,800 (1.3%) individuals had 2 or more hospital admissions in the previous year, 6308 (0.4%) had a Barthel Index score greater than 60, and 7271 (0.4%) demonstrated cognitive impairment (Pfeiffer score≥5). Nursing records (FHP, NANDA, NOC, and NIC) were present for 57.9% (980,437/1,691,075) of the population, although only 2.7% (46,380/1,691,075) had completed the full nursing process. On average, 3.84 (SD 3.58) FHP were assessed per person, together with 3.07 (SD 2.9) NANDA nursing diagnoses, 2.77 (SD 2.96) NOC outcomes, and 4.12 (7.04) NIC interventions.

The mean cost per patient was €1283.08 (SD €3007; €1=US $1.19 as of December 31, 2017), with higher values observed among women. Hospital admissions and consultations accounted for the highest cost components. Costs were heterogeneously distributed across care settings, with those for consultations (mean €355.07, SD €586.87) and pharmaceutical dispensing (mean €345.53, SD €946.09) being especially notable. Table 1 presents details of the variables included in the models; the complete cohort profile can be found in previous publications [14,24].

XSL•FO
RenderX

Table . Characteristics of the study population and variables included in the predictive models.

Variables	Total[a]	Women	Men
Sociodemographic and clinical, n (%)			
Pensioners	392,166 (23.2)	216,630 (12.8)	175,536 (10.4)
≥2 hospital admissions	22,846 (1.4)	12,082 (0.7)	10,764 (0.6)
Barthel Index <60	5971 (0.4)	4315 (0.3)	1656 (0.1)
Pfeiffer score ≥5	6834 (0.4)	5178 (0.3)	1656 (0.1)
AMG[b] system, mean (SD)			
Complexity weight	5.82 (5.89)	6.25 (6.06)	5.10 (5.62)
Mean number medications	4.35 (5.10)	5.19 (5.47)	3.48 (4.55)
Nursing process, n (%)			
Nursing process records	980,437 (57.9)	533,231 (61.8)	447,206 (54.0)
Completed nursing process	46,380 (2.7)	29,405 (3.4)	16,975 (2.1)
At least 1 FHP[c] assessed	415,068 (24.5)	247,162 (28.6)	167,906 (20.3)
At least 1 active NANDA[d]	916,024 (54.2)	489,738 (56.7)	426,286 (51.5)
At least 1 active NOC[e]	772,730 (45.7)	419,496 (48.6)	353,234 (42.7)
At least 1 active NIC[f]	787,510 (46.6)	427,370 (49.5)	360,140 (43.5)
Cost (€; €1=US $1.19)[g], mean (SD)			
Mean total health care cost	1283.08 (3007)	1377.04 (2868.27)	1185.14 (3142.05)
Hospital admissions	463.11 (2207.39)	461.55 (2060.64)	464.72 (2350.62)
Outpatient consultations	355.07 (586.87)	418.80 (623.38)	288.64 (538.24)
Emergency department visits	119.35 (324.68)	131.10 (332.68)	107.09 (315.67)
Pharmaceutical dispensing	345.53 (946.09)	365.55 (914.18)	324.67 (977.81)
Hospital care	571.90 (2274)	597.73 (2118.51)	544.99 (2425.37)
Primary care	246.85 (421)	290.96 (454.43)	200.87 (378.69)
Contracted care	118.8 (808)	122.76 (827.41)	114.59 (786.31)

[a]The cohort included 1,691,075 individuals (863,071 women and 828,004 men).

[b]AMG: Adjusted Morbidity Groups.

[c]FHP: Functional Health Pattern.

[d]NANDA: North American Nursing Diagnosis Association.

[e]NOC: Nursing Outcomes Classification.

[f]NIC: Nursing Interventions Classification.

[g]Health care cost components are presented by care setting (primary care, hospital care, and contracted care) and by service type (hospital admissions, outpatient consultations, and emergency department visits), as well as pharmaceutical dispensing.

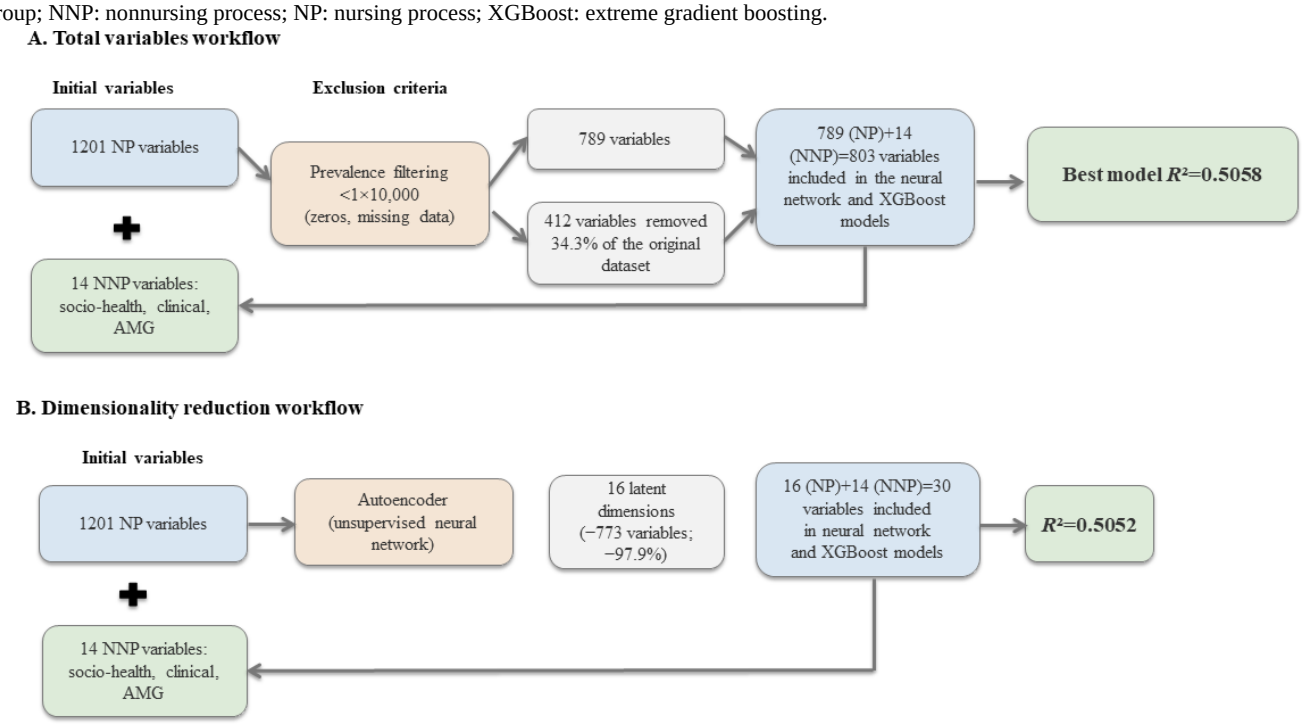
Predictive Performance of the Baseline and Extended Models for Health Care Cost Prediction

The neural network architecture with the best performance was a feedforward network with 4 hidden layers (128, 64, 32, and 16 nodes) and an output corresponding to the logarithm of the total cost. For the XGBoost model, Bayesian hyperparameter optimization was applied to identify the combinations with the best performance, and an ensemble model was subsequently generated by averaging the predictions.

The initial feature set included sociodemographic and clinical variables, together with 1201 nursing process variables. After those with a positive prevalence of fewer than one per 10,000 patients were excluded, 789 nursing variables and 14 sociodemographic, clinical, and AMG system variables were retained (Figure 1).

XSL•FO
RenderX

Figure 1. Workflow of nursing-variable selection and dimensionality reduction for the neural network and XGBoost models. AMG: Adjusted Morbidity Group; NNP: nonnursing process; NP: nursing process; XGBoost: extreme gradient boosting.



In the baseline models, excluding the nursing variables, the R^2 was 0.4996 for the neural network, 0.5030 for XGBoost, and 0.5023 for the ensemble model. After the nursing process variables were incorporated, the predictive performance increased to 0.5026 (MSE 3.8561), 0.5065 (MSE 3.7936), and 0.5058 (MSE 3.7981), respectively, indicating improvements across all 3 algorithms.

Predictive Performance by Cost Component According to Care Setting and Service Type

Across most cost components, the inclusion of the nursing process variables improved predictive performance, with variations depending on the care setting and model type (Table 2). For hospital care, the greatest increase was observed in the ensemble model, with R^2 increasing from 0.2161 to 0.2396 (+0.0235), while the improvements in the neural network and XGBoost models were similar (+0.0230 and +0.0223, respectively). For primary care, improvements were more moderate but consistent; the ensemble model increased from 0.4388 to 0.4442 (+0.0054), and comparable gains were observed for the neural network (+0.0050) and XGBoost (+0.0054). For emergency department visits, the predictive performance also increased, with that of the ensemble model increasing from 0.1431 to 0.1506 (+0.0075) and similar gains observed for XGBoost (+0.0074). For contracted care, improvements were more limited in magnitude but were observed across all models, with the ensemble model's performance increasing from 0.0389 to 0.0443 (+0.0054).

XSL•FO
RenderX

Table . Predictive performance (R^2) of machine learning–models with and without nursing process variables across health care cost components.

Health care cost components and models	No (R^2)	Yes (R^2)	Differences
Care settings			
Primary care			
Neural network R^2	0.4359	0.4409	+0.0050
XGB[a] R^2	0.4394	0.4448	+0.0054
Ensemble R^2	0.4388	0.4442	+0.0054
Hospital care			
Neural network R^2	0.2133	0.2363	+0.0230
XGB R^2	0.2163	0.2386	+0.0223
Ensemble R^2	0.2161	0.2396	+0.0235
Contracted care			
Neural network R^2	0.0382	0.0402	+0.0020
XGB R^2	0.0380	0.0452	+0.0072
Ensemble R^2	0.0389	0.0443	+0.0054
Service types			
Hospital admissions			
Neural network R^2	0.0624	0.0618	−0.0006
XGB R^2	0.0636	0.0652	+0.0016
Ensemble R^2	0.0641	0.0650	+0.0009
Outpatient consultations			
Neural network R^2	0.4568	0.4608	+0.0040
XGB R^2	0.4606	0.4652	+0.0046
Ensemble R^2	0.4598	0.4642	+0.0044
Emergency department visits			
Neural network R^2	0.1407	0.1473	+0.0066
XGB R^2	0.1434	0.1508	+0.0074
Ensemble R^2	0.1431	0.1506	+0.0075
Other cost components			
Pharmaceutical costs			
Neural network R^2	0.6565	0.6612	+0.0047
XGB R^2	0.6588	0.6631	+0.0043
Ensemble R^2	0.6584	0.6631	+0.0047

[a]XGB: extreme gradient boosting.

When analyzed by service type, outpatient consultations and emergency department visits showed the most consistent improvements. For outpatient consultations, the R^2 of the ensemble model increased from 0.4598 to 0.4642 (+0.0044), whereas for emergency department visits, it increased from 0.1431 to 0.1506 (+0.0075). For hospital admissions, the changes were minimal; the ensemble model R^2 increased slightly from 0.0641 to 0.0650 (+0.0009), whereas that of the neural network marginally decreased (−0.0006).

Predictive Models Using Latent Variables Derived From the Autoencoder

Finally, an autoencoder was applied to reduce the 789 nursing variables to 16 latent dimensions while preserving relevant data patterns. These compressed variables were incorporated with the sociodemographic, clinical, and AMG-related variables. With this combination, the models achieved R^2 values of 0.5013 for the neural network, 0.5062 for XGBoost, and 0.5052 for the ensemble model. Thus, dimensionality reduction markedly simplified the nursing feature space while yielding predictive performance that was broadly comparable to that of the models

XSL•FO
RenderX

using the larger nursing-variable set, but without a substantial performance improvement.

Discussion

Principal Findings

The integration of variables derived from nursing methodology into ML-based prediction of total and component-level health care costs resulted in small numerical improvements in total cost prediction, whereas changes observed across cost components were heterogeneous and, in many cases, of small magnitude. These results suggest that nursing information may contain an additional predictive signal, although the magnitude of this contribution was limited, thereby reflecting differences in care delivery structures and in the relationship between care processes and resource use. Accordingly, the balance between added model complexity and predictive benefit should be interpreted with caution, as the contribution of nursing information varies according to the care setting and service type considered.

To date, the predominant methodologies used in health care cost analysis have systematically excluded data related to nursing interventions and outcomes, despite evidence that nursing terminologies can improve both the quality and the efficiency of resource management [30]. Their standardization enables the documentation and quantification of nursing work and provides a holistic view of the patient that is not usually captured by traditional biomedical data. Recent studies have shown that incorporating nursing information into clinical databases facilitates a better understanding of the impact of nursing practice on patient outcomes [21] and contributes to improving patient safety [31], with a potential indirect effect on cost reduction.

However, the literature on the application of ML to nursing data has focused primarily on the development of intelligent clinical tools, such as virtual assistants, fall prediction models, and patient monitoring systems, rather than on economic outcomes [32]. In this context, the present study represents, to our knowledge, a novel approach by systematically integrating nursing process information into cost prediction models using ML techniques, thereby extending previous research in which the added economic value of nursing data has rarely been evaluated.

The use of ML algorithms has demonstrated usefulness in health care cost prediction by overcoming limitations inherent in traditional statistical models, such as data skewness and distributional bias [6,33]. In general, the variables most commonly used to predict health care expenditure are medical diagnoses and prescribed medications [2,34]. For example, in the study by Drewe-Boss et al [2], neural network–based models relied primarily on medical diagnoses coded according to the *International Classification of Diseases*, 10th Revision (ICD-10) and the Anatomical Therapeutic Chemical (ATC) classification system. In contrast, models built using ridge regression placed greater weight on Gebührenordnungspositionen codes, which are broadly comparable to the diagnosis-related groups used in Spain and other European countries.

Similarly, comorbidity may be a better predictor of pharmaceutical consumption and other health care products than clinical complexity, which encompasses psychological, social, and functional dimensions [35]. The findings of the present study are partially consistent with this pattern because the inclusion of nursing variables increased the predictive capacity for total health care costs, although the magnitude of this increase was modest. Nevertheless, given the limited magnitude of the observed improvement, its operational relevance for population risk adjustment or resource planning should be interpreted with caution.

Health care costs are complex and influenced by multiple factors. The incorporation of measures of well-being into prediction models has been shown to improve the estimation of future health care costs by 5.7% to 13% [36]. Another study revealed that each one-point increase in well-being indices was associated with a 1% reduction in health care expenditure [37]. Similarly, the inclusion of social determinants in predictive models has demonstrated additional benefits: in populations characterized by high levels of poverty, inequality, or lack of coverage, social determinants reduced cost underestimation and improved predictive performance by 3%, corresponding to approximately US $200 per person per year [18].

The integration of this type of data into ML models represents a promising avenue for improving the efficiency and sustainability of health care systems [38]. Along these lines, nursing data capture dimensions related to functional status, care needs, and human responses to health problems that are traditionally absent from cost models and could complement the social and clinical determinants commonly used.

In the analyses performed, pharmaceutical costs showed the highest predictive accuracy, with R^2 values increasing from 0.6584 to 0.6631 following the inclusion of nursing variables. This finding suggests that these variables may capture information related to therapeutic adherence, medication management, and educational processes, which are consistent with the role of nursing practice. Although the observed improvement was modest, its limited magnitude means that any interpretation regarding its practical usefulness should remain cautious. In contrast, in other areas, such as outpatient consultations or primary care, changes were small and, in some cases, showed a downward trend, indicating a limited effect of nursing variables in these specific settings. Overall, these patterns indicate that the contribution of nursing methodology is not uniform and may be more relevant in functions closely linked to clinical follow-up and pharmacological management.

In primary care, a setting in which nursing plays a central role, particularly in the management of chronic conditions, an improvement in predictive performance was also observed (R^2 increasing from 0.4388 to 0.4442). This result suggests that long-term care and continuous follow-up, areas with greater nursing involvement, may benefit more from the systematic incorporation of the nursing process into predictive models.

In contrast, the XGBoost models consistently showed better predictive performance than the neural networks in the present analysis, although the literature has reported mixed findings. For example, in a study on pharmaceutical cost prediction,

XSL•FO
RenderX

boosted tree–based models outperformed neural networks (area under the curve [AUC]=0.74) [39], whereas another study on hospital cost prediction reported better performance for neural networks (R^2=0.813) than for decision trees (R^2=0.713) [40]. Nevertheless, XGBoost is widely regarded as one of the most robust algorithms for tabular data and for scenarios with missing values [41]. The consistency of its performance in the present study further supports its suitability for complex health care data with multiple categorical and sparse variables, such as those derived from the nursing process.

One of the main uses of predictive models lies in identifying relevant features within large datasets and in understanding their contribution to outcome prediction. However, when the number of variables is high, the presence of irrelevant or redundant dimensions may adversely affect model performance by introducing noise and unnecessarily increasing complexity. In the present study, the modeling strategy was designed to assess the incremental predictive value of the nursing data block as a whole, rather than to determine the individual relevance of specific nursing diagnoses, outcomes, interventions, or functional patterns.

In this context, autoencoders have become established as effective tools for dimensionality reduction, redundancy elimination, and the generation of latent representations that preserve essential information [42]. In the present study, autoencoder implementation enabled the condensing of 789 nursing methodology variables into 16 latent variables. Although this reduction preserved relevant patterns and substantially simplified model complexity, it did not translate into a substantial improvement in predictive performance, indicating that dimensionality reduction alone was not sufficient to strengthen predictive performance in a meaningful way. Therefore, the autoencoder should be interpreted here as a strategy for representation simplification rather than as a predictive advantage. Future studies could explore which of these latent dimensions are most informative and whether they capture distinct patient profiles.

Among the most widely used health care management tools are risk-adjustment models such as ACGs, Diagnosis-Related Groups, and AMGs, which are extensively applied in Spain and other countries. These models are integrated into electronic health records with the aim of classifying morbidity, predicting hospital admissions or mortality, and estimating health care costs, thereby supporting efficient resource allocation and informing managerial decision-making. However, because they are largely based on linear statistical approaches, they present important limitations when applied in settings characterized by complex, nonlinear, and high-dimensional data, as is common in clinical information systems [15]. These limitations may result in the underestimation of actual health care costs and reduced predictive accuracy in scenarios of high care complexity.

In contrast, ML algorithms have been shown to outperform traditional models across several key metrics. For example, in mortality prediction, they achieve AUC values ranging from 0.92 to 0.94, compared with 0.91 for classical models, and in the prediction of hospital readmissions, performance improves from 0.69 to 0.75-0.76 [43]. In light of these findings, the present results suggest that combining ML algorithms with variables derived from the nursing process could increase the precision and practical use of current risk adjustment models, particularly in settings characterized by a high nursing workload.

Implications for Practice

The findings of the present study indicate that the systematic incorporation of variables derived from the nursing process into clinical information systems may contribute to improving the accuracy of health care cost prediction models and could enhance population stratification tools. Integrating these data may enable more accurate identification of patient groups with greater resource needs and facilitate the planning of more efficient, evidence-based interventions.

In addition, recognizing the predictive value of nursing documentation highlights its contribution to health care system management and underscores the importance of promoting structured, interoperable, and high-quality nursing records in clinical practice.

Strengths and Limitations

The present study has several limitations that should be considered when interpreting the findings. First, the observation period was limited to a single year, which restricts the ability to analyze long-term cost trajectories. Although previous studies have indicated that predictive performance improves with longer follow-up periods [2,44], the use of a large population-based cohort (N=1,691,075) provides a robust foundation and reduces the risk of overfitting, thereby strengthening the external validity of the analysis.

Second, comparability with other studies is limited because of the heterogeneity in the algorithms, validation strategies, and performance metrics reported in the literature. In addition, no studies have been identified that have systematically integrated nursing process variables into health care cost prediction models, which hinders direct comparison but also highlights the innovative nature of this work.

Another limitation relates to dimensionality reduction using autoencoders. Although this approach enabled the synthesis of 789 nursing variables into 16 latent components, the clinical interpretability of these components was lower, and no substantial improvement in predictive performance was observed. Nevertheless, these latent representations offer opportunities for future research aimed at identifying patient profiles and patterns associated with resource use.

Finally, the rapid evolution of AI techniques means that some approaches may become outdated over time. However, the methods applied represent well-established standards in ML and are appropriate for the stated objectives. Replicating the present study in other health care systems would help strengthen the generalizability of the findings and broaden their applicability.

Conclusions

The incorporation of variables derived from the nursing process into ML-based health care cost prediction models was associated

XSL•FO
RenderX

with small incremental improvements in the explanatory capacity for total health care costs, whereas effects observed across cost components were heterogeneous and of smaller magnitude. These findings suggest that structured nursing information may complement the sociodemographic, clinical, and morbidity-related variables commonly used in health care expenditure estimation.

The results highlight the relevance of moving toward the integration of nursing process information into population-stratification systems and risk-adjustment models, which are key tools for health care planning and efficient resource allocation. The inclusion of dimensions related to functional status, care needs, and human responses broadens the traditional perspective based on clinical and administrative data, thereby supporting more comprehensive approaches to cost prediction.

The potential of ML to improve the estimation of health care resource use should not be limited exclusively to conventional biomedical predictors. The incorporation of structured information derived from nursing practice may enhance the understanding of the determinants of health care costs and support the development of more holistic, person-centered predictive models aligned with contemporary approaches to integrated care.

Acknowledgments

No generative AI tools were used in the preparation of this manuscript. Authors remain fully responsible for the accuracy, originality, and integrity of all the content in the manuscript, including all the references and their citations.

Funding

The authors gratefully acknowledge the financial support provided by the Instituto de Investigación Sanitaria y Biomédica de Alicante (ISABIAL) to cover the open access publication costs of this article within the programme 2026-0243 // UGP-26-05

Data Availability

The datasets generated and/or analyzed during the study are not publicly available because of the inclusion of sensitive and confidential data but are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: MCC-S, MIO-C, VMG-C
Investigation: MCC-S, MIO-C
Formal analysis: MCC-S, VMG-C
Supervision: MCC-S, MIO-C, VMG-C
Writing – original draft: MCC-S, MIO-C, VMG-C
Writing – review & editing: MCC-S, VMG-C, MIO-C

Conflicts of Interest

None declared.

Multimedia Appendix 1
Cost variables used to calculate total costs.
[DOCX File, 14 KB - nursing_v9i1e93638_app1.docx]

References

1. Santamaría Benhumea AM, Herrera Villalobos JE, Sil Jaimes PA, Santamaría Benhumea NH, Flores Manzur M, del Arco Ortiz A. Estructura, sistemas y análisis de costos de la atención médica hospitalaria [Article in Spanish]. Medicina E Investigación 2015 Jul;3(2):134-140. [doi: 10.1016/j.mei.2015.06.001]
2. Drewe-Boss P, Enders D, Walker J, Ohler U. Deep learning for prediction of population health costs. BMC Med Inform Decis Mak 2022 Feb 3;22(1):32. [doi: 10.1186/s12911-021-01743-z] [Medline: 35114978]
3. Kaplan RS, Porter ME. How to solve the cost crisis in health care. Harv Bus Rev 2011 Sep;89(9):46-52. [Medline: 21939127]
4. Badawy M, Ramadan N, Hefny HA. Healthcare predictive analytics using machine learning and deep learning techniques: a survey. J Electr Syst Inf Technol 2023;10(1):40. [doi: 10.1186/s43067-023-00108-y]
5. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017 Feb 2;542(7639):115-118. [doi: 10.1038/nature21056] [Medline: 28117445]
6. Morid MA, Kawamoto K, Ault T, Dorius J, Abdelrahman S. Supervised learning methods for predicting healthcare costs: systematic literature review and empirical evaluation. AMIA Annu Symp Proc 2018;2017:1312-1321. [Medline: 29854200]
7. Perez-Lebel A, Varoquaux G, Le Morvan M, Josse J, Poline JB. Benchmarking missing-values approaches for predictive models on health databases. Gigascience 2022;11:giac013. [doi: 10.1093/gigascience/giac013] [Medline: 35426912]

XSL•FO
RenderX

8. Ru B, Tan X, Liu Y, et al. Comparison of machine learning algorithms for predicting hospital readmissions and worsening heart failure events in patients with heart failure with reduced ejection fraction: modeling study. JMIR Form Res 2023 Apr 17;7:e41775. [doi: 10.2196/41775] [Medline: 37067873]
9. Wemmert C, Weber J, Feuerhake F, Forestier G. Deep learning for histopathological image analysis. In: Elloumi M, editor. Deep Learning for Biomedical Data Analysis: Springer; 2021:153-169. [doi: 10.1007/978-3-030-71676-9_7]
10. The Johns Hopkins ACG® system: excerpt from version 11.0 technical reference guide. : Johns Hopkins University, Bloomberg School of Public Health; 2014 URL: https://www.hopkinsacg.org/document/acg-system-version-11-technical-reference-guide/ [accessed 2026-08-25]
11. 3M™ Clinical Risk Groups: measuring risk, managing care. : 3M Health Information Systems / 3M United Kingdom PLC; 2016 URL: https://multimedia.3m.com/mws/media/1356109O/crg-measuring-risk-managing-care-ukv1.pdf [accessed 2025-02-22]
12. Monterde D, Vela E, Clèries M, grupo colaborativo GMA. Adjusted morbidity groups: a new multiple morbidity measurement of use in primary care [Article in Spanish]. Aten Primaria 2016 Dec;48(10):674-682. [doi: 10.1016/j.aprim.2016.06.003] [Medline: 27495004]
13. Rose S. A machine learning framework for plan payment risk adjustment. Health Serv Res 2016 Dec;51(6):2358-2374. [doi: 10.1111/1475-6773.12464] [Medline: 26891974]
14. Company-Sancho MC, González-Chordá VM, Orts-Cortés MI. Variability in healthcare expenditure according to the stratification of adjusted morbidity groups in the Canary Islands (Spain). Int J Environ Res Public Health 2022 Apr 1;19(7):4219. [doi: 10.3390/ijerph19074219] [Medline: 35409900]
15. Kan HJ, Kharrazi H, Chang HY, Bodycombe D, Lemke K, Weiner JP. Exploring the use of machine learning for risk adjustment: acomparison of standard and penalized linear regression models in predicting health care costs in older adults. PLoS One 2019;14(3):e0213258. [doi: 10.1371/journal.pone.0213258] [Medline: 30840682]
16. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet 2012 May 2;13(6):395-405. [doi: 10.1038/nrg3208] [Medline: 22549152]
17. Bertsimas D, Bjarnadóttir MV, Kane MA, et al. Algorithmic prediction of health-care costs. Oper Res 2008 Dec;56(6):1382-1392. [doi: 10.1287/opre.1080.0619]
18. Irvin JA, Kondrich AA, Ko M, et al. Incorporating machine learning and social determinants of health indicators into prospective risk adjustment for health plan payments. BMC Public Health 2020 May 1;20(1):608. [doi: 10.1186/s12889-020-08735-0] [Medline: 32357871]
19. Fitzpatrick T, Rosella LC, Calzavara A, et al. Looking beyond income and education: socioeconomic status gradients among future high-cost users of health care. Am J Prev Med 2015 Aug;49(2):161-171. [doi: 10.1016/j.amepre.2015.02.018] [Medline: 25960393]
20. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2:3. [doi: 10.1186/2047-2501-2-3] [Medline: 25825667]
21. Macieira TGR, Chianca TCM, Smith MB, et al. Secondary use of standardized nursing care data for advancing nursing science and practice:a systematic review. J Am Med Inform Assoc 2019 Nov 1;26(11):1401-1411. [doi: 10.1093/jamia/ocz086] [Medline: 31188439]
22. Welton JM, Harper EM. Measuring nursing care value. Nurs Econ 2016;34(1):7-14. [Medline: 27055306]
23. Gengo E Silva Butcher RC, Jones DA. An integrative review of comprehensive nursing assessment tools developed based on Gordon's eleven functional health patterns. Int J Nurs Knowl 2021 Oct;32(4):294-307. [doi: 10.1111/2047-3095.12321] [Medline: 33620162]
24. Company-Sancho MC, González-Chordá VM, Isabel Orts-Cortés M. The nursing process and total health cost variability: an analysis using machine learning. BMC Nurs 2025 Jul 1;24(1):738. [doi: 10.1186/s12912-025-03304-5] [Medline: 40598233]
25. Goodfellow I, Bengio Y, Courville A. Deep Learning: MIT Press; 2016.
26. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: Association for Computing Machinery; 2016:785-794. [doi: 10.1145/2939672.2939785]
27. Mahajan P, Uddin S, Hajati F, Moni MA. Ensemble learning for disease prediction: a review. Healthcare (Basel) 2023 Jun 20;11(12):1808. [doi: 10.3390/healthcare11121808] [Medline: 37372925]
28. Berahmand K, Daneshfar F, Salehi ES, Li Y, Xu Y. Autoencoders and their applications in machine learning: a survey. Artif Intell Rev 2024;57(2):28. [doi: 10.1007/s10462-023-10662-6]
29. Bhatti MHR, Javaid N, Mansoor B, Alrajeh N, Aslam M, Asad M. New hybrid deep learning models to predict cost from healthcare providers in smart hospitals. IEEE Access 2023;11:136988-137010. [doi: 10.1109/ACCESS.2023.3336424]
30. Müller-Staub M, Lavin MA, Needham I, van Achterberg T. Nursing diagnoses, interventions and outcomes - application and impact on nursing practice: systematic review. J Adv Nurs 2006 Dec;56(5):514-531. [doi: 10.1111/j.1365-2648.2006.04012.x] [Medline: 17078827]
31. Douma MC, Rejeb MB, Zardoub N, et al. Impact of implementing electronic nursing records on quality and safety indicators in care. Libyan J Med 2024 Dec 31;19(1):2421625. [doi: 10.1080/19932820.2024.2421625] [Medline: 39570988]

XSL•FO
RenderX

32. Yelne S, Chaudhary M, Dod K, Sayyad A, Sharma R. Harnessing the power of AI: a comprehensive review of its impact and challenges in nursing science and healthcare. Cureus 2023 Nov;15(11):e49252. [doi: 10.7759/cureus.49252] [Medline: 38143615]
33. Duncan I, Loginov M, Ludkovski M. Testing alternative regression frameworks for predictive modeling of health care costs. N Am Actuar J 2016 Jan;20(1):65-87. [doi: 10.1080/10920277.2015.1110491]
34. Powers CA, Meyer CM, Roebuck MC, Vaziri B. Predictive modeling of total healthcare costs using pharmacy claims data: a comparison of alternative econometric cost modeling techniques. Med Care 2005 Nov;43(11):1065-1072. [doi: 10.1097/01.mlr.0000182408.54390.00] [Medline: 16224298]
35. Mera Flores AM, del Busto Bonifaz S, Bernal Sobrino JL. Evaluación de tres sistemas de ajuste de riesgo como predictores del consumo de medicamentos y productos sanitarios en unidades polivalentes de hospitalización [Article in Spanish]. Rev Esp Salud Pública 2016;90:e40018 [FREE Full text]
36. Wells AR, Guo X, Coberley CR, Pope JE. Integrating well-being information and the multidimensional adaptive prediction process to estimate individual-level future health care expenditure levels. Popul Health Manag 2016 Dec;19(6):429-438. [doi: 10.1089/pop.2015.0184] [Medline: 27267664]
37. Harrison PL, Pope JE, Coberley CR, Rula EY. Evaluation of the relationship between individual well-being and future health care utilization and cost. Popul Health Manag 2012 Dec;15(6):325-330. [doi: 10.1089/pop.2011.0089] [Medline: 22356589]
38. Orji U, Ukwandu E. Machine learning for an explainable cost prediction of medical insurance. Mach Learn Appl 2024 Mar;15:100516. [doi: 10.1016/j.mlwa.2023.100516]
39. Jödicke AM, Zellweger U, Tomka IT, et al. Prediction of health care expenditure increase: how does pharmacotherapy contribute? BMC Health Serv Res 2019 Dec 11;19(1):953. [doi: 10.1186/s12913-019-4616-x] [Medline: 31829224]
40. Lee SM, Kang JO, Suh YM. Comparison of hospital charge prediction models for colorectal cancer patients: Neural network vs. decision tree models. J Korean Med Sci 2004 Oct;19(5):677-681. [doi: 10.3346/jkms.2004.19.5.677] [Medline: 15483343]
41. Jolicoeur-Martineau A, Fatras K, Kachman T. Generating and imputing tabular data via diffusion and flow-based gradient-boosted trees. Presented at: 27th International Conference on Artificial Intelligence and Statistics (AISTATS 2024); May 2-4, 2024; Valencia, Spain p. 1288-1296 URL: https://proceedings.mlr.press/v238/jolicoeur-martineau24a/jolicoeur-martineau24a.pdf [accessed 2026-08-13]
42. Garzon M, Deng LY, Kumar N. What is dimensionality reduction (DR)? In: Garzon M, Yang CC, Venugopal D, Kumar N, Jana K, Deng LY, editors. Dimensionality Reduction in Data Science: Springer; 2022:67-77. [doi: 10.1007/978-3-031-05371-9_3]
43. Rajkomar A, Oren E, Chen K, et al. Scalable and accurate deep learning with electronic health records. NPJ Digit Med 2018;1:18. [doi: 10.1038/s41746-018-0029-1] [Medline: 31304302]
44. Tamang S, Milstein A, Sørensen HT, et al. Predicting patient "cost blooms" in Denmark: a longitudinal population-based study. BMJ Open 2017 Jan 11;7(1):e011580. [doi: 10.1136/bmjopen-2016-011580] [Medline: 28077408]

Abbreviations

ACG: Adjusted Clinical Group
AMG: Adjusted Morbidity Group
ATC: Anatomical Therapeutic Chemical
AUC: area under the curve
CRG: Clinical Risk Group
FHP: Functional Health Patterns
ICD-10: *International Classification of Diseases, 10th Revision*
ML: machine learning
MSE: mean squared error
NANDA: North American Nursing Diagnosis Association
NIC: Nursing Interventions Classification
NOC: Nursing Outcomes Classification
XGBoost: extreme gradient boosting

XSL•FO
RenderX

Edited by E Borycki; submitted 16.Feb.2026; peer-reviewed by SM Rashidul Hasan, SS Chai; revised version received 11.May.2026; accepted 17.Jun.2026; published 04.Sep.2026.

Please cite as:
Company-Sancho MC, González-Chordá VM, Orts-Cortés MI
Nursing Process Data for Health Care Cost Prediction Using Machine Learning: Longitudinal Study
JMIR Nursing 2026;9:e93638
URL: https://nursing.jmir.org/2026/1/e93638
doi:10.2196/93638

© María Consuelo Company-Sancho, Víctor M González-Chordá, María Isabel Orts-Cortés. Originally published in JMIR Nursing (https://nursing.jmir.org), 4.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predictors of Treatment Adherence in Outpatients With Chronic Obstructive Pulmonary Disease: Cross-Sectional Study

Hien Nguyen[1,2,3*], RN, PhD; Trieu Nguyen[4*], RN, PhD; Hoai Nguyen[4*], RN, MSN; Giang Luong[5*], RN, MSN; Doanh Vu[5*], RN, MNS; Thong Tran[6,7,8*], PhD, MD

[1]Dermatology and Burn Department, Bach Mai Hospital, Hanoi, Vietnam

[2]Nam Dinh University of Nursing, Ninh Binh, Vietnam

[3]Tan Tao University, Tay Ninh Province, Vietnam

[4]Faculty of Nursing, Da Nang University of Medical Technology and Pharmacy, 99 Hung Vuong street, Hai Chau District, Da Nang, Vietnam

[5]Faculty of Nursing, Hai Phong University of Medicine and Pharmacy, Hai Phong, Vietnam

[6]Center for Emergency Medicine, Bach Mai Hospital, Hanoi, Vietnam

[7]Department of Emergency and Critical Care Medicine, Hanoi Medical University, Hanoi, Vietnam

[8]Faculty of Medicine, University of Medicine and Pharmacy, Vietnam National University, Hanoi, Vietnam

[*]all authors contributed equally

Corresponding Author:
Trieu Nguyen, RN, PhD
Faculty of Nursing, Da Nang University of Medical Technology and Pharmacy, 99 Hung Vuong street, Hai Chau District, Da Nang, Vietnam

Abstract

Background: Gaps in knowledge, inhalation skills, and adherence continue to hinder effective chronic obstructive pulmonary disease (COPD) management.

Objective: The study was conducted to describe the status of COPD-related knowledge, practical skills for inhalation and breathing exercises, and treatment adherence among patients with COPD and to identify the relationships between these variables.

Methods: A cross-sectional study was conducted among 420 outpatients with COPD. Data were collected in 2 rounds, including interviews to assess patients' knowledge and treatment adherence and checklist-based observations of patients' practical skills. Multivariable logistic regression analysis was used to evaluate the associations between knowledge and practical skills and the status of treatment adherence.

Results: A total of 420 outpatients with COPD were included; 331 (78.8%) participants had inadequate COPD-related knowledge, 268 (63.8%) demonstrated adequate practical skills, and 155 (36.9%) achieved overall treatment adherence. In multivariable logistic regression, higher knowledge scores were associated with better treatment adherence (adjusted odds ratio [aOR] 1.20, 95% CI 1.06 - 1.31; *P*=.001). Participants who correctly performed inhaler-use techniques were more likely to be adherent than those with incorrect technique (aOR 3.10, 95% CI 2.10 - 4.70; *P*=.001). Likewise, adequate breathing-exercise skills were associated with higher adherence (aOR 1.70, 95% CI 1.08 - 2.71; *P*=.02).

Conclusions: Higher COPD-related knowledge and better practical skills in inhaler use and breathing exercises were associated with better treatment adherence among outpatients with COPD. Because of the cross-sectional design, these findings indicate associations rather than causal relationships. These observed associations may help inform future intervention studies evaluating patient education and skill-based training, although their effectiveness should be confirmed in longitudinal or interventional studies.

(JMIR Nursing 2026;9:e83132) doi:10.2196/83132

KEYWORDS

breathing exercises; chronic obstructive pulmonary disease; COPD; inhaler use; knowledge; treatment adherence

Introduction

Chronic obstructive pulmonary disease (COPD) is becoming a major global health challenge. According to the 2024 report of the Global Initiative for Chronic Obstructive Lung Disease (GOLD), there were about 3 million deaths globally each year, and by 2060, the world is estimated to have about 5.4 million deaths related to COPD [1]. In Vietnam, the prevalence of COPD among people aged 40 years and older is 4.2%, and approximately 1.3 million people with COPD currently require diagnosis and treatment [2]. Because of the prevalence,

XSL•FO
RenderX

prolonged progression, high treatment costs, and complex complications of COPD, this disease places a substantial burden on patients and health care systems worldwide and is associated with high morbidity and mortality [1].

Poor treatment adherence in COPD may be influenced by multiple factors, including inadequate disease-related knowledge, limited confidence in treatment, incorrect inhaler technique, and insufficient understanding of symptom management and breathing exercises [3]. These barriers may contribute to exacerbations, prolonged hospitalization, and poorer quality of life [1]. One of the major challenges in COPD management is ensuring that patients receive adequate education and support to understand and implement recommended self-management practices. Teaching and using an action plan in the event of worsening respiratory symptoms have been shown to reduce severity and shorten recovery time [4,5]. For patients to participate effectively in self-management programs, they must understand their disease as well as how to use medications and perform breathing exercises properly at home.

Treatment adherence is critical to the effective management of COPD and is key to addressing the growing burden of the disease [4]. For patients with COPD in the stable stage, adherence to medication and respiratory rehabilitation through the practice of breathing exercises are key to managing and alleviating symptoms, improving lung function, respiratory muscle strength, and exercise capacity, as well as reducing dyspnea and improving the quality of life [1]. However, many studies have shown that long-term adherence to chronic disease treatment is low, and it is estimated that only 50% of patients comply with treatment [6]. In addition, after the COVID-19 pandemic was brought under control in Vietnam, bed occupancy in respiratory departments of healthcare facilities has increased, coinciding with a large number of hospitalized patients with COPD exacerbations. One of the main reasons for this situation is that patients do not adhere to treatment.

To date, evidence regarding the relationship between treatment adherence, disease-related knowledge, and practical self-management skills among outpatients with COPD in Vietnam remains limited. This study aimed to assess COPD-related knowledge, practical skills in inhaler use and breathing exercises, and treatment adherence among outpatients with COPD and to examine the associations between knowledge, practical skills, and treatment adherence.

Methods

Study Design and Settings

A cross-sectional study was conducted among outpatients with COPD who were recruited from 3 hospitals, including Hospital C, Da Nang Hospital, and Da Nang Hospital for Lung Diseases, from April 2021 to August 2022 in Vietnam. This manuscript was prepared in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guideline for cross-sectional studies.

Participants and Sampling

The sample size was calculated based on the following formula:

$$n=(Z1-\alpha 2)2\times P\times(1-P)d2$$

Where $Z1-\alpha/2$ is the corresponding value at the selected confidence level (equal to 1.96 if the confidence level is 95%), P is the treatment adherence rate (50%) [7], and d is the absolute error of 5%.

On the basis of the formula, the required sample size for the study was 384. In addition, the research team estimated that about 10% of participants would withdraw after signing the consent form or would not complete the survey form; therefore, the minimum sample size was set at 420. All participants were recruited using convenience sampling.

The inclusion criteria were patients diagnosed with COPD; receiving home drug therapy with inhaled medications; able to speak, read, and understand Vietnamese; and willing to participate in the study.

The exclusion criteria were patients with a history of bronchial asthma, allergic rhinitis, lung surgery, or other acute respiratory diseases; experiencing COPD exacerbations or exacerbations of comorbidities or who had experienced medication changes within the previous 3 months; or with mental disorders or other serious illnesses.

On the basis of outpatient clinic records from the 3 participating hospitals, a list of 653 patients with a physician-confirmed diagnosis of COPD who attended follow-up outpatient visits and were receiving inhaled maintenance treatment between April 2021 and August 2022 was compiled. The research team contacted these patients by telephone, explained the purpose and importance of the study, and invited them to participate. A total of 497 patients agreed to participate. Reasons for nonparticipation mainly included advanced age, poor health status, living far from the hospital, or lack of family support for transportation. After applying the inclusion and exclusion criteria at the time of data collection, 420 participants were enrolled in the final analysis.

Measures

The questionnaire included 4 parts.

Part 1: Sociodemographic and Clinical Characteristics

The demographic questionnaire designed by the research team comprised 7 items (age, sex, education level, employment, years since diagnosis, smoking, and disease severity). Disease severity was assessed using the modified Medical Research Council (mMRC) questionnaire and the COPD Assessment Test (CAT). Both instruments have been translated into Vietnamese and officially published by the Ministry of Health of Vietnam and have been widely used in COPD research in Vietnam [8]. Patients were classified as having mild disease when mMRC scores were 0 to 1 and CAT scores were <10 and were classified as severe when mMRC scores were ≥2 and CAT scores were ≥10.

Part 2: Knowledge About COPD

Knowledge about COPD was measured using the Chronic Obstructive Pulmonary Disease Knowledge Questionnaire developed by Maples et al [9]. This scale consisted of 13 questions; each item was scored as 1 (correct) or 0 (incorrect

XSL•FO
RenderX

or unknown). Overall, patients were considered to have adequate knowledge if their total score was ≥10, whereas a total score of <10 indicated inadequate disease knowledge [9]. The Vietnamese version of the scale has been evaluated in 90 people with COPD, and the results showed good content validity and internal consistency (Kuder-Richardson formula 20 was 0.883 and 0.866 before and after the test, respectively). The test-retest reliability after 4 weeks was high (r=0.996; P<.001) [10].

Part 3: Practical Skills Assessment

Patients' practical skills were assessed using 2 observational checklists developed based on the 2020 Vietnamese Ministry of Health guideline for the diagnosis and treatment of COPD [11]. These assessments covered 2 domains: inhaler-use technique and breathing-exercise technique.

The inhaler-use checklist assessed patients' ability to use metered-dose inhalers (MDIs) with 6 steps and a Turbuhaler in 6 steps. For each step, patients received 2 points for correct performance, 1 point for incorrect performance, and 0 points for failure. According to Ngo et al [8], patients were classified as having "passed" if they performed all steps correctly, and as having "failed" if at least 1 step was performed incorrectly or not performed.

The checklist for breathing exercises included pursed-lip breathing (4 steps) and diaphragmatic breathing (4 steps). For each step performed correctly, patients received 2 points, while incorrect performance received 1 point, and wrong or failed steps received 0 points. The overall assessment was classified as "passed" if the patient performed all steps of the checklist correctly and fully, and "failed" when at least one step was wrong or not performed [12].

Part 4: Treatment Adherence

Treatment adherence was defined as the extent to which patients followed health care providers' recommendations regarding 2 main components of COPD treatment: inhaled medication use and breathing exercises.

Adherence to inhaled medication was assessed using the 10-item Test of Adherence to Inhalers (TAI-10), developed by Plaza et al [13]. Each item was rated on a 5-point Likert scale ranging from 1 ("worst adherence") to 5 ("best adherence"), yielding a total score from 10 to 50. Patients with a total score of 46 to 50 were classified as adherent to inhaled medication, whereas those with a score of ≤45 were classified as nonadherent. The TAI-10 has demonstrated good reliability in previous studies, with a Cronbach α of 0.871 and a test-retest reliability coefficient of 0.832 (P<.01) [14].

Adherence to breathing exercises was assessed based on both the correctness of breathing-exercise performance and the maintenance of daily practice frequency according to recommendations from the GOLD and the Vietnamese Ministry of Health COPD guideline. A checklist was used to evaluate whether patients correctly performed the prescribed breathing exercises. Patients were classified as adherent to breathing exercises only if they met both of the following criteria: (1) correctly performed all required steps of the breathing exercises and (2) maintained daily practice for at least 10 to 15 minutes, with a gradual increase according to their physical capacity. Patients who performed the breathing exercises correctly but did not maintain daily practice or who practiced daily but did not perform the exercises correctly were classified as nonadherent to breathing exercises. The assessment criteria for breathing-exercise adherence were developed based on guideline recommendations and expert consensus from 5 respiratory physicians and nurses at Hospital C, Da Nang.

Overall treatment adherence was defined as adherence to both inhaled medication and breathing exercises. Patients were classified as adherent overall only if they were adherent to both components; those who were adherent to only one component or to neither component were classified as nonadherent overall. The complete set of questionnaire items is presented in Multimedia Appendix 1.

Statistical Analysis

The study used SPSS (version 23.0; IBM Corp) for data analysis. Descriptive statistics were used to describe variables. Univariate analysis (chi-square test and independent-sample 2-tailed t test) was performed to examine the association between independent variables and treatment adherence. Variables with P<.05 (or P<.20 where appropriate) in univariate analysis were considered for inclusion in the multivariable model. Multivariable logistic regression analysis was then conducted to identify independent factors associated with treatment adherence, including demographic characteristics, clinical variables, knowledge, and practical skills. Adjusted odds ratios (aORs) with 95% CIs were reported. P<.05 was considered statistically significant.

Ethical Considerations

The Protocol Approval Committee and the Medical Ethics Committee of Nam Dinh University of Nursing approved the study (1681/GCN-HDDD) on August 2, 2021. Participants provided written informed consent before participation.

Results

General Characteristics and Clinical Features of Patients With COPD

Among 420 participants, most were male (n=335, 79.8%). The mean age was 59.5 (SD 8.2) years. Most participants had a post–high school education level (n=346, 82.4%). Regarding employment status, 202 (48.1%) participants were retired, whereas only 33 (7.8%) were engaged in free trade. The mean time since diagnosis was 11.39 (SD 5.46) years, and 51.4% (n=216) of the participants were smokers (Table 1).

XSL•FO
RenderX

Table . General characteristics and clinical features of patients with chronic obstructive pulmonary disease (COPD; n=420).

Contents	Patients
Age (y), mean (SD)	59.5 (8.2)
Sex, n (%)	
Male	335 (79.8)
Female	85 (20.2)
Education level, n (%)	
High school or under	74 (17.6)
Post–high school	346 (82.4)
Employment, n (%)	
Worker and farmer	51 (12.2)
Civil servants	134 (31.9)
Retired	202 (48.1)
Free trade	33 (7.8)
Time since diagnosis (y), mean (SD)	11.39 (5.46)
Time since diagnosis (y), n (%)	
<5	84 (20)
5 - 10	152 (36.2)
>10	184 (43.8)
Smoking, n (%)	
No	204 (48.6)
Yes	216 (51.4)
Disease severity, n (%)	
Mild (mMRC[a] score 0 - 1 and CAT[b] score <10)	22 (5.2)
Severe (mMRC score ≥2 and CAT score ≥10)	398 (94.8)

[a]mMRC: modified Medical Research Council.

[b]CAT: COPD Assessment Test.

Patients' Knowledge About COPD

Among 420 participants, 331 (78.8%) had inadequate knowledge about COPD, whereas 89 (21.2%) had adequate knowledge. Item-level analysis showed that knowledge gaps were mainly related to smoking cessation and medication use (Table 2). Specifically, the items with the lowest correct response rates were "Quitting smoking will prevent the disease from getting worse" (n=53, 12.6%) and "Medications prevent the disease from getting worse" (n=60, 14.3%). The item with the highest correct response rate was "Shortness of breath is a symptom of COPD" (n=375, 89.2%; Multimedia Appendix 2).

Table . Patient's knowledge, practice skills, and treatment adherence about chronic obstructive pulmonary disease (COPD; n=420).

Contents	Correct answers, n (%)
Total score of knowledge about COPD	
Adequate knowledge	89 (21.2)
Inadequate knowledge	331 (78.8)
Overall practical skills performance	
Passed	268 (63.8)
Failed	152 (36.2)
Overall treatment adherence	
Adherence	155 (36.9)
Nonadherence	265 (63.1)

XSL•FO
RenderX

Practical Skills of Patients With COPD

Overall, of 420 participants, 152 (36.2%) failed the overall practical skills assessment, whereas 268 (63.8%) passed (Table 2). For the MDI technique, 112 (26.7%) participants failed, with the most common error being failure to exhale slowly before inhalation (77/112, 68.8%). For the Turbuhaler technique, 40 (9.5%) participants failed, and incorrect inhalation technique was the most frequent error (28/40, 70%).

Regarding breathing-exercise skills, 278 (66.2%) participants correctly performed both breathing exercises during the practical skill assessment, whereas 142 (33.8%) did not (Multimedia Appendix 3).

Treatment Adherence of Patients With COPD

Overall treatment adherence was observed in 155 (36.9%) of 420 participants, whereas 265 (63.1%) were classified as nonadherent (Table 2). Adherence to inhaled medication was observed in 234 (55.7%) participants, and adherence to breathing exercises was observed in 265 (63.1%) participants (Multimedia Appendix 4).

Factors Associated With Treatment Adherence of Patients With COPD

In univariate analysis, disease severity, smoking status, knowledge scores, inhaler-use skills, and breathing-exercise skills were significantly associated with treatment adherence (P<.05). Age, sex, education level, employment status, and duration of illness were not significantly associated with treatment adherence. Variables with P<.20 in univariate analysis were entered into the multivariable logistic regression model (Multimedia Appendix 5).

Knowledge scores were significantly associated with treatment adherence (adjusted odds ratio [aOR] 1.20, 95% CI 1.06 - 1.31; P=.002). Participants who passed the inhaler-use skill assessment had higher odds of overall treatment adherence than those who failed (aOR 3.10, 95% CI 2.01 - 4.75; P=.001). Similarly, participants who passed the breathing-exercise skill assessment had higher odds of overall treatment adherence than those who failed (aOR 1.70, 95% CI 1.08 - 2.71; P=.02; Table 3). Among these factors, correct inhaler-use technique showed the strongest association with treatment adherence.

Table . Factors associated with treatment adherence in the multivariable logistic regression model[a].

Factors	B	aOR[b] (95% CI)	*P* value
Age (years)	−0.02	0.98 (0.95 - 1.01)	.14
Sex			
Male	−0.41	0.66 (0.34 - 1.28)	.22
Female	Reference	1.0	—[c]
Disease severity			
Mild	0.74	2.10 (0.88 - 5.01)	.09
Severe	Reference	1.0	—
Smoking			
Yes	0.6	1.8 (0.85 - 2.23)	.08
No	Reference	1.0	—
Knowledge score	0.17	1.2 (1.06 - 1.31)	.002
Practical skills			
Inhaler-use skills			
Passed	1.2	3.1 (2.01 - 4.75)	<.001
Failed	Reference	1.0	—
Breathing-exercise skills			
Passed	0.5	1.7 (1.08 - 2.71)	.02
Failed	Reference	1.0	—

[a]Goodness-of-fit test of the logistic regression model: Hosmer-Lemeshow test, χ^2_8=7.6; P>.05.

[b]aOR: adjusted odds ratio.

[c]Not applicable.

Discussion

Principal Findings

This study found that COPD-related knowledge, observed practical skills, and treatment adherence among outpatients with COPD were generally suboptimal. Nearly 4 in 5 participants lacked adequate COPD-related knowledge, more than one-third demonstrated poor practical skills, and only 36.9% (155/420) achieved overall treatment adherence. In addition, better knowledge and adequate inhaler-use and breathing-exercise

XSL•FO
RenderX

skills were significantly associated with higher treatment adherence.

COPD has become a major cause of morbidity and mortality worldwide. Currently, COPD cannot be completely cured, but early treatment and adherence to treatment recommendations can reduce symptoms, slow the progression of lung damage, and improve patients' quality of life [15].

The study found that 78.8% (331/420) of patients with COPD lacked knowledge about the disease; this proportion was higher than that reported by Yang et al [16]. In addition, when comparing the correct response rates, our study and the study by Yang et al [16] showed similar patterns regarding the highest and lowest correct response rates. Specifically, the highest correct response rate was for "shortness of breath as a symptom" at 89.2% (375/420) in our study and 77.7% in the study by Yang et al [16]. This similarity can be explained by dyspnea being one of the most typical and frequent symptoms of COPD. Most people with COPD experience dyspnea when they are physically active, even in the early stages, and as the disease progresses, this problem gradually increases in intensity and severity and may occur even at rest. Two items with the lowest correct response rates were "Medications to prevent the disease from getting worse" with a correct response rate of only 14.3% (60/420) in our study compared with 10.7% in the study by Yang et al [16] and "Quitting smoking will prevent the disease from getting worse" with 12.6% (53/420) and 11.8%, respectively [16]. These knowledge assessments were related to the progression of COPD. Although most patients with COPD know that regular medication use and smoking cessation are beneficial for disease management, many do not seem to realize that with COPD, lung function cannot be fully restored [16,17]. Currently, there are no drugs or methods that can completely reverse the decline in lung function in people with COPD [17]. These findings highlight areas in which patients demonstrated substantial knowledge gaps. Educational strategies addressing these topics could be evaluated in future intervention studies.

Our study also found that 63.8% (268/420) of patients demonstrated adequate inhaler-use skills, which was higher than some studies conducted in Italy (50%) and in Korea (46.7%) but lower than that reported in a study in Pakistan (93.7%) [18-20]. Among inhaled bronchodilators used by patients, the most common errors were related to the use of the MDI and Turbuhaler, with error rates of 26.7% and 9.5%, respectively. Compared with previous studies, these rates were lower than those reported in the study in Pakistan (60% and 41%, respectively) [20] and the study by Molimard et al [21] (46.9% and 32.1%, respectively). In addition, when assessing the errors that patients often made when using the MDI and Turbuhaler, our study showed that the highest rate was for "Haven't exhaled very slowly yet," at 68.8% (77/112) for the MDI and 50% (20/40) for the Turbuhaler. This rate was higher than that reported by Molimard et al [21] in a study of 2935 patients with COPD hospitalized with acute exacerbations.

Most patients practiced both breathing exercises adequately (278/420, 66.2%); almost all patients passed the skills assessment when performing pursed-lip breathing (346/420, 82.4%), while diaphragm breathing exercises were still low, with 56%. This may be because diaphragmatic breathing involves more complex steps than pursed-lip breathing. These findings suggest that breathing exercises remain an important area for patient education; however, whether educational interventions improve adherence should be confirmed in future interventional studies. Additionally, our results were similar to previous research [22,23].

Treatment adherence in this study was defined as adherence to 2 key components of the prescribed COPD treatment regimen: inhaled medication use and breathing exercises. The overall treatment adherence rate was low, with only 36.9% (155/420) of participants meeting both criteria. One possible explanation is that many previous studies assessed adherence primarily in terms of medication use alone, whereas our study applied a broader and more stringent definition that required adherence to both inhaled medication and breathing exercises. This composite approach may partly explain the lower overall adherence rate observed in our sample. In addition, the study was conducted during the COVID-19 pandemic period, when limited access to health care services, reduced follow-up visits, and fear of infection may also have negatively affected treatment adherence. When inhaled medication adherence was considered separately, 55.7% (234/420) of patients were classified as adherent according to the TAI-10. This finding differs from those reported in previous studies. A study in Greece reported a high rate of nonadherence to inhaled medication (74.1%) [24], whereas studies from several European, Latin American, and Asian countries reported lower nonadherence rates [25-27]. These differences may be related to variations in study populations, health care systems, adherence assessment tools, and the operational definitions of adherence used across studies. The inclusion of breathing exercises in the definition of overall treatment adherence was based on the clinical management of COPD, which involves not only pharmacological treatment but also nonpharmacological interventions. In the present study, adherence to breathing exercises was defined not merely by correct technique but by both correct performance and maintenance of daily practice according to the prescribed recommendations. These findings suggest that inhaler technique and breathing-exercise skills may represent potential targets for future nursing interventions. However, the present cross-sectional study cannot determine whether such interventions improve treatment adherence.

In addition, higher COPD knowledge was significantly associated with greater odds of treatment adherence, suggesting that patients who better understand their disease and its management may be more likely to follow prescribed treatment recommendations. This finding is consistent with previous studies [28]. Participants who passed the inhaler-use skill assessment had 3.1-fold greater odds of overall treatment adherence than those who failed, representing the strongest association observed in the present study and consistent with previous findings [29]. Similarly, participants who passed the breathing-exercise skill assessment had 1.7 times higher odds of overall treatment adherence than those who failed, which is also in line with earlier research [16]. These findings should not be interpreted as meaning that practical skills are equivalent to adherence. Rather, they suggest that correct performance of

XSL•FO
RenderX

inhaler-use techniques and breathing exercises may reflect a patient's ability to understand, retain, and implement treatment instructions, which in turn may support better adherence behavior. In other words, practical skills in this study should be considered correlates of treatment adherence rather than direct measures of adherence itself. Because this was a cross-sectional study, the temporal direction of these associations cannot be determined, and causal relationships cannot be inferred.

Although knowledge and practical skills were significantly associated with treatment adherence in the multivariable model, these associations should be interpreted with caution. In the context of COPD management in Vietnam, adherence behavior is likely influenced not only by patients' knowledge and self-management skills but also by broader contextual factors such as medication affordability, health insurance coverage, family support, and the quality of patient–health care provider communication. These factors may affect a patient's ability to obtain medications, attend follow-up visits, understand treatment instructions, and maintain regular breathing exercises and may therefore be closely intertwined with both treatment adherence and the measured knowledge and skill variables. Because these contextual variables were not collected and could not be included in the regression model, residual confounding and omitted variable bias cannot be excluded. As a result, the independent contributions of COPD knowledge, inhaler-use skills, and breathing-exercise skills may have been overestimated to some extent. Therefore, the present findings should be interpreted as associations within the measured model rather than as evidence of fully independent causal effects.

Strengths and Limitations

Treatment adherence in this study was assessed using a combination of methods, including the self-reported TAI-10 for inhaled medication adherence and a structured assessment of breathing-exercise adherence. In addition, patients' inhaler-use and breathing-exercise skills were evaluated by direct observation using standardized checklists, which provided complementary information on practical performance.

However, this study has several limitations. First, data collection was conducted shortly after the COVID-19 pandemic was brought under control in Da Nang City, which created challenges in accessing and recruiting the target population. Although the final sample included 420 participants from 3 hospitals, caution is still needed when interpreting the representativeness of the findings. Second, treatment adherence in COPD is multidimensional and may include not only adherence to inhaled medications and breathing exercises but also smoking cessation, dietary recommendations, physical activity, and other self-management behaviors. Due to resource constraints and study feasibility, the present study focused only on adherence to inhaled medications and breathing exercises. Therefore, the findings may not fully reflect overall adherence to all recommended COPD management behaviors. Third, the study did not assess the quality of physician-patient communication or patient–health care provider relationships, which may substantially influence patients' understanding of treatment, motivation, and adherence behaviors. Future studies should incorporate communication-related factors to provide a more comprehensive explanation of treatment adherence among patients with COPD. Fourth, some invited patients declined participation because of older age, frailty, transportation difficulties, or limited family support. In addition, patients who had experienced COPD exacerbations, exacerbations of comorbidities, or medication changes within the previous 3 months were excluded from the study. These factors may have introduced selection bias toward a relatively more stable and healthier subgroup of patients and may therefore have led to an overestimation of treatment adherence compared with the broader COPD population receiving follow-up care. Fifth, the study did not specify a minimum duration since COPD diagnosis or initiation of inhaled treatment. As a result, participants may have differed in their level of experience with disease management, inhaler use, and breathing exercises, which could have affected the homogeneity of the sample and should be considered when interpreting the findings. Finally, because this was a cross-sectional study, temporal relationships between COPD knowledge, practical skills, and treatment adherence could not be established. Therefore, causal inferences cannot be drawn. Longitudinal cohort studies and intervention trials are required to determine whether improving knowledge and practical skills leads to better treatment adherence.

Conclusions

Higher COPD-related knowledge and better practical skills in inhaler use and breathing exercises were associated with better treatment adherence among outpatients with COPD. Because this study used a cross-sectional design, these findings demonstrate associations rather than causal relationships. The observed associations indicate that patient education and skill-based training warrant further investigation as potential nursing interventions; however, their effectiveness in improving treatment adherence should be confirmed in longitudinal cohort studies and randomized controlled trials before causal conclusions are drawn.

Funding

The authors declared no financial support was received for this work.

Authors' Contributions

Conceptualization: TN
Data curation: TN, Hoai Nguyen, TT, Hien Nguyen, GL, DV
Formal analysis: TN, Hoai Nguyen, TT, Hien Nguyen, GL, DV
Validation: TN, Hoai Nguyen, TT, Hien Nguyen, GL, DV

XSL•FO
RenderX

Writing—original draft: TN, Hoai Nguyen, TT, Hien Nguyen, GL, DV
Writing—review and editing: TN, Hoai Nguyen, TT, Hien Nguyen, GL, DV

Conflicts of Interest

None declared.

Multimedia Appendix 1
Responses to the Test of Adherence to Inhalers questionnaire.
[DOCX File, 21 KB - nursing_v9i1e83132_app1.docx]

Multimedia Appendix 2
Practical skills of patients with chronic obstructive pulmonary disease.
[DOCX File, 20 KB - nursing_v9i1e83132_app2.docx]

Multimedia Appendix 3
Treatment adherence of patients with chronic obstructive pulmonary disease.
[DOCX File, 19 KB - nursing_v9i1e83132_app3.docx]

Multimedia Appendix 4
The level of treatment adherence based on the clinical characteristics of patients with chronic obstructive pulmonary disease.
[DOCX File, 24 KB - nursing_v9i1e83132_app4.docx]

Multimedia Appendix 5
Study questionnaire and assessment tools used for patients with chronic obstructive pulmonary disease (COPD), including COPD knowledge questionnaire, inhaler-use technique checklists, breathing-exercise practice checklists, and adherence to inhaled medications assessed using the Test of Adherence to Inhalers.
[DOCX File, 25 KB - nursing_v9i1e83132_app5.docx]

References

1. Global strategy for prevention, diagnosis and management of COPD: 2024 report. : Global Initiative for Chronic Obstructive Lung Disease; 2024 URL: https://goldcopd.org/2024-gold-report/ [accessed 2026-09-13]
2. GBD 2023 Asia Chronic Respiratory Disease Collaborators. Burden of chronic respiratory disease in Asia, 1990-2023: a systematic analysis for the Global Burden of Disease Study 2023. Lancet Respir Med 2026 Mar;14(3):233-255. [doi: 10.1016/S2213-2600(25)00404-7] [Medline: 41579872]
3. Świątoniowska N, Chabowski M, Polański J, Mazur G, Jankowska-Polańska B. Adherence to therapy in chronic obstructive pulmonary disease: a systematic review. Adv Exp Med Biol 2020;1271:37-47. [doi: 10.1007/5584_2019_477] [Medline: 32016912]
4. Alharbi MG, Kalra HS, Suri M, et al. Pulmonary rehabilitation in management of chronic obstructive pulmonary disease. Cureus 2021 Oct;13(10):e18414. [doi: 10.7759/cureus.18414] [Medline: 34646587]
5. Bischof AY, Cordier J, Vogel J, Geissler A. Medication adherence halves COPD patients' hospitalization risk - evidence from Swiss health insurance data. NPJ Prim Care Respir Med 2024 Mar 7;34(1):1. [doi: 10.1038/s41533-024-00361-2] [Medline: 38453930]
6. Dunbar-Jacob J, Zhao J. Medication adherence measurement in chronic diseases: a state-of-the-art review of the literature. Nurs Rep 2025 Oct 16;15(10):370. [doi: 10.3390/nursrep15100370] [Medline: 41149685]
7. Adherence to long-term therapies: evidence for action. World Health Organization. 2003. URL: https://iris.who.int/items/bf8058c0-03b2-4b47-838f-5534849927fb [accessed 2024-05-15]
8. Ngo CQ, Phan DM, Vu GV, et al. Inhaler technique and adherence to inhaled medications among patients with acute exacerbation of chronic obstructive pulmonary disease in Vietnam. Int J Environ Res Public Health 2019 Jan 10;16(2):185. [doi: 10.3390/ijerph16020185] [Medline: 30634631]
9. Maples P, Franks A, Ray S, Stevens AB, Wallace LS. Development and validation of a low-literacy Chronic Obstructive Pulmonary Disease knowledge Questionnaire (COPD-Q). Patient Educ Couns 2010 Oct;81(1):19-22. [doi: 10.1016/j.pec.2009.11.020] [Medline: 20044232]
10. Nguyen TT, Upvall M, Nguyen HT, Nguyen HT, Tran LV, Nguyen PT. Validity and reliability of the Vietnamese version of Chronic Obstructive Pulmonary Disease-Questionnaire for measuring knowledge of chronic obstructive pulmonary disease patients. J Health Sci Med Res 2023;41(1). [doi: 10.31584/jhsmr.2022881]

XSL•FO
RenderX

11. Ministry of Health's 2020 guidelines for the diagnosis and treatment of asthma (part 1) [Article in Vietnamese]. B nh hen. 2020. URL: https://www.benhhen.vn/tintuc/huong-dan-chan-doan-va-dieu-tri-hen-phe-quan-cua-bo-y-te-nam-2020-ky-1.html [accessed 2024-05-15]
12. Thao LT. The survey on adherence rate of breathing exercises and relationship with quality of life in patients with chronic obstructive pulmonary disease. Int J Collab Res Intern Med Public Health 2022;14(1) [FREE Full text] [doi: 10.35248/1840-4529.22.14.341]
13. Plaza V, Fernández-Rodríguez C, Melero C, et al. Validation of the "Test of the Adherence to Inhalers" (TAI) for asthma and COPD patients. J Aerosol Med Pulm Drug Deliv 2016 Apr;29(2):142-152. [doi: 10.1089/jamp.2015.1212] [Medline: 26230150]
14. Muneswarao J, Hassali MA, Ibrahim B, et al. Translation and validation of the Test of Adherence to Inhalers (TAI) questionnaire among adult patients with asthma in Malaysia. J Asthma 2021 Sep;58(9):1229-1236. [doi: 10.1080/02770903.2020.1776728] [Medline: 32493083]
15. Liu W, Liu Y, Li X. Impact of exercise capacity upon respiratory functions, perception of dyspnea, and quality of life in patients with chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis 2021;16:1529-1534. [doi: 10.2147/COPD.S311221] [Medline: 34103910]
16. Yang H, Wang H, Du L, Wang Y, Wang X, Zhang R. Disease knowledge and self-management behavior of COPD patients in China. Medicine (Baltimore) 2019 Feb;98(8):e14460. [doi: 10.1097/MD.0000000000014460] [Medline: 30813148]
17. Andreas S, Hering T, Mühlig S, Nowak D, Raupach T, Worth H. Smoking cessation in chronic obstructive pulmonary disease: an effective medical intervention. Dtsch Arztebl Int 2009 Apr;106(16):276-282. [doi: 10.3238/arztebl.2009.0276] [Medline: 19547629]
18. Vitacca M, Paneroni M, Fracassi M, et al. Inhaler technique knowledge and skills before and after an educational program in obstructive respiratory disease patients: a real-life pilot study. Pulmonology 2023;29(2):130-137. [doi: 10.1016/j.pulmoe.2020.04.010] [Medline: 33268032]
19. Jang JG, Ahn JH, Jin HJ. Incidence and prognostic factors of respiratory viral infections in severe acute exacerbation of chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis 2021;16:1265-1273. [doi: 10.2147/COPD.S306916] [Medline: 33994783]
20. Sanaullah T, Khan S, Masoom A, Mandokhail ZK, Sadiqa A, Malik MI. Inhaler use technique in chronic obstructive pulmonary disease patients: errors, practices and barriers. Cureus 2020 Sep 21;12(9):e10569. [doi: 10.7759/cureus.10569] [Medline: 33101814]
21. Molimard M, Raherison C, Lignot S, et al. Chronic obstructive pulmonary disease exacerbation and inhaler device handling: real-life assessment of 2935 patients. Eur Respir J 2017 Feb;49(2):1601794. [doi: 10.1183/13993003.01794-2016] [Medline: 28182569]
22. Sørensen D, Svenningsen H. Adherence to home-based inspiratory muscle training in individuals with chronic obstructive pulmonary disease. Appl Nurs Res 2018 Oct;43:75-79. [doi: 10.1016/j.apnr.2018.07.005] [Medline: 30220368]
23. Le TT, Nguyen TK, Chu TG. Evaluation of the practice breathing exercises of patients with chronic obstructive pulmonary disease. Cantho J Med Pharm 2023(55):8-13. [doi: 10.58490/ctump.2022i55.339]
24. Moradkhani B, Mollazadeh S, Niloofar P, Bashiri A, Oghazian MB. Association between medication adherence and health-related quality of life in patients with chronic obstructive pulmonary disease. J Pharm Health Care Sci 2021 Nov 15;7(1):40. [doi: 10.1186/s40780-021-00222-x] [Medline: 34775992]
25. Schreiber J, Sonnenburg T, Luecke E. Inhaler devices in asthma and COPD patients - a prospective cross-sectional study on inhaler preferences and error rates. BMC Pulm Med 2020 Aug 20;20(1):222. [doi: 10.1186/s12890-020-01246-z] [Medline: 32819337]
26. Şanlıtürk D, Ayaz-Alkaya S. The effect of a theory of planned behavior education program on asthma control and medication adherence: a randomized controlled trial. J Allergy Clin Immunol Pract 2021 Sep;9(9):3371-3379. [doi: 10.1016/j.jaip.2021.03.060] [Medline: 33930620]
27. Bissell P, May CR, Noyce PR. From compliance to concordance: barriers to accomplishing a re-framed model of health care interactions. Soc Sci Med 2004 Feb;58(4):851-862. [doi: 10.1016/s0277-9536(03)00259-4] [Medline: 14672598]
28. Janjua S, Pike KC, Carr R, Coles A, Fortescue R, Batavia M. Interventions to improve adherence to pharmacological therapy for chronic obstructive pulmonary disease (COPD). Cochrane Database Syst Rev 2021 Sep 8;9(9):CD013381. [doi: 10.1002/14651858.CD013381.pub2] [Medline: 34496032]
29. Elander A, Gustafsson M. Inhaler technique and self-reported adherence to medications among hospitalised people with asthma and COPD. Drugs Real World Outcomes 2020 Dec;7(4):317-323. [doi: 10.1007/s40801-020-00210-x] [Medline: 33052539]

Abbreviations

aOR: adjusted odds ratio
CAT: COPD Assessment Test
COPD: chronic obstructive pulmonary disease

XSL•FO
RenderX

GOLD: Global Initiative for Chronic Obstructive Lung Disease
MDI: metered-dose inhaler
mMRC: modified Medical Research Council
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TAI-10: 10-item Test of Adherence to Inhalers

Edited by E Borycki; submitted 28.Aug.2025; peer-reviewed by T Yang, V Chandrasekaran; revised version received 30.Jul.2026; accepted 07.Sep.2026; published 25.Sep.2026.

Please cite as:
Nguyen H, Nguyen T, Nguyen H, Luong G, Vu D, Tran T
Predictors of Treatment Adherence in Outpatients With Chronic Obstructive Pulmonary Disease: Cross-Sectional Study
JMIR Nursing 2026;9:e83132
URL: https://nursing.jmir.org/2026/1/e83132
doi:10.2196/83132

© Hien Nguyen, Trieu Nguyen, Hoai Nguyen, Giang Luong, Doanh Vu, Thong Tran. Originally published in JMIR Nursing (https://nursing.jmir.org), 25.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Relational Playbook Nurse Leadership Development Program Using the Whistle Systems Employee Recognition Platform: Feasibility Mixed Methods Study

Marguerite Daus[1,2], RN, PhD; Brigid Connelly[2], MA; Drew Carter[3], MBA; Heather M Gilmartin[2,4], PhD

[1]Division of General Internal Medicine, Department of Medicine, University of Colorado Anschutz Medical Campus, 13001 E 17th Place, Aurora, CO, United States

[2]Denver/Seattle Center of Innovation for Veteran-Centered and Value Driven Care, United States Department of Veterans Affairs, VA Eastern Colorado Healthcare System, Aurora, CO, United States

[3]Whistle Systems, St. Louis, MO, United States

[4]Department of Health Systems, Management and Policy, School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO, United States

Corresponding Author:
Marguerite Daus, RN, PhD
Division of General Internal Medicine, Department of Medicine, University of Colorado Anschutz Medical Campus, 13001 E 17th Place, Aurora, CO, United States

Abstract

Background: Leadership development programs in health care often fail due to their lack of adaptability to the schedules of busy clinicians. This study addressed the need for scalable, flexible programs tailored to nurse leaders.

Objective: This study evaluated the acceptability, appropriateness, and feasibility of the Relational Playbook, an evidence-based leadership development program developed in the Veterans Health Administration delivered through the Whistle Systems employee recognition web application and mobile app.

Methods: A 1-year, single-team pilot was deployed using descriptive survey data and qualitative interview analysis. The Relational Playbook's educational content and interventions were hosted on the Whistle platform, which integrates behavioral science and gamification strategies. Content was delivered weekly via app-based nudge notifications and email. Engagement metrics included activity completion rates. User experience data were collected through weekly reflection surveys (with Likert-scale responses and open-text options); monthly check-ins; and a postimplementation acceptability, appropriateness, and feasibility survey and interview. Descriptive statistics summarized engagement levels and trends, and qualitative data were analyzed using content analysis to identify recurring concepts. Quantitative and qualitative data were analyzed sequentially for comprehensive insights.

Results: The section chief and 4 practicing cardiology nurse practitioners from a large academic medical center participated. The nurse practitioner section chief deemed the Whistle platform an acceptable, appropriate, and feasible technology for delivering the Relational Playbook content. They valued the weekly nudges, microlearning content, and flexibility of the web application and mobile app. The Relational Playbook content supported their personal growth and fostered positive shifts in attitudes toward work.

Conclusions: Delivering leadership development content through the Whistle platform is an acceptable approach to support the growth and well-being of busy nurse leaders. The small sample and absence of a comparison group limit generalizability.

(JMIR Nursing 2026;9:e79188) doi:10.2196/79188

KEYWORDS

leadership; nursing; mobile app; development; workforce

Introduction

The well-being of nurse leaders and the nursing workforce is an urgent concern worsened by the COVID-19 pandemic, increasing patient complexity, evolving hospital systems, and high workloads [1]. Nurse leaders play a pivotal role in improving nurse well-being, patient care, and clinical outcomes by managing frontline clinical staff [2]. However, leadership is a challenging role that requires skills to foster interdisciplinary teamwork, continuous learning, and high reliability. Most nurse leadership training occurs through face-to-face didactic education or on-the-job training that falls short of true leadership development [3]. Digital leadership programs are available; however, most lack scientific rigor and impact evaluation [4].

XSL•FO
RenderX

With many nurse leaders nearing retirement, developing the next generation is essential to sustaining the profession and ensuring high-quality care [5].

The Relational Playbook is an innovative leadership development program grounded in adult learning principles, including experiential learning [6] and situated learning theories [7]. The Relational Playbook is designed to equip frontline clinical leaders with the skills to foster a culture of learning and high reliability within clinical teams. The Relational Playbook integrates evidence-based concepts and practices from positive psychology, team science, servant leadership, the Veterans Health Administration (VHA) Whole Health model, and clinical team training [8-10]. The Relational Playbook's key innovation lies in bringing these principles together into a single, cohesive program tailored specifically for frontline health care leaders. These principles are presented in an e-book with five chapters on (1) creating a positive culture, (2) teamwork, (3) leading teams, (4) creating joy in work, and (5) communication and high reliability. The Relational Playbook contains brief asynchronous learning modules, 11 kick-off interventions, and 39 additional evidence-based interventions. Table 1 provides more details on the Relational Playbook chapters, their resource topics, and their kick-off interventions. Frontline leaders complete weekly self-directed education and then select and implement specific Relational Playbook interventions into existing meetings or trainings. Each chapter builds on the previous one and results in the development of supportive learning environments.

Table . Relational Playbook chapters, modules, and kick-off interventions.

Relational Playbook chapter	Resource topics	Kick-off interventions
Chapter 1: Creating a Positive Culture	• Positive culture • Assessing team well-being • Appreciative inquiry	• "Three Good Things" practice [11] • Appreciative inquiry questions [12]
Chapter 2: Teamwork	• Building a team • Relationships at work • Difficult relationships at work • Hiring for high-performing teams	• "Walk in My Shoes" exercise [13] • Ice breaker questions [13]
Chapter 3: Leading Teams	• Wellness-centered leadership • Servant leadership • Essential leadership skills	• "Stop, Start, Continue" method [13] • "Situation-Behavior-Impact" feedback [14]
Chapter 4: Creating Joy in Work	• Burnout • Joy and happiness • Gratitude	• Understanding what matters [15] • "Was It Worth It?" method [16] • Gratitude huddle [17]
Chapter 5: Communication and High Reliability	• Effective communication • High-reliability practices	• Start-of-day huddles [18] • Debriefs [18]

Pilot research with the Relational Playbook has suggested improvements in employee engagement and retention while reducing burnout and turnover, which are critical workforce challenges [8]. The Relational Playbook aligns with multiple priority areas in health care, including the shift toward learning health systems to improve patient safety. In 2022, the Relational Playbook was registered as an invention with the VHA Technology Transfer Program (VHA ID 2022-474) to foster partnerships with external digital technology innovators and leverage emerging technology to expand and scale the program.

The Relational Playbook team collaborated with Whistle Systems, a company specializing in digital programs and trainings to sustain employee behavior change and improve workplace culture. Leveraging a mobile-first design, the Whistle platform integrates evidence-based strategies such as microlearning, gamification, and strategic nudges to optimize user engagement and adherence [19]. The platform delivers real-time feedback through notifications, on-demand resources, and a user-centric interface to enhance accessibility. Whistle has shown measurable success in improving employee engagement and reducing turnover across sectors, including aviation, finance, and construction [20]. The partnership aimed to adapt Relational Playbook content to the Whistle platform and assess whether this innovative technology is an acceptable, appropriate, and feasible tool for delivering the Relational Playbook to nurse leaders.

Methods

Overview

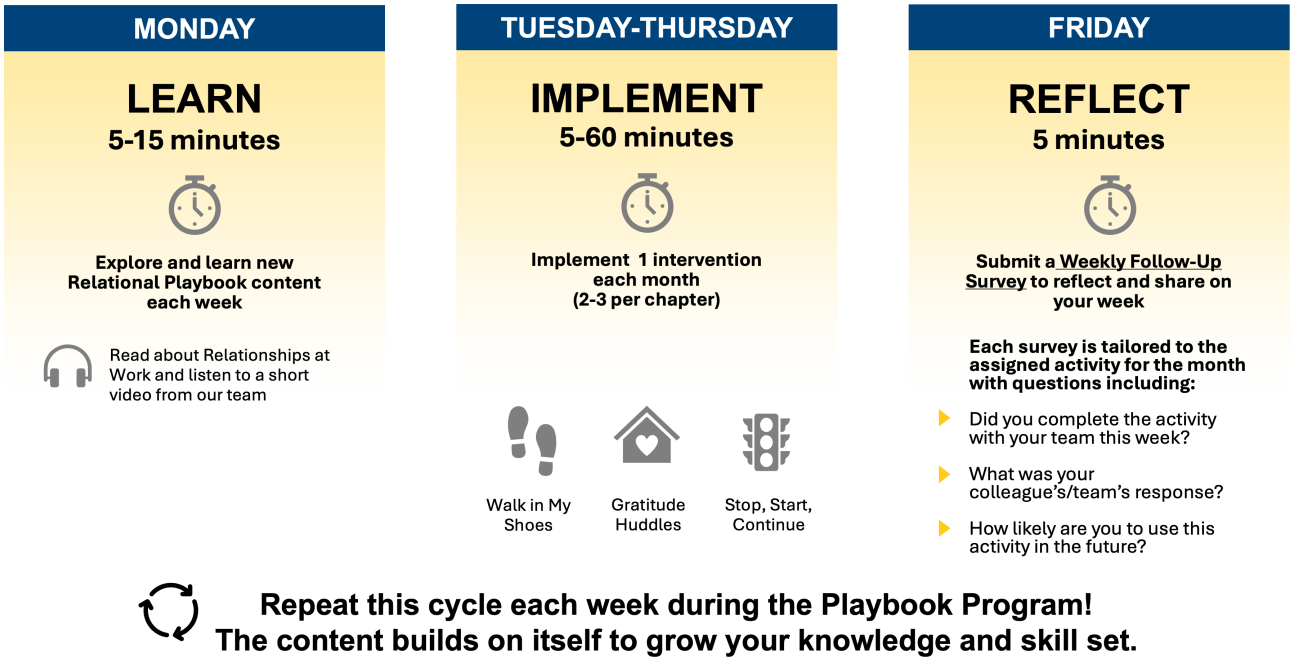
We conducted a 1-year, single-team feasibility study of the Relational Playbook delivered on the Whistle platform within a real-world clinical setting using descriptive survey data and qualitative interview analysis. A cardiology nurse practitioner (NP) team (n=5) at a large academic medical center volunteered to pilot the Relational Playbook on Whistle driven by an interest in leadership development training to improve team dynamics. The NP section chief was the primary participant, and their team members were invited to engage with the platform to enhance team understanding and participation. To maintain participant confidentiality, detailed demographic information was not reported. The inclusion criterion was a formal supervisory role in the department. Participation was voluntary and considered part of the employees' work.

XSL•FO
RenderX

The Relational Playbook and Whistle team established a cooperative research and development agreement enabling collaboration between the VHA and private companies. The first author adapted the Relational Playbook content for delivery on the Whistle platform with input from the Relational Playbook developers (HMG and BC) and Whistle engineers. The educational content was reformatted into a microlearning flash card model incorporating short text, colored images, videos, and kick-off interventions (Multimedia Appendix 1). The Whistle-enabled Relational Playbook used multiple behavioral science mechanisms, including celebratory feedback (confetti) to trigger dopaminergic reward responses, progress indicators (completion bars) to leverage the goal gradient effect, microlearning modules to reduce cognitive load and enhance perceived task simplicity, and strategic nudges to serve as action triggers and mitigate decision inertia. The Whistle web platform hosted the content, with accessible iPhone and Android mobile apps.

The Relational Playbook on Whistle begins with participants completing the 13-item Learning Environment Assessment Tool (Multimedia Appendix 2), an abbreviated version of the validated 64-item Learning Environment Survey [10,21]. The Learning Environment Assessment Tool evaluates key aspects of supportive learning environments through statements such as "The cardiology team demonstrates trust and mutual respect with each other," "The cardiology team is comfortable asking for help and feedback from others," and "The cardiology team can control their own practice and regularly participate in decisions about their work." Each item is rated on a 3-point scale (1="rarely," 2="sometimes," and 3="almost always"). Responses are automatically summed in Qualtrics (Qualtrics International Inc) and ranked from lowest to highest. Using these data, the research team assigned initial (lowest ranking) and subsequent chapters to the participants in order. The Relational Playbook consists of 5 chapters, each delivered over a 2-month interval across 1 year (Figure 1). Participants receive weekly email and app notifications ("nudges") linking to flash card–based learning modules. Each module concludes with a brief comprehensive quiz and celebratory feedback to reinforce engagement. Modules end with details about the next chapter's kick-off intervention to implement.

Figure 1. The Relational Playbook program weekly activities. This figure shows what a week in the Relational Playbook looks like for a cardiology nurse practitioner team. On Monday, they are instructed to take 5 to 15 minutes and learn new content by completing learning modules and watching short videos. The rest of the week, they implement their learnings and activities through 1 intervention each week, such as "Walk in My Shoes" or "gratitude huddles." At the end of the week, they submit a weekly follow-up survey that is tailored to the assigned activity for the month to see whether they progressed with the activity and their team's response.



The evaluation data were collected through surveys and interviews from the NP section chief (eg, primary participant) to measure implementation, engagement, and adoption outcomes. Primary implementation outcomes were the acceptability, appropriateness, and feasibility of the Relational Playbook on Whistle assessed using the Acceptability of Intervention Measure, Feasibility of Intervention Measure, and Intervention Appropriateness Measure [22] (Multimedia Appendix 3). These surveys assess the acceptability, appropriateness, and feasibility of the Relational Playbook on Whistle through statements such as "The Relational Playbook on Whistle meets my approval," "The Relational Playbook on Whistle seems suitable," and "The Relational Playbook on Whistle seems implementable." Each item is rated on a 5-point scale (1="completely disagree"; 5="completely agree") and analyzed using mean scores [22]. The secondary outcomes of engagement and adoption were assessed using all participant platform visit data and survey responses (n=5), supplemented by a follow-up interview with the NP section chief to contextualize the findings. Whistle's reporting tools were used to track platform visits, module completions, survey completions, and responses.

Quantitative data, including platform use and survey responses, were summarized both by chapter and for the Relational

XSL•FO
RenderX

Playbook overall. Chapter and Playbook completion rates were calculated as the percentage of completed modules relative to the total number of modules using Microsoft Excel. Platform visit data were reviewed to confirm that all participants accessed the Relational Playbook content on Whistle at least once. For qualitative data, we used a rapid qualitative matrix approach [21]. Two team members (BC and MD) met to discuss data and reach consensus on the concepts; given the informal nature and small sample (n=1), all data were analyzed instead of stopping at a point of saturation. The initial matrix summary was developed by 1 team member using identified concepts and illustrative quotes from the data. A second team member independently reviewed and refined the matrix to ensure accuracy, completeness, and consistency in data representation. To enhance rigor, discrepancies were discussed and resolved collaboratively. The full analytic team then conducted a review of the finalized matrix, engaging in consensus building to identify cross-cutting concepts and key insights.

Ethical Considerations

This study was deemed an exempt human research study by the Colorado Multiple Institutional Review Board (17-1153) and did not require informed consent. All participant data were handled in accordance with institutional privacy and confidentiality guidelines.

Results

The NP section chief and 4 cardiology NP team members participated in the study, with the NP section chief acting as the primary implementer and evaluator due to their formal supervisory role. The NP team provides clinical care across various inpatient and outpatient settings (eg, heart failure clinic and structural heart and valve clinic) and meets virtually each month for updates and professional development.

Acceptability, Appropriateness, and Feasibility and Platform Visits

The NP section chief, the primary participant, strongly agreed (5/5) on all measures that the Relational Playbook on the Whistle platform was acceptable (Acceptability of Intervention Measure), appropriate (Intervention Appropriateness Measure), and feasible (Feasibility of Intervention Measure). They engaged with 86% (73/85) of the learning content and reflection surveys and adopted all 11 kick-off interventions at least once with their team. The 4 cardiology NP team members engaged fully with the introduction, and their engagement declined over the subsequent 5 chapters. During scheduled check-ins, NP team members reported that their busy schedules, their direct care responsibilities, and not leading their own teams made it harder for them to implement the Relational Playbook interventions as they were not directly applicable to their roles. They found the learning components valuable and noted that these resources enhanced their understanding of and engagement in the section chief's activities. However, the lack of applicability in their daily work led to disengagement with the platform. Figure 2 shows the completion rates by chapter for the Whistle-enabled Relational Playbook.

The NP section chief provided generally positive feedback on the kick-off interventions and reported that their team adopted a more positive outlook, identifying what mattered most to them, such as "'family,' 'friends/relationships,' 'fulfilling my responsibilities,' and 'being good at my job.'" They had specific successes with using debriefs:

> *Worked well with a co-worker to solve a problem!*

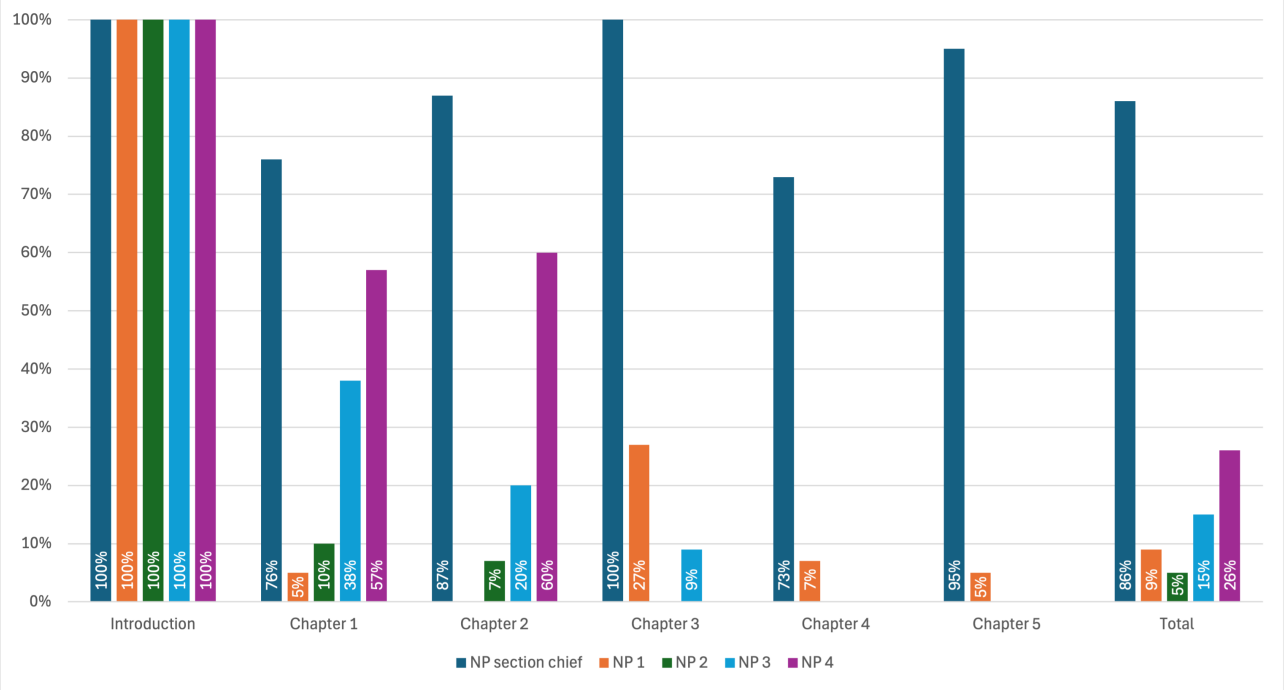
However, the *Walk in My Shoes* and *Was It Worth It* interventions were more challenging to implement, as the NP section chief explained:

> *...I think it needs to be the right timing and people.*

They also reported adapting some interventions, such as replacing the "Stop, Start, Continue" group discussion with an online survey and incorporating the "What was a win this week" question from chapter 4 as a meeting icebreaker.

XSL•FO
RenderX

Figure 2. The Relational Playbook on Whistle completion rates. This figure displays the chapter completion rates for each participant alongside their total completion rate for the entire program. While all participants completed the introduction chapter, participation rates declined for everyone but the nurse practitioner (NP) section chief for chapters 1 to 5.



Interview Results

Platform Usability and Learning Content

The NP section chief primarily used the website, rating it as easy to use (4 out of 5 on a verbal ease-of-use scale) and preferable to the mobile app, which required frequent log-ins and multiple clicks to access content. They appreciated the ability to navigate seamlessly through the program and liked the push notifications (ie, nudges). The Relational Playbook content was described as "bite size enough to be done in one sitting and...easily digested." However, they observed that the Relational Playbook was designed for teams that "work together in a clinical way." Considering that the NP group did not physically work together, they suggested modifying the content to better reflect virtual team dynamics.

Relational Playbook Implementation

The NP section chief rated the implementation of the Relational Playbook practices as moderately easy (3 out of 5 on a verbal ease-of-implementation scale). They confirmed that the practices fostered a more positive team outlook, stating the following:

> *...there's some hard days.... And we focus a lot on the negative...instead of what went right. I liked that about the Playbook.*

They emphasized that the learning content only took 5 to 10 minutes to complete. Overall, the NP section chief described the Relational Playbook on Whistle as a valuable program:

> *You can always improve as a leader—I highly recommend it.*

Adaptations and Sustainment

The NP section chief suggested adaptations before expanding the program: (1) tailor Relational Playbook content for virtual teams that do not meet routinely in person, (2) condense the program to 6 months (from 1 year), (3) reduce nudges to every other week (from weekly), and (4) develop an educational module for team members without formal leadership roles. At the conclusion of the 1-year program, the NP section chief reported ongoing use of the huddles and appreciative inquiry practices.

Discussion

This feasibility study demonstrated that the Relational Playbook on Whistle is an acceptable, appropriate, and feasible intervention for a nurse leader. The NP section chief engaged with 86% (73/85) of the educational content and implemented all 11 kick-off interventions, confirming the platform's usability. Participant feedback emphasized the Relational Playbook's strengths, including its concise, "bite-sized" content; intuitive navigation; and direct relevance to clinical practice. By applying Relational Playbook practices, the NP section chief fostered a culture of learning and positivity within the team. The NP clinical team accessed the platform and participated intermittently. This may reflect the absence of opportunity for practicing NPs to put the leadership interventions into practice, the unique challenges of virtual teams that do not routinely work in person (noted above), or the assumption that leadership development is only for those who have formal leadership titles. Their reports of finding the education components valuable reinforce the section chief's recommendation for an educational module for team members without formal leadership roles.

The Whistle platform behavioral science features guided participants to set learning intentions, assume responsibility for their goals, and receive feedback on progress. While our study data did not allow for a granular analysis of each feature's individual contribution to engagement, this is an area that we

XSL•FO
RenderX

will explore in future research. Rapid application of newly acquired skills in practice represents the gold standard of leadership development programs. Our single-team study process and findings align with recent work by Güntner et al [23], which reported that a web-based leadership transfer intervention positively influenced leaders' mindsets and self-regulated learning. The significance of these studies lies in demonstrating that digital microlearning interventions can effectively support leadership development in high-demand clinical environments, offering a scalable and cost-efficient alternative to traditional programs. Future studies of the Relational Playbook on Whistle will advance the evidence base for digital leadership training programs, ensuring positive outcomes in a cost-effective manner.

The partnership with Whistle Systems was an opportunity to integrate an evidence-based leadership development program into a digital technology innovation that delivered microlearning content using behavioral science and gamification principles. This technology overcame time constraints, a common barrier for busy nurse leaders [24,25], and promoted engagement and new habits. The Whistle platform's usability was rated favorably by the NP section chief, with the website preferred over the mobile app due to frequent re–log-ins and navigation challenges. Nudges effectively promoted engagement by linking directly to assigned content; however, a decrease in frequency was suggested to reduce response burden.

Additional behavior-driving tools offered by the Whistle platform were not leveraged in this feasibility study that could increase engagement and effectiveness among team members. These include real-time automated acknowledgments; hospital-branded Visa cards for monetary reward; peer or community recognition engagement through Whistle's "Town-Square" social feature; or the platform's artificial intelligence engine Robin, which tailors nudges informed by behavioral personalization algorithms to effectively prompt individuals while considering their unique motivational drivers [19]. Future iterations of the Relational Playbook on Whistle will include an educational module for team members in clinical roles to motivate them to learn and engage in culture change, content for virtual and hybrid teams, and additional Whistle tools.

This study's strengths include the real-world evaluation of the Relational Playbook on Whistle and its focus on clinical leadership development. The feasibility approach provided early insights into implementation outcomes, user experience, and potential program impact. However, several limitations should be noted: the small sample size and single-team design, reliance on 1 primary participant, absence of a control group, and lack of objective leadership or patient-related outcomes reduce the ability to attribute observed changes to the intervention. These factors limit generalizability, and findings should be interpreted with caution. Future research will address these limitations by including a larger, more diverse sample; detailed demographic data; a comparison group; and multisource ratings to capture changes in leader behavior and team culture before and after implementation.

The Relational Playbook on Whistle shows strong promise as an acceptable, appropriate, and feasible nurse leadership development program capable of addressing critical workforce challenges such as burnout, team dynamics, and leadership readiness. To maximize its impact, future iterations should include adaptations for virtual teams, streamline program delivery, actively engage all team members, assess the impact on patient safety, and evaluate cost-effectiveness compared to traditional leadership programs. Scaling the Relational Playbook across diverse clinical settings will require strategic collaboration with technology partners to ensure accessibility, flexibility, and sustained implementation. By leveraging digital platforms for leadership development, health care organizations can accelerate skill acquisition, strengthen team culture, and build resilient leaders in an increasingly complex care environment.

Acknowledgments

The authors thank the cardiology nurse practitioner team who participated in this study. The content is solely the responsibility of the authors and does not necessarily represent the official views of the US Department of Veterans Affairs, the US government, or the Agency for Healthcare Research and Quality.

Funding

This material is based on work supported (or supported in part) by the US Department of Veterans Affairs (VA), Veterans Health Administration. The Office of Research and Development, Health Services Research and Development Service, provides support through the Seattle-Denver Center of Innovation for Veteran-Centered and Value-Driven Care. HMG is supported by Career Development Award 1IK2HX002567-01A1 from the VA Office of Research and Development. In-kind funding for use of the Whistle web application and mobile app was provided by Whistle Systems. The VA and Whistle Systems had no role in the study design; collection, management, analysis, or interpretation of the data; or decision to publish. Whistle Systems contributed to the Introduction, Methods, and Discussion sections of the manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request after US Department of Veterans Affairs permissions are obtained for release.

XSL•FO
RenderX

Authors' Contributions

MD and HMG conceived the intervention and study design. MD, HMG, and BC designed, collected, analyzed, and interpreted the mixed methods data. All authors contributed to writing and approved the final manuscript.

Conflicts of Interest

MD, BC, and HMG receive salaries from the US Department of Veterans Affairs. The other author declares no other conflicts of interest.

Multimedia Appendix 1
Screenshots from the Whistle platform showing its unique features, with the Relational Playbook's educational content reformatted into a microlearning flash card model incorporating short text, colored images, videos, and kick-off interventions.
[DOCX File, 849 KB - nursing_v9i1e79188_app1.docx]

Multimedia Appendix 2
The 13-item Learning Environment Assessment Tool (LEAT) is an abbreviated version of the validated 64-item Learning Environment Survey. The LEAT assesses key aspects of supportive learning environments, and each item is rated on a 3-point scale (1="rarely," 2="sometimes," and 3="almost always").
[DOCX File, 27 KB - nursing_v9i1e79188_app2.docx]

Multimedia Appendix 3
Primary implementation outcomes assessed using the Acceptability of Intervention Measure, Feasibility of Intervention Measure, and Intervention Appropriateness Measure tailored to the Relational Playbook on Whistle. Each item is rated on a 5-point scale (1="completely disagree"; 5="completely agree") and analyzed using mean scores.
[DOCX File, 49 KB - nursing_v9i1e79188_app3.docx]

References

1. Blake N. Future of nursing 2020-2030: supporting the health and well-being of nurses. AACN Adv Crit Care 2022 Mar 15;33(1):99-102. [doi: 10.4037/aacnacc2022286] [Medline: 35259218]
2. Al Sabei SD, Ross AM. The relationship between nursing leadership and patient readmission rate: a systematic review. Can J Nurs Res 2023 Sep;55(3):267-278. [doi: 10.1177/08445621231152959] [Medline: 36734052]
3. Fennimore L, Wolf G. Nurse manager leadership development: leveraging the evidence and system-level support. J Nurs Adm 2011 May;41(5):204-210. [doi: 10.1097/NNA.0b013e3182171aff] [Medline: 21519206]
4. Leroy H, Anisman-Razin M, Detert J. Leadership development is failing us here's how to fix it. Massachusetts Institute of Technology. 2025. URL: https://sloanreview.mit.edu/article/leadership-development-is-failing-us-heres-how-to-fix-it [accessed 2025-05-25]
5. Cummings GG, Lee S, Tate K, et al. The essentials of nursing leadership: a systematic review of factors and educational interventions influencing nursing leadership. Int J Nurs Stud 2021 Mar;115:103842. [doi: 10.1016/j.ijnurstu.2020.103842] [Medline: 33383271]
6. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development: Prentice-Hal; 2014.
7. Lave J, Wenger E. Situated Learning: Legitimate Peripheral Participation: Cambridge University Press; 1991.
8. Gilmartin HM, Connelly B, Daus M, et al. Implementation of the acute inpatient medicine-high reliability, learning environment, and workforce development initiative (AIM-HI) in rural veterans health administration hospitals: a mixed methods evaluation protocol. J Hosp Med 2025 Jan;20(1):33-41. [doi: 10.1002/jhm.13474] [Medline: 39149834]
9. Gilmartin HM, Connelly B, Hess E, et al. Developing a relational playbook for cardiology teams to cultivate supportive learning environments, enhance clinician well-being, and veteran care. Learn Health Syst 2024 Apr;8(2):e10383. [doi: 10.1002/lrh2.10383] [Medline: 38633018]
10. Gilmartin HM, Hess E, Mueller C, et al. Learning environments, reliability enhancing work practices, employee engagement, and safety climate in VA cardiac catheterization laboratories. Health Serv Res 2022 Apr;57(2):385-391. [doi: 10.1111/1475-6773.13907] [Medline: 35297037]
11. Sexton JB, Adair KC. Forty-five good things: a prospective pilot study of the Three Good Things well-being intervention in the USA for healthcare worker emotional exhaustion, depression, work-life balance and happiness. BMJ Open 2019 Mar 20;9(3):e022695. [doi: 10.1136/bmjopen-2018-022695] [Medline: 30898795]
12. Richer MC, Ritchie J, Marchionni C. "If we can't do more, let's do it differently!": using appreciative inquiry to promote innovative ideas for better health care work environments. J Nurs Manag 2009 Dec;17(8):947-955. [doi: 10.1111/j.1365-2834.2009.01022.x] [Medline: 19941568]
13. Maslach C, Leiter MP. The Burnout Challenge: Managing People's Relationships with Their Jobs: Harvard University Press; 2022.

XSL•FO
RenderX

14. Gentry B, Young S. Use situation-behavior-impact (SBI) to understand intent.: Center for Creative Leadership; 2025. URL: https://www.ccl.org/articles/leading-effectively-articles/closing-the-gap-between-intent-vs-impact-sbii [accessed 2025-05-25]
15. Perlo J, Feeley D. Applying community organizing principles to restore joy in work. NEJM Catal 2018;4(3). [doi: 10.1056/CAT.18.0165]
16. Thanarajasingam G, Basch E, Mead-Harvey C, et al. An exploratory analysis of the "Was It Worth It?" Questionnaire as a novel metric to capture patient perceptions of cancer treatment. Value Health 2022 Jul;25(7):1081-1086. [doi: 10.1016/j.jval.2021.11.1368] [Medline: 35779938]
17. Greater good magazine. Greater Good Science Center. 2021. URL: https://ggsc.berkeley.edu/what_we_do/greater_good_magazine [accessed 2025-05-25]
18. Scoville R, Little K, Rakover J, Luther K, Mate K. Sustaining improvement. East London Foundation Trust. 2016. URL: https://qi.elft.nhs.uk/resource/sustaining-improvement [accessed 2025-05-25]
19. Whistle rewards. Whistle Systems Inc. 2025. URL: https://www.wewhistle.com/ [accessed 2025-05-25]
20. Carter D, Dornfeld C, Newman L. The right tool for the job: improving construction projects using rewards and incentives. Whistle Systems. 2023. URL: https://wewhistle.com/wp-content/uploads/2024/10/The-Right-tool-for-the-job-Oct-2023-3.pdf [accessed 2025-05-25]
21. Gilmartin HM, Hess E, Mueller C, Plomondon ME, Waldo SW, Battaglia C. A pilot study to assess the learning environment and use of reliability enhancing work practices in VHA cardiac catheterization laboratories. Learn Health Syst 2021 Apr;5(2):e10227. [doi: 10.1002/lrh2.10227] [Medline: 33889736]
22. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
23. Güntner AV, Heimann AL, Kleinmann M, Ingold PV. The combined effect of a goal-oriented leadership app and leaders' mindset in optimising training transfer. European Journal of Work and Organizational Psychology 2025 Mar 4;34(2):175-187. [doi: 10.1080/1359432X.2024.2412359]
24. Salemo C, Fitzpatrick MM. Nurse leader development programs: a best-practice framework. Nurs Manage 2023 Sep 1;54(9):35-46. [doi: 10.1097/nmg.0000000000000050] [Medline: 37647558]
25. Chen W, Modanloo S, Graham ID, Hu J, Lewis KB, Gifford W. A mixed-methods systematic review of interventions to improve leadership competencies of managers supervising nurses. J Nurs Manag 2022 Nov;30(8):4156-4211. [doi: 10.1111/jonm.13828] [Medline: 36194186]

Abbreviations

NP: nurse practitioner
VHA: Veterans Health Administration

Edited by E Borycki; submitted 16.Jun.2025; peer-reviewed by FA Dhabbari, J Himstedt, K Morris, WO Pereira; revised version received 01.Dec.2025; accepted 23.Dec.2025; published 02.Feb.2026.

Please cite as:
Daus M, Connelly B, Carter D, Gilmartin HM
The Relational Playbook Nurse Leadership Development Program Using the Whistle Systems Employee Recognition Platform: Feasibility Mixed Methods Study
JMIR Nursing 2026;9:e79188
URL: https://nursing.jmir.org/2026/1/e79188
doi:10.2196/79188

© Marguerite Daus, Brigid Connelly, Drew Carter, Heather M Gilmartin. Originally published in JMIR Nursing (https://nursing.jmir.org), 2.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

WeChat-Based Nursing Interventions in Women's Mobile Health: Systematic Review

Jun Xiao[1], MD, PhD; Rui Li[2], BSN; Yangyang Wu[3], BSA; Tong Wu[4,5], MD, PhD

[1]Department of Thyroid and Breast Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[2]Emergency Department, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[3]Department of Animal Science and Technology, Sichuan Agricultural University, Chengdu, China

[4]Key Laboratory of Cancer Invasion and Metastasis (Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[5]Department of Obstetrics and Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, No. 1095, Jiefang Avenue, Wuhan, China

Corresponding Author:
Tong Wu, MD, PhD
Key Laboratory of Cancer Invasion and Metastasis (Ministry of Education), Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Abstract

Background: Mobile health (mHealth) technology offers new approaches to improve women's health by providing personalized monitoring and real-time guidance. As one of the most widely used social media platforms in China, WeChat has shown great potential in mHealth practice, yet systematic evidence on its application in women's health care remains insufficient.

Objective: This study aims to systematically review WeChat-based nursing interventions in women's mHealth in order to clarify the application status, intervention modalities, target populations, and effectiveness outcomes.

Methods: Searches were conducted in IEEE Xplore, Web of Science, PubMed, Scopus, ACM Digital Library, and Cochrane Central Register of Controlled Trials between January 2011 and December 2024. Two independent reviewers screened studies; extracted data on study design, intervention forms, target diseases, and outcome indicators; and assessed methodological quality. The Cohen κ coefficient was used to evaluate interreviewer agreement. Publication trends, institutional collaborations, author contributions, and research hotspots were analyzed using VOSviewer (Leiden University) and InCites (Clarivate) for bibliometric analysis.

Results: A total of 31 eligible studies published from 2014 to 2024 were included. Most studies were randomized controlled trials (n=27). Intervention modalities mainly included WeChat groups (n=22), official accounts (n=18), applets (n=4), and private chats (n=9), mostly used in combination. The top focused health issues were prenatal care (n=5), breast cancer (n=5), gynecological cancer (n=5), and gestational diabetes mellitus (n=4). Six studies adopted multidisciplinary teams. Cohen κ was 0.71, indicating substantial agreement. Publications grew rapidly after 2018, peaking in 2021 and 2024. A total of 40 institutions participated, with Xi'an Jiaotong University having the highest citation impact. Most studies were at high risk of bias due to a nonblinding design.

Conclusions: WeChat-based nursing interventions improve personalized health information access, self-management ability, treatment compliance, and real-time doctor-patient communication for women. This is the first systematic review to evaluate WeChat mHealth interventions in women's health, filling the research gap. Future research should focus on improving methodological quality, exploring cross-cultural adaptability, conducting long-term follow-up, and integrating wearable devices and electronic health records to further optimize WeChat-based women's health services.

Trial Registration: PROSPERO CRD420251051789; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251051789

(JMIR Nursing 2026;9:e82679) doi:10.2196/82679

KEYWORDS

WeChat; mobile health; mHealth; bibliometric; applet; women

Introduction

Women face unique health issues at different life stages [1]. In adolescence, the onset of menstruation may bring challenges like dysmenorrhea or polycystic ovary syndrome [2], while reproductive years often involve fertility management, prenatal care, and pregnancy-related conditions [3]. During midlife, menopause introduces hormonal fluctuations that can trigger

XSL•FO
RenderX

hot flashes, osteoporosis, and mood disorders [4]. Even beyond menopause, postreproductive years confer increased risks of cardiovascular disease, breast cancer, and cognitive decline [5,6]. Furthermore, women tend to be more vulnerable to anxiety and depression than men, owing to a mix of biological, psychological, and social elements [7]. Therefore, safeguarding women's health is an important task in the public health field. It not only concerns their own quality of life but also correlates with the overall development of society.

With the quick development of information technology, mobile health (mHealth) offers a new avenue for improving women's health. It refers to leveraging mobile devices, like smartphones and wearables, to provide accessible and tailored health care services [8]. For instance, women may use the Oura ring to track their sleep, temperature, and heart rate during the menstrual cycle, which helps find early health warning signs [9]. A high usage of the SmartMoms app in pregnant women exerts a positive effect on moderate to vigorous physical activity. Additionally, mHealth tools can provide nutrition guidance and exercise advice for pregnant and postmenopausal women [10]. Some may even promote screening attendance and promote timely preventive health care behaviors [11]. Overall, mHealth is transforming women's health care by supporting their varied needs and making health interventions more accessible and effective.

In China, WeChat stands out as a major social media platform with a substantial user population. According to Tencent's 2025 annual earnings report [12], WeChat had more than 1.3 billion monthly active users as of March 2025, covering most of China's internet users. This phenomenon renders the WeChat platform as an effective mHealth platform for women's health [13,14], with several key advantages. First, its large user base allows accurate delivery of health information and services. Second, users can share health care information online [15,16]. Third, applets can be developed into practical health management tools [17]. However, despite its convenience, it still lacks systematic reviews on the WeChat app in women's health care.

Previous reviews have evaluated the effects of mHealth interventions on women's health outcomes, but they did not distinguish between different mHealth delivery modalities [18,19]. In particular, WeChat-based interventions were not separately analyzed in prior syntheses. Meanwhile, existing reviews that involve WeChat have centered on chronic conditions such as cancer [20], mental health [21,22], and diabetes [23], without focusing on the full life cycle of women's health. Under this context, it is essential to systematically explore the application, effectiveness, and implementation of WeChat-based mHealth interventions in women's health to fill this important evidence gap.

This systematic review aims to comprehensively and systematically collect and analyze relevant research on women's mHealth apps based on the WeChat platform, summarizing their application status, effectiveness, and existing problems. By comprehensively evaluating WeChat's apps in all aspects of women's health monitoring, disease prevention, and health management, it provides a scientific basis for further optimizing the application of WeChat in women's mHealth. At the same time, this study helps to identify research gaps and development trends in this field, providing directions for subsequent research and practice. Thus, it can better use the WeChat platform to improve women's health levels, filling the gaps in existing research, and has important theoretical and practical significance.

Methods

Study Design

This review was registered in the PROSPERO International Prospective Register of Systematic Reviews (CRD420251051789) and followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines closely in its conduct and reporting (Checklist 1) [24].

Research Question

This systematic review aimed to explore the research question: How effective are WeChat-based nursing interventions in enhancing health outcomes for women at various life stages, and what is their current implementation status? The design of the review adhered to the PICO framework:

1. Population (P): Women of all ages and life stages (adolescence, reproductive years, pregnancy, postpartum, menopause, and postmenopausal periods) with various women-related health conditions.
2. Intervention (I): WeChat-based nursing interventions (WeChat groups, official accounts, applets, and private chat) for health education, care management, psychological support, and multidisciplinary collaborative care.
3. Comparison (C): Standard routine care, usual hospital follow-up, or no additional WeChat-based intervention.
4. Outcomes (O): Clinical indicators, quality of life, self-management ability, treatment compliance, psychological status, satisfaction, and behavioral outcomes.

Database Selection and Search Strategy

An extensive search plan was formulated by using free-text words and subject headings. The following databases were explored from January 2011 to December 2024: IEEE Xplore, Web of Science, PubMed, Scopus, ACM Digital Library, the Cochrane Central Register of Controlled Trials, and the China National Knowledge Infrastructure (CNKI). An effective retrieval strategy was established using a 2-step process. The first step involved collecting and organizing relevant search terms from related systematic reviews and meta-analyses [25,26]. Afterward, 3 authors (RL, YW, and TW) carefully examined and improved these terms to guarantee thorough inclusion of all essential terms. Using Boolean logic operators, the concluding search strategy was designed: ("wechat" or "weixin" or "official account" or "public account" or "applet" or "mini program") and (medic* OR illness* OR disease* OR health* OR pharma* OR drug* OR therap*) and (woman OR women OR girl OR female). The full search strategy for each database is detailed in Table 1. Additionally, we did not impose any restrictions on health-related keywords in the Cochrane Central Register of Controlled Trials database, because we

XSL•FO
RenderX

assumed that all documents in this database are related to human health by default, so that we can cover a wider range of potential literatures. Since numerous relevant studies have been published in Chinese, we specifically included CNKI. Other ways to identify studies involved expert recommendations and manually searching the references of included studies and pertinent systematic reviews.

Table . Specific search terms of databases.

Database	Search term	Studies, n
IEEE Xplore	((((All Metadata:“Wechat") OR (All Metadata:"weixin") OR (All Metadata:"official account") OR (All Metadata:"applet") OR (All Metadata:"Mini Program"))) AND ((All Metadata:medic*) OR (All Metadata:illness*) OR (All Metadata:disease*) OR (All Metadata:health*) OR (All Metadata:pharma*) OR (All Metadata:drug*) OR (All Metadata:therap*))) AND ((All Metadata:"woman") OR (All Metadata:"women") OR (All Metadata:"girl") OR (All Metadata:"female”))	6
Web of Science	((TS=(“Wechat” or “weixin” or “official account” or “applet” or “Mini Program”)) AND TS=(medic* OR illness* OR disease* OR health* OR pharma* OR drug* OR therap*)) AND TS=(“woman” or “women” or “girl” or “female”)	629
PubMed	((“Wechat” OR “weixin” OR “official account” OR “applet” OR “Mini Program”) AND (medic* OR illness* OR disease* OR health* OR pharma* OR drug* OR therap*)) AND (“woman” or “women” or “girl” or “female”)	575
ACM Digital Library	[[All: “wechat”] OR [All: “weixin”] OR [All: “official account”] OR [All: “applet”] OR [All: “mini program”]] AND [[All: medic*] OR [All: illness*] OR [All: disease*] OR [All: health*] OR [All: pharma*] OR [All: drug*] OR [All: therap*]] AND [[All: “woman”] OR [All: “women”] OR [All: “girl”] OR [All: “female”]]	583
Cochrane Central Register of Controlled Trials	“Wechat” or “weimin” or “official account” or “applet” or “Mini Program” in Title Abstract Keyword AND “woman” or “women” or “girl” or “female” in Title Abstract Keyword - (Word variations have been searched)	351
Scopus	TITLE-ABS-KEY (“Wechat” OR “weixin” OR “official account” OR “applet” OR “Mini Program”) AND TITLE-ABS-KEY (medic* OR illness* OR disease* OR health* OR pharma* OR drug* OR therap*) AND TITLE-ABS-KEY (“woman” OR “women” OR “girl” OR “female”)	900
China National Knowledge Infrastructure	(SU%='微信' OR SU%='公众号' OR SU%='小程序' OR SU%='服务号') AND (SU%='女性' OR SU%='妇女' OR SU%='女生' OR SU%='女童') AND (SU%='健康' OR SU%='医疗' OR SU%='疾病' OR SU%='卫生' OR SU%='护理' OR SU%='康复' OR SU%='健康教育' OR SU%='健康管理')	347

Retrieved records were first imported into EndNote X21. Duplicate records were identified using the automatic deduplication tool, and a manual check was conducted to remove any duplicates left due to differences in author names, publication formats, or metadata entries.

Screening Process

The titles and abstracts were independently reviewed by two reviewers (JX and RL) following a step-by-step process using the established inclusion and exclusion criteria (Table 2). Any discrepancies were resolved through discussion until consensus; if unresolved, a third senior reviewer (TW) adjudicated the final

XSL•FO
RenderX

decision. Cohen κ was used to assess the agreement between the reviewers.

Table . Inclusion and exclusion criteria for title and abstract screening.

Framework	Inclusion criteria	Exclusion criteria
Population	• Women of all age groups and life stages with related health conditions	• Non-female-specific diseases (asthma, appendicitis, etc)
Intervention	• WeChat-based mHealth interventions in women (health education, symptom monitoring, care management, psychological intervention, etc) • Focused on nursing-led or nursing-coordinated interventions, determined by the actual intervention content, the first/corresponding author's nursing department affiliation, and author contribution statements	• Investigated the relationship between WeChat use and other factors (mental health, etc) • Computational analytical studies (content analysis, framework development, etc) • Merely used WeChat for questionnaire distribution and data collection, with no active health intervention • Irrelevant to WeChat-based mHealth interventions
Comparison	• Standard care and follow-up	• Not available
Outcomes	• Clinical indicators, psychological status, self-management, compliance, and other health-related outcomes	• Not available
Study design	• Original research articles • Interventional studies (randomized controlled trials and quasi-experiments) • No language restrictions	• Clinical trial protocols • Noninterventional designs (cross-sectional, cohort, case series, etc) • CNKI[a]-sourced literature outside Chinese core journal databases (Peking University Core, CSCD[b], Chinese Science and Technology Core Journals)

[a]CNKI: China National Knowledge Infrastructure.

[b]CSCD: Chinese Science Citation Database.

Quality Appraisal

The Cochrane risk-of-bias tool [27] and Risk of Bias in Nonrandomized Studies of Interventions (ROBINS-I) [28] were used to assess the quality of randomized controlled trials (RCTs) and quasi-experiments, respectively. Both instruments encourage assessment of internal and external validity and aid in evaluating methodological quality and clarity in reporting. Two independent reviewers (JX and RL) conducted the quality assessment for all studies. Any disagreements were resolved through discussion or by involving a third reviewer (TW).

Data Preprocessing and Extraction

To support subsequent bibliometric analysis, records from all databases were standardized and manually adjusted to align with the Web of Science format. For studies with missing bibliometric information (eg, incomplete affiliation, citation counts, or journal metadata), supplementary information was manually retrieved from the original full text.

Before bibliometric analysis, data were systematically cleaned and standardized using the *bibliometrix* R packages. To ensure consistency, author names were standardized in terms of abbreviations, initials, and sequence. For example, different versions like "Zhang, San," "Zhang, S.," and "San Zhang" were unified into a consistent format. Synonyms, aliases, and different singular/plural versions of keywords were also unified. For example, "mini-program," "mini program," and "applet" were unified into a single term. Other core bibliographic information, such as first/corresponding authors, countries, affiliations, and journals, were also standardized. Ultimately, the cleaned bibliographic records were converted into a structured data frame, making them ready for further bibliometric analysis.

Given the diverse women's health conditions, the small number of studies per condition, and clinical heterogeneity across included studies, a narrative descriptive synthesis was prespecified as the primary outcome analysis method. One reviewer (JX) independently extracted all data from each included study using the predefined form. Extracted items included author, publication year, country, study setting, study design, objectives, sample size, participant characteristics, intervention details, outcome measures, and key findings. A second reviewer (RL) independently verified all extracted data for accuracy and completeness. Discrepancies between the two reviewers (JX and RL) were resolved through discussion and consensus. If no agreement was reached, a third reviewer (TW) made the final decision. For missing study-level data, we only extracted and analyzed available valid data. No imputation or statistical filling was performed for missing data. The review adhered strictly to the registered protocol, with no deviations from the study plan.

XSL•FO
RenderX

Bibliometric Metrics

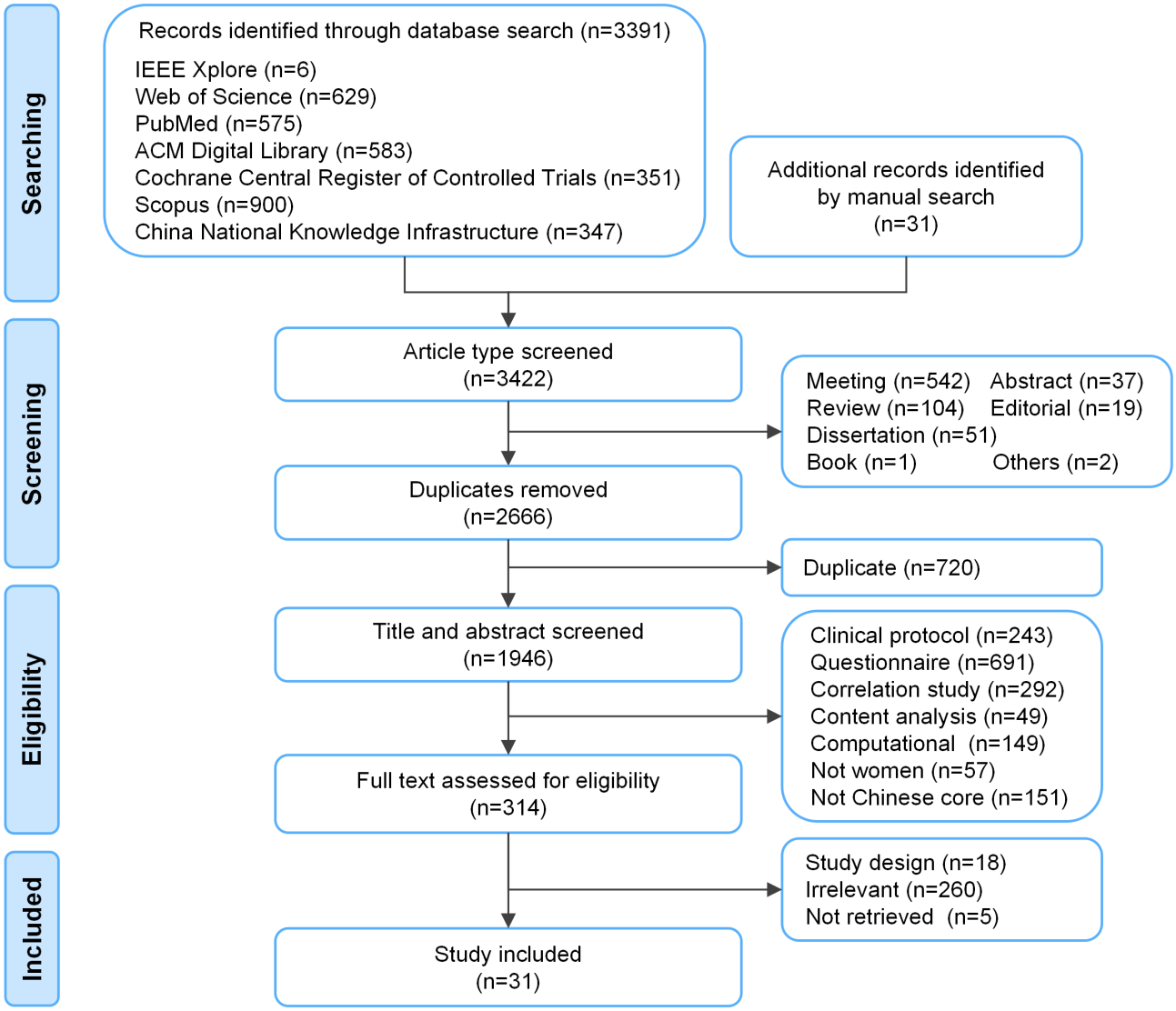
VOSviewer (Leiden University) and InCites (Clarivate) were used to quantify the knowledge structure and emerging research trends. VOSviewer 1.6.18 was used to cluster institutions and researchers. To standardize terminological variations, a custom thesaurus file was used. Association strength normalization was applied, with attraction and repulsion values set at 3 and −1. Clustering was conducted using a resolution parameter of 1.0 and a minimum cluster size of 1, resulting in a network with 21 items, 103 links, and a total link strength of 106. The total citation impact was divided by the number of years since publication to determine the yearly citation impact. The H-index signifies that an author has H articles cited no less than H times.

Results

Process of Screening

Figure 1 presents a PRISMA flow diagram summarizing the screening process and exclusions. The electronic databases yielded 3391 records, and an extra 31 were found by manually searching journals, bringing the total initial number of records to 3422. After excluding nonoriginal articles and duplicates, the remaining 1946 records proceeded to title and abstract assessment. Of these, 1632 were excluded at this stage, and 314 full-text articles were assessed for eligibility. Finally, a set of 31 studies was included in the data extraction. Cohen κ was calculated to be 0.71, which suggests a substantial level of agreement.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.



Basic Characteristics of Eligible Studies

The review systematically included 31 studies that appeared in publications between 2014 and 2024 (Table 3). Overall, 17 provinces participated in this research discourse, with the majority of publications originating from Beijing (n=5) [29-33], Shanghai (n=5) [34-38], Shanxi (n=3) [39-41], Liaoning (n=3) [42-44], and Zhejiang (n=3) [45-47]. Regarding the research type, 27 were RCTs and 4 were quasi-experiments [34,48-50], respectively.

XSL•FO
RenderX

Table . Study design, intervention modalities, and main outcomes of included studies.

Reference	Study design	Modality	WeChat intervention	Grouping and duration	Primary outcome
Prenatal care					
Guan et al [42] (2014)	RCT[a]		Establish a WeChat group and send prenatal examination times and prenatal education.	① Intervention (n=109) ② Usual care (n=109)	Higher rate of vaginal delivery, fewer negative emotions, and postpartum blood loss in ① than ②.
Huang et al [34] (2019)	QE[b]		Establish an official account to push intervention information.	① Intervention (n=69) ② Usual care (n=69) *t*=13 weeks	Higher health belief in ① than ②. No significant behavior compliance and knowledge level.
Chen et al [48] (2020)	QE		Share gestational knowledge, send course schedules, and push individualized prenatal care information.	① Intervention group (n=83) ② Regular examination (n=85) *t*=Until delivery	Higher number of prenatal examinations in ① than ②.
Gu et al [37] (2024)	RCT		Make 8 appointments for midwife-led antenatal home visit services in the outpatient department and handling routine matters.	① Intervention (n=485) ② Usual care (n=499) *t*=Until delivery	Lower cesarean section rate, chances of epidural analgesia, and NICU[c] admission, less amount of bleeding in the third stage of labor and 2 hours postpartum, and less amount of total bleeding in ① than ②.
Wang et al [51] (2024)	RCT		Based on the various stages of pregnancy, the intervention plan was formulated, incorporating motivation, ability, and trigger steps in each segment.	① Online and offline management (n=29) ② Standard care (n=48)	Reduced gestational weight gain, increased the appropriate weight gain rate, and improved perinatal outcomes in ① than ②.
Breast cancer					
Zhou et al [39] (2020)	RCT		A multimodal nursing program offers guidance, education, and support on physical, mental, and social dimensions.	① Intervention (n=56) ② Routine care (n=55) *t*=6 months	Higher health-related quality of life during early rehabilitation in ① than ②. Comparable effects on pain, fatigue, and sleep disturbances in ① and ②.
Xu et al [52] (2021)	RCT		Establish a WeChat medical service team to create case management files and conduct health education.	① Intervention (n=63) ② Routine care (n=63)	Improved self-efficacy, quality of life, and satisfaction in ① than ②.
Xu et al [32] (2022)	Self-before-and-after		Join a WeChat group and explore the official account concerning complex decongestive therapy. The intensive phase spanned 1 month, and the maintenance phase was perpetual with weekly online instructions.	Received complex decongestive therapy with WeChat (n=61) *t*=3 months	Reduced excess arm volume, relieved lymphedema symptoms, and improved the quality of life after treatment.

XSL•FO
RenderX

Reference	Study design	Modality	WeChat intervention	Grouping and duration	Primary outcome
Wu et al [47] (2021)	RCT		Group management and public account for education of patient's family.	① Intervention (n=43) ② Usual care (n=43) *t*=3 months	Improved function of the affected limb, psychosocial adjustment ability, functional exercise compliance, quality of life, and nursing satisfaction in ① than ②.
Wang et al [38] (2024)	RCT		A nursing program with multiple modes delivers details on physical, psychological, and social rehabilitation.	① Intervention (n=62) ② Routine care (n=63) *t*=Until 6 months post-surgery	Reduced fear of cancer recurrence and increased health-related quality of life in ① than ②.
Gynecological cancer					
Zou et al [53] (2018)	RCT		WeChat preoperative education includes psychological nursing, disease, HIFU,[d] and the possible complications.	① Intervention (n=251) ② Usual care (n=175)	Reduced anxiety before and after surgery, decreased pain after surgery, and greater satisfaction with the treatment in ① than ②.
Lu et al [46] (2021)	RCT		Establish an educational management group and provide targeted psychological intervention.	① Baofukang (n=50) ② Baofukang + Lactobacillus (n=50) ③ Baofukang + Lactobacillus + WeChat (n=50) *t*=3 months	Higher rate of lesion reversal, total effective rate, cognitive score in ③ than ① and ②.
Han et al [40] (2021)	RCT		Multidisciplinary nursing collaboration is maintained continuously through WeChat, telephone, and face-to-face discussions.	① Intervention (n=66) ② Regular care (n=66) *t*=3 months	Attenuated anxiety and depression, higher physical functioning, general health, and mental health in ① than ②.
Li et al [41] (2023)	RCT		Deliver psychoeducation, skills training, and counseling through the official account "Hand in Hand."	① WeChat couple-based psychosocial support (n=49) ② Eight WeChat articles (n=49) *t*=2 months	Higher relationship satisfaction and quality of life in ① than ②. Comparable effect on sexual function in ① and ②.
Tian et al [33] 2024	RCT		Join the nutrition management group on the "Good Nutrition" applet and send a private message for consultation.	① Intervention (n=48) ② Usual care (n=48)	Improved nutritional status in ① than ②. Significant altered nutrition-inflammation composite indices in ① and ②.
Gestational diabetes mellitus					
Tian et al [30] (2021)	RCT		Patients share self-management experiences based on provided learning materials.	① Intervention (n=147) ② Standard care (n=162) *t*=Until delivery	More effective for blood glucose control in ① than ②. Pregnancy outcomes showed no significant differences.

XSL•FO
RenderX

Reference	Study design	Modality	WeChat intervention	Grouping and duration	Primary outcome
Guo et al [54] (2021)	RCT		Establish a WeChat group, and push GDM[e]-related knowledge, conduct one-to-one guidance weekly and followed up online.	① Intervention (n=70) ② Routine care (n=70) *t*=Until delivery	Blood glucose levels, mental state, self-management behavior, quality of life, and pregnancy outcomes were better in ① compared to ②.
Deng et al [31] (2022)	RCT		Remind participants to record diet and exercise diaries and post photos.	① Intervention (n=47) ② Routine care (n=47) *t*=Until delivery	Fewer GDM in ① than ②. Weight gain did not differ significantly.
Wang et al [49] (2024)	QE		Pregnant women were guided offline on diet, exercise, etc, and used applet to upload data for online guidance.	① Intervention (n=26) ② Routine care (n=26) *t*=Until delivery	Reduced 2-hour postprandial glucose, improved self-management ability and blood glucose management satisfaction in ① than ②.
Breastfeeding					
Wu et al [29] (2020)	RCT		Push breastfeeding knowledge and promotion information through “Ke Xue Wei Yang.”	① Intervention (n=170) ② Routine care (n=174) *t*=Until 6 months postpartum	During the first month after birth, the rate of exclusive breastfeeding was greater in ① compared to ②.
Yu et al [55] (2021)	RCT		Deliver breastfeeding education and parenting insights, foster the mother-child relationship, and provide guidance on life specifics.	① Intervention (n=30) ② Usual care (n=30) *t*=3 months	The scores for breastfeeding efficacy and satisfaction, along with the exclusive breastfeeding rate, are superior in ① than in ②.
Polycystic ovary syndrome					
Sang et al [56] (2022)	RCT		Patients report personal information and communicate with medical staff through WeChat.	① Intervention (n=39) ② Standard care (n=40) *t*=3 months	Improved weight loss and oocyte quality, and increased the live birth rate in ① than ②.
Dilimulati et al [36] (2024)	RCT		An applet with 3 modules: watching 2 videos about healthy lifestyle per week, recording health information, and self-selected reading. A doctor coached the patients weekly.	① Intervention (n=40) ② Metformin (n=40) *t*=3 months	Comparable insulin resistance in ① and ②. ① better than ② in reducing waist circumference, waist-to-hip ratio, total fat mass, and dehydroepiandrosterone sulfate. ① fewer than ② in adverse events.
Pregnant with psychological distress					
Yang et al [45] (2019)	RCT		Join 4 mindfulness courses and a WeChat group to share experiences with researchers.	① Intervention (n=62) ② Routine care (n=61) *t*=2 months	Reduced depressive and anxious symptoms, as well as improved mindfulness skills in ① than ②.
Zhang et al [57] (2023)	RCT		Reminder to participate in thematic lessons in video format and homework, such as mindful breathing and body scan practices.	① Intervention (n=80) ② Usual care (n=80) *t*=6 weeks	Reduced maternal depression and anxiety in ① than ②. More favorable temperament scores of infants in ① than ②.

XSL•FO
RenderX

Reference	Study design	Modality	WeChat intervention	Grouping and duration	Primary outcome
Postpartum care					
Li et al [58] (2020)	RCT		Push health education information via "maternal health care standard during pregnancy."	① WeChat (n=350) ② Specialist team (n=350) ③ WeChat-specialist team (n=350) ④ Routine care (n=350) *t*=Until 7 weeks postpartum	Enhanced quality of maternal health care services in ①②③ than ④. ④ had the largest effect on both primary and secondary outcomes.
Chu et al [35] (2024)	RCT		Deliver daily training reminders and see videos guiding pelvic floor muscle training.	① Intervention (n=76) ② Control (n=72) *t*=6 weeks	Increased adherence rate, improved peak surface electromyography, and contraction endurance in ① than ②. Comparable stress urinary incontinence symptoms in ① and ②.
Others					
Wang et al [43] (2021)	RCT (gynecological inflammatory disease)			① Intervention (n=160) ② Usual care (n=160) *t*=3 months	The quality of nursing, understanding of gynecological health knowledge, and satisfaction level are higher in ① compared to ②.
Su et al [44] (2016)	RCT (primary dysmenorrhea)			① Intervention (n=90) ② Routine care (n=90) *t*=3 months	Improved knowledge, attitude, behavior, and reduced pain intensity in ① than ②.
Li et al [59] (2017)	RCT (infertility)			① Intervention (n=300) ② Usual care (n=300) *t*=3 months	Improved understanding of assisted reproduction, greater effectiveness in managing one's own cycle, self-efficacy, and satisfaction, lower degree of anxiety in ① than ②.
Yu et al [50] (2023)	QE (pelvic organ prolapse)			① Intervention (n=40) ② Routine care (n=40)	Improved cognition, compliance, pelvic floor function, muscle strength, and quality of life in ① than ②.

[a]RCT: randomized controlled trial.

[b]QE: quasi-experiment.

[c]NICU: neonatal intensive care unit.

[d]HIFU: high-intensity focused ultrasound.

[e]GDM: gestational diabetes mellitus.

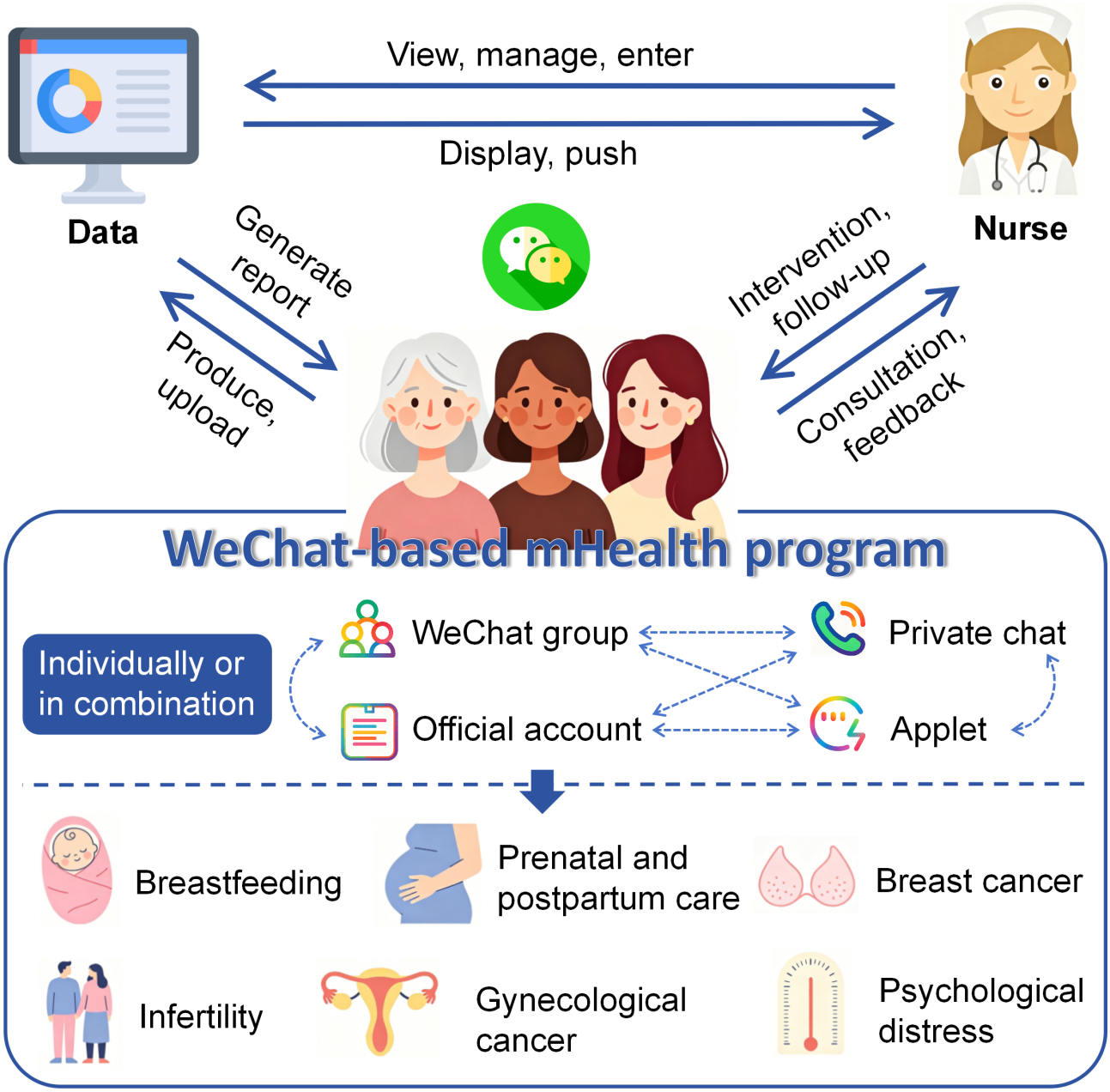
Intervention Method

In most studies, the control group received standard care, which included routine patient education and health management. For pregnant women, standard care typically involves regular classes at the maternity school and scheduled telephone follow-ups (Figure 2) [49]. The mHealth group comprises standard care plus WeChat interventions, which fall into 4 categories: WeChat group, official account, private chats, and applet. Twenty-two studies established WeChat groups, and participants could freely share experiences, discuss difficulties, and interact with peers and researchers at any time [30-33,37-40,43-48,50-56,58]. In these groups, all inquiries and consultations concerning the intervention were promptly handled by researchers. Nine studies used private chats for personalized follow-up communication, often in conjunction with group-based activities [33,35,38,40,42,44,54,56,59]. Eighteen studies delivered health-related information via official accounts to enhance the

XSL•FO
RenderX

knowledge of diseases [29,32,34,35,37,39,41,42,44,47,48,50-52,54,55,58,59]. Some of these studies released original content, while others shared and reposted existing credible popular science articles from other sources. Four studies were developed to achieve more individual interaction [33,36,49,57]. An applet usually contains several modules, including a symptom assessment instrument and a health education knowledge base. They could also provide educational videos and check-in tasks [36]. Notably, 19 of the 31 studies used 2 or more intervention components in combination, indicating that combined WeChat-based strategies are common in mHealth interventions.

Figure 2. Conceptual logic of interventional studies, illustrating the full workflow, participant interaction patterns, and service components of WeChat-based nursing interventions. mHealth: mobile health.



Conditions Involved

The most studied diseases include prenatal care (n=5) [34,37,42,48,51], breast cancer (n=5) [32,38,39,47,52], and gynecological cancer (n=5) [33,40,41,46,53], followed by gestational diabetes mellitus (n=4) [30,31,49,54], breastfeeding (n=2) [29,55], polycystic ovary syndrome (n=2) [36,56], pregnant with psychological distress (n=2) [45,57], and postpartum care (n=2) [35,58]. There is only one study for gynecological inflammatory diseases [43], primary dysmenorrhea [44], pelvic organ prolapse [50], and infertility [59].

Main Effectiveness of WeChat-Based Nursing Interventions

Prenatal Care

A total of 5 studies investigated WeChat-based interventions for prenatal care. Huang et al [34] reported no significant improvement in knowledge level or behavioral compliance despite increased health beliefs, whereas another two studies

XSL•FO
RenderX

showed improvements in both knowledge and adherence [42,48]. Three studies found a lower cesarean section rate [37,42,51], and two of them reported reduced postpartum hemorrhage [37,42].

Breast Cancer

All 5 studies demonstrated that WeChat-based nursing interventions exerted favorable effects on psychological recovery and quality of life on women with breast cancer. Liang et al [32] exclusively assessed objective lymphedema-related outcomes and demonstrated clear relief of lymphedema. Wang et al [38] uniquely measured fear of cancer recurrence, which was significantly reduced in the intervention group. Xu et al [52] focused on self-efficacy and treatment-related discomfort, both of which showed significant benefits. Despite no studies reporting conflicting or opposite results, 2 studies consistently found that pain, fatigue, and sleep disturbance showed no statistically significant between-group differences [38,39].

Gestational Diabetes Mellitus

Altogether, 4 studies evaluated WeChat-based interventions for glycemic management in women with gestational diabetes mellitus [30,31,49,54]. All studies confirmed significant improvements in glycemic control (reduced fasting and 2-hour postprandial blood glucose levels, higher glycemic qualification rates), enhanced self-management ability, and lower incidence of adverse maternal and neonatal outcomes. Guo et al [54] specifically observed the improved psychological status after WeChat intervention. Tian et al [30] found that earlier intervention (23 - 24 weeks) led to more stable glucose control.

Other Conditions

Two studies explored WeChat-based interventions to support breastfeeding. Both investigations yielded similar positive findings: elevated exclusive breastfeeding rates, strengthened breastfeeding self-efficacy, and boosted parental satisfaction relative to standard care [29,55]. However, Yu et al [55] centered on feeding safety in infants with congenital heart disease, while Wu et al [29] provided a longitudinal, population-level assessment of breastfeeding behaviors in healthy mother-infant dyads.

Two studies confirmed that digitally supported lifestyle modification significantly reduced body weight and BMI, and enhanced reproductive outcomes in patients with polycystic ovary syndrome relative to standard care [36,56]. These convergent results underscore the efficacy of WeChat platforms in facilitating sustainable lifestyle changes and clinical improvement.

Two studies revealed that digital mindfulness programs significantly alleviated depressive and anxiety symptoms and enhanced mindfulness capabilities [45,57]. Notably, the study of Zhang et al [57] was the first to reveal favorable ripple effects on infant neurodevelopment, with significantly better infant temperament at 6 weeks; these benefits were partially mediated by reduced pregnancy-specific anxiety.

In terms of postpartum rehabilitation, Chu et al [35] concentrated on postpartum pelvic floor muscle training, demonstrating that daily WeChat reminders more than doubled the adherence rate. By contrast, Li et al [58] evaluated comprehensive maternal health care quality, revealing that WeChat-based health education combined with specialist team support produced the highest satisfaction, followed by specialist support alone and WeChat alone, all superior to usual care.

Multidisciplinary Participant

Six studies established multidisciplinary teams for mHealth initiatives [31,39,40,45,49,51]. In addition to physicians and nurses, midwives were typically included to support pregnant women. Psychologists or psychological counselors are incorporated in mindfulness interventions [45]. Within a diet and exercise program, nutritionists or dietitians would calculate the energy requirements for each pregnant woman, while exercise specialists provided tailored advice as necessary [31]. Pharmacists were tasked with developing medication plans for patients. To assess the quality of life in patients with cervical cancer, one study included a sex counselor [40]. When medication was required, researchers collaborated with endocrinologists to achieve optimal glucose control [49].

Trend of Publication and Citation

The total number of publications shows a fluctuating trend. Prior to 2018, only a small number of articles were released annually [42,44,59]. From 2018 to 2024, there were at least 3 articles published each year, and the number of articles published was the highest in 2021 and 2024 (n=9 both). In terms of research fields, the majority of documents were categorized under obstetrics and gynecology (n=9), health care sciences services (n=7), general internal medicine (n=5), and medical informatics (n=5). Overall, this analysis emphasizes the interdisciplinary nature of the application of WeChat in women's mHealth, which has attracted increasing interest and experienced substantial growth in recent years.

Academic Output of Institutions

Analyzing institutional publications and citations offers a thorough insight into the scientific ecosystem. Altogether, 40 institutions participated in this field (Table 4). Multiple top-tier Chinese universities, including Peking University [32], Fudan University [37], Zhejiang University [45], and Shandong University [57], participated actively, alongside international institutions such as the University of Oxford [29], the National University of Singapore [38], and the University of Texas System [57]. Notably, Xi'an Jiaotong University distinguished itself as a productive contributor to the research with the highest citations (n=109) [39-41].

XSL•FO
RenderX

Table . List of organizations that published more than one paper.

Organization	Paper, n	Domestic document, n	International document, n	Times cited, n	Category Normalized Citation Impact
Shanghai Jiao Tong University	3	2	1	26	2.80
Xi'an Jiaotong University	3	2	1	109	3.19
Chinese Academy of Medical Sciences – Peking Union Medical College	3	3	0	56	1.55
Peking Union Medical College Hospital	3	3	0	56	1.55
Zhejiang University	2	2	0	71	2.99
Tongji University	2	2	0	16	2.29
Peking Union Medical College	2	2	0	53	1.98

High-Impact Highly Cited Studies

Among the included studies, several publications emerged as highly cited and representative works in this field. Zhou et al [39] conducted the most frequently referenced study, which assessed a WeChat-based multimodal nursing program for women recovering from breast cancer surgery, and found notable enhancements in quality of life. Another work demonstrated that an 8-week WeChat-based mindfulness intervention effectively reduced anxiety and depression in pregnant women, and established a feasible model for psychological care via WeChat [45]. Wu et al [29] demonstrated that WeChat official account-based health education significantly improved exclusive breastfeeding rates. Tian et al [30] showed that WeChat group-mediated health education and lifestyle support yielded better glycemic control than standard prenatal care among women with gestational diabetes mellitus. Collectively, these highly cited studies represented key milestones and laid the methodological foundation for WeChat-based mHealth interventions in women's health.

Risk-of-Bias Assessment

In 20 studies, the randomization process was judged to be at low risk of bias, whereas 7 studies presented an unclear risk of bias in this domain (Figure 3 and Table 5) [40,43,47,52-55]. Blinding participants and personnel to group assignments was not feasible in all studies except one due to the interventions' characteristics [45]. Lack of blinding may have altered participants' health information-seeking behaviors in the control group, potentially biasing the outcome assessment. With regard to missing outcome data, all included studies were considered to have a low risk of bias. Regarding the measurement of outcomes and selection of the reported result, all studies were rated as having a high risk of bias, primarily because blinding of outcome assessors was not performed, and the reporting of outcomes was insufficiently transparent. Overall, the risk of bias across studies was predominantly high, which should be considered when interpreting the findings.

XSL•FO
RenderX

Figure 3. Risk of bias assessment by the revised Cochrane risk-of-bias tool for randomized trials “Rob2” [29-33,35-47,51-59].

XSL•FO
RenderX

Table . Risk of bias in quasi-experiment studies with ROBINS-I[a] tool.

Reference	Bias due to confounding	Bias in selection of patients into the study	Bias in classification of intervention	Bias due to deviations from intended interventions	Bias due to missing data	Bias in measurement of outcomes	Bias in selection of the reported result	Overall risk
Huang et al [34] (2019)	Serious	Low	Moderate	Moderate	Low	Low	Moderate	Serious
Chen et al [48] (2020)	Serious	Low	Moderate	Moderate	Low	Low	Moderate	Serious
Yu and Guo [50] (2023)	Serious	Low	Moderate	Moderate	Low	Low	Moderate	Serious
Wang et al [49] (2024)	Serious	Low	Moderate	Moderate	Low	Low	Moderate	Serious

[a]ROBINS-I: Risk of Bias in Nonrandomized Studies of Interventions.

Discussion

To our knowledge, this is the first thorough review examining the impact of mHealth via WeChat on women. Although mHealth has been widely applied in various fields of obstetrics and gynecology, including prenatal care, postpartum health management, and gynecological disease prevention [60-62], the specific application of WeChat-based mHealth interventions in this population remains underexplored. In the included studies, WeChat-based mHealth was reported to be associated with greater access to personalized health information [63], observed to support self-management capabilities [64], and documented to enable real-time communication with health care providers [65]. This technology suggested beneficial effects on women's proactive engagement in health monitoring and improving overall reproductive health outcomes.

As a popular social media platform in China, WeChat shows promising potential as an mHealth platform. It has been applied to support chronic disease management by providing personalized health guidance and support [63]. People do not need to download software specifically, and all functions are on WeChat, which people use every day. Specifically, WeChat's group chat may help build health support groups and promote interaction and experience sharing among participants [33,66,67]. WeChat's public account provides an effective channel for health education. Through public accounts, health professionals can regularly publish relevant information on health knowledge, disease prevention, and healthy lifestyles. This model was reported to reach large audiences and allow timely updates, with potential links to health literacy [68]. Finally, applets integrate health monitoring tools and personalized health plans, potentially supporting long-term health management and behavior tracking. This personalized intervention can provide customized health advice and support based on users' specific health conditions and needs, thereby improving the effectiveness of health management [66]. Altogether, WeChat, through its diverse functions, provides strong support for the implementation of mHealth, becoming an important tool for promoting health management and health education.

Our findings revealed that the effective implementation of mHealth involved collaboration among multidisciplinary teams. Obstetricians play a central role in assessing and managing the health of pregnant women and adjusting treatment plans. Midwives provide support and care to ensure the safety and comfort of pregnant women [69]. In addition, the involvement of internal medicine physicians is crucial for managing chronic conditions in pregnant women, such as gestational diabetes or hypertension. Dietitians help pregnant women maintain a healthy weight and nutritional status through personalized dietary advice, thereby promoting the health of both mother and baby [31,51]. This multidisciplinary management model may improve the quality of obstetric care and strengthen communication and collaboration among health care providers [70].

The highly cited studies serve as pivotal clinical cornerstones for WeChat-based mHealth interventions in women's health. These landmarks cover breast cancer, perinatal mental disorders, breastfeeding difficulties, and gestational diabetes mellitus. In most cases, these conditions demand continuous and personalized self-management, which conventional outpatient care struggles to fulfill efficiently. Breast cancer patients often face various physiological and psychological challenges during and after treatment, and remote rehabilitation effectively improves patients' mental health and cognitive function [71]. Pregnant women are vulnerable to anxiety and depression and need timely psychological intervention [72]. Breastfeeding is one of the optimal ways for infants to obtain optimal nutrition; however, in practice, many mothers face various challenges. Therefore, sustained health education is crucial for the success of breastfeeding [73]. Meanwhile, women with gestational diabetes depend on regular glucose monitoring, dietary coaching, and behavioral supervision. Collectively, WeChat overcomes the temporal and spatial limitations of traditional care, enabling persistent, tailored, and responsive support.

Most eligible studies were RCTs, and RCTs are widely regarded as the gold standard in evidence-based medical research. Since mHealth interventions delivered via WeChat are difficult to blind in practice, performance bias and detection bias were almost unavoidable in the included trials. However, measurement bias and selective reporting bias are two common issues in these mHealth-related RCTs. Selective reporting bias suggests that researchers may selectively report certain outcomes while omitting or altering others to highlight statistically significant results. In leading journals in dentistry, the incidence

XSL•FO
RenderX

of selective reporting bias is 51.5%, with the most common form being inconsistent assessment times for primary outcomes [74]. Even 85% of trials did not fully adhere to prespecified results in the antipsychotic drugs field [75]. Researchers and policymakers must enhance their understanding of the various assumptions, biases, and limitations to elevate the quality of RCTs. Moreover, RCTs ought to be combined with other research techniques to create a more thorough evidence base [76]. In particular, researchers should use objective, validated tools to measure key outcomes rather than relying solely on subjective self-reported data. For instance, researchers leveraged sensor data from smartphones and smartwatches combined with machine learning algorithms to detect emotional states and transitions [77]. By integrating deep neural networks with visual indicators, the mental health status of cancer patients can be objectively evaluated, achieving a high level of classification accuracy [78]. Only with improved methodological quality can this field establish a robust evidence base and achieve stable, scalable, and trustworthy mHealth models.

Our analysis has some limitations. First, current research mostly focuses on a single cultural context, while under multicultural settings, the effectiveness of WeChat mHealth apps may vary. Culturally adapted behavioral health interventions are more effective in improving health behaviors and outcomes [79]. The process of cultural adaptation is complex and requires multiple stages, including information collection, preliminary design, testing, refinement, and final implementation. Therefore, researchers can explore the adaptability and effectiveness of WeChat mHealth apps in different cultural contexts through cross-cultural comparative studies [80]. Second, despite many studies proving the short-term benefits of WeChat mHealth, conducting long-term follow-up studies can offer a deeper understanding of its enduring impact on health management and how it changes over different periods [81,82]. Third, combining WeChat with other technological platforms (such as wearable devices and electronic health record systems) could further enhance the effectiveness of these interventions [83].

WeChat‐based mHealth interventions show promising but still uncertain potential in improving access to personalized health information, enhancing self-management, and facilitating real-time doctor-patient communication. Given the high risk of bias across most included studies, the overall strength of evidence remains low and does not support definitive effectiveness. Multidisciplinary collaboration is crucial for effective mHealth implementation. Current limitations include a single-cultural context focus and a lack of long-term follow-up. Future studies should adopt more rigorous methodologies, including improved blinding, transparent outcome reporting, and validated objective measures. Cross-cultural adaptability, long-term effectiveness, and integration with wearable technologies should also be examined in subsequent research. This review fills a research gap and provides directions for optimizing WeChat-based mHealth interventions for women.

Acknowledgments

The authors declare the use of generative artificial intelligence in the research and writing process. According to the GAIDeT taxonomy (2025), proofreading and editing tasks were delegated to Doubao under full human supervision. Responsibility for the final manuscript lies entirely with the authors.

Funding

This work is financially supported by grants from the National Natural Science Foundation of China (82301849) and the Hubei Provincial Natural Science Foundation of China (2026AFB711).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: RL, YW
Data curation: JX
Formal analysis: RL, YW, TW
Funding acquisition: RL, TW
Methodology: JX, RL, YW, TW
Supervision: JX, TW
Validation: JX
Visualization: JX, YW, TW
Writing – original draft: RL, YW
Writing – review & editing: JX, TW
All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Checklist 1
PRISMA checklist.
[PDF File, 99 KB - nursing_v9i1e82679_app1.pdf]

References

1. Cho HN, Choi E, Seo DH, et al. The Korean Study of Women's Health-Related Issues (K-Stori): rationale and study design. BMC Public Health 2017 Jun 29;17(1):609. [doi: 10.1186/s12889-017-4531-1] [Medline: 28662652]
2. Kostopoulou E, Anagnostis P, Bosdou JK, Spiliotis BE, Goulis DG. Polycystic ovary syndrome in adolescents: pitfalls in diagnosis and management. Curr Obes Rep 2020 Sep;9(3):193-203. [doi: 10.1007/s13679-020-00388-9] [Medline: 32504286]
3. Moutos CP, Ly P, Moutos DM, Clark SM. Pregnancy after 40: recommendations for counseling, evaluation, and management from preconception to delivery. Obstet Gynecol Surv 2022;77(2):111-121. [doi: 10.1097/OGX.0000000000000967]
4. McElhany K, Aggarwal S, Wood G, Beauchamp J. Protective and harmful social and psychological factors associated with mood and anxiety disorders in perimenopausal women: a narrative review. Maturitas 2024 Dec;190:108118. [doi: 10.1016/j.maturitas.2024.108118] [Medline: 39317031]
5. Haring B, Leng X, Robinson J, et al. Cardiovascular disease and cognitive decline in postmenopausal women: results from the Women's Health Initiative Memory Study. J Am Heart Assoc 2013 Dec 18;2(6):e000369. [doi: 10.1161/JAHA.113.000369] [Medline: 24351701]
6. Printz C. Older women with a family history of breast cancer face increased risk of the disease. Cancer 2018 Jul;124(13):2673-2673. [doi: 10.1002/cncr.31591]
7. Shawon MSR, Hossain FB, Hasan M, Rahman MR. Gender differences in the prevalence of anxiety and depression and care seeking for mental health problems in Nepal: analysis of nationally representative survey data. Glob Ment Health (Camb) 2024;11:e46. [doi: 10.1017/gmh.2024.37] [Medline: 38690568]
8. Wood CS, Thomas MR, Budd J, et al. Taking connected mobile-health diagnostics of infectious diseases to the field. Nature 2019 Feb;566(7745):467-474. [doi: 10.1038/s41586-019-0956-2] [Medline: 30814711]
9. Alzueta E, de Zambotti M, Javitz H, et al. Tracking sleep, temperature, heart rate, and daily symptoms across the menstrual cycle with the Oura ring in healthy women. Int J Womens Health 2022;14:491-503. [doi: 10.2147/IJWH.S341917] [Medline: 35422659]
10. Souza SCS, da Silva DF, Nagpal TS, et al. The short-term effect of a mHealth intervention on gestational weight gain and health behaviors: the SmartMoms Canada pilot study. Physiol Behav 2022 Dec;257:113977. [doi: 10.1016/j.physbeh.2022.113977]
11. Vidal C, Garcia M, Benito L, Milà N, Binefa G, Moreno V. Use of text-message reminders to improve participation in a population-based breast cancer screening program. J Med Syst 2014 Sep;38(9):118. [doi: 10.1007/s10916-014-0118-x] [Medline: 25073694]
12. Tencent announces 2025 annual and fourth quarter results. : Tencent Holdings Limited; 2026 Mar 18 URL: https://static.www.tencent.com/uploads/2026/03/18/e6a646796d0d869acc76271c9ee1a6a5.pdf [accessed 2026-06-27]
13. Zhang J, Wu Q, Wang X, Zhang Y, Helena van Velthoven M. High response rate and low cost of WeChat-based data collection of infant and young child feeding data in rural Qinghai Province, China. J Glob Health 2022 Oct 21;12:11011. [doi: 10.7189/jogh.12.11011] [Medline: 36269216]
14. Wang J, Gao F, Li J, et al. The usability of WeChat as a mobile and interactive medium in student-centered medical teaching. Biochem Mol Biol Educ 2017 Sep;45(5):421-425. [doi: 10.1002/bmb.21065] [Medline: 28581156]
15. Ma R, Deng Z, Wu M. Effects of health information dissemination on user follows and likes during COVID-19 outbreak in China: data and content analysis. Int J Environ Res Public Health 2020 Jul 14;17(14):5081. [doi: 10.3390/ijerph17145081] [Medline: 32674510]
16. He C, Wu S, Zhao Y, et al. Social media-promoted weight loss among an occupational population: cohort study using a WeChat mobile phone app-based campaign. J Med Internet Res 2017 Oct 23;19(10):e357. [doi: 10.2196/jmir.7861] [Medline: 29061555]
17. Sandborg J, Henriksson P, Larsen E, et al. Participants' engagement and satisfaction with a smartphone app intended to support healthy weight gain, diet, and physical activity during pregnancy: qualitative study within the HealthyMoms trial. JMIR Mhealth Uhealth 2021 Mar 5;9(3):e26159. [doi: 10.2196/26159] [Medline: 33666554]
18. Sakamoto JL, Carandang RR, Kharel M, et al. Effects of mHealth on the psychosocial health of pregnant women and mothers: a systematic review. BMJ Open 2022 Feb 15;12(2):e056807. [doi: 10.1136/bmjopen-2021-056807] [Medline: 35168981]
19. He Y, Huang C, He Q, Liao S, Luo B. Effects of mHealth-based lifestyle interventions on gestational diabetes mellitus in pregnant women with overweight and obesity: systematic review and meta-analysis. JMIR Mhealth Uhealth 2024 Jan 17;12:e49373. [doi: 10.2196/49373] [Medline: 38231555]
20. Ni X, Lou Y, Hu W, et al. Development of mobile health-based self-management support for patients with lung cancer: a stepwise approach. Nurs Open 2022 May;9(3):1612-1624. [doi: 10.1002/nop2.1185] [Medline: 35128821]
21. Huang L, Li VW, Yang T, et al. Mobile health applications for depression in China: a systematic review. Cureus 2022 Jul;14(7):e27299. [doi: 10.7759/cureus.27299] [Medline: 35903483]

XSL•FO
RenderX

22. Zou P, Huang A, Luo Y, Tchakerian N, Zhang H, Zhang C. Effects of using WeChat/WhatsApp on physical and psychosocial health outcomes among oncology patients: a systematic review. Health Informatics J 2023;29(1). [doi: 10.1177/14604582231164697] [Medline: 36949645]
23. Yang J, Yang H, Wang Z, et al. Self-management among type 2 diabetes patients via the WeChat application: a systematic review and meta-analysis. J Clin Pharm Ther 2021 Feb;46(1):4-16. [doi: 10.1111/jcpt.13264] [Medline: 33085789]
24. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg 2021 Apr;88:105906. [doi: 10.1016/j.ijsu.2021.105906] [Medline: 33789826]
25. McLean KA, Knight SR, Diehl TM, et al. Readiness for implementation of novel digital health interventions for postoperative monitoring: a systematic review and clinical innovation network analysis. Lancet Digit Health 2023 May;5(5):e295-e315. [doi: 10.1016/S2589-7500(23)00026-2] [Medline: 37100544]
26. Ramachandran HJ, Jiang Y, Tam WWS, Yeo TJ, Wang W. Effectiveness of home-based cardiac telerehabilitation as an alternative to phase 2 cardiac rehabilitation of coronary heart disease: a systematic review and meta-analysis. Eur J Prev Cardiol 2022 May 25;29(7):1017-1043. [doi: 10.1093/eurjpc/zwab106] [Medline: 34254118]
27. Crocker TF, Lam N, Jordão M, et al. Risk-of-bias assessment using Cochrane's revised tool for randomized trials (RoB 2) was useful but challenging and resource-intensive: observations from a systematic review. J Clin Epidemiol 2023 Sep;161:39-45. [doi: 10.1016/j.jclinepi.2023.06.015] [Medline: 37364620]
28. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919. [doi: 10.1136/bmj.i4919] [Medline: 27733354]
29. Wu Q, Huang Y, Liao Z, van Velthoven MH, Wang W, Zhang Y. Effectiveness of WeChat for improving exclusive breastfeeding in Huzhu County China: randomized controlled trial. J Med Internet Res 2020 Dec 3;22(12):e23273. [doi: 10.2196/23273] [Medline: 33270026]
30. Tian Y, Zhang S, Huang F, Ma L. Comparing the efficacies of telemedicine and standard prenatal care on blood glucose control in women with gestational diabetes mellitus: randomized controlled trial. JMIR Mhealth Uhealth 2021 May 25;9(5):e22881. [doi: 10.2196/22881] [Medline: 33783365]
31. Deng Y, Hou Y, Wu L, Liu Y, Ma L, Yao A. Effects of diet and exercise interventions to prevent gestational diabetes mellitus in pregnant women with high-risk factors in China: a randomized controlled study. Clin Nurs Res 2022 Jun;31(5):836-847. [doi: 10.1177/10547738211055576] [Medline: 34775875]
32. Liang X, You M, Wen C, et al. Self-administration of complex decongestive therapy facilitated by the mobile application WeChat improves lymphedema and quality of life in breast cancer survivors: an observational study. Ann Transl Med 2022 Feb;10(3):146. [doi: 10.21037/atm-21-6662] [Medline: 35284545]
33. Tian X, Liu Y, Zhang J, et al. Efficacy of a WeChat-based, multidisciplinary, full-course nutritional management program on the nutritional status of patients with ovarian cancer undergoing chemotherapy: randomized controlled trial. JMIR Mhealth Uhealth 2024 Nov 4;12:e56475. [doi: 10.2196/56475] [Medline: 39496160]
34. Huang L, Sun Y, Zhang J. A WeChat intervention program to improve self-monitoring of fetal movement among pregnant women: development and validation [Article in Chinese]. Nurs J Chin People's Liberation Army 2019(4):38-41. [doi: 10.3969/j.issn.1008-9993.2019.04.009]
35. Chu L, Jin X, Wu S, Tong X, Li H, Chen X. Effect of pelvic floor muscle training with smartphone reminders on women in the postpartum period: a randomized controlled trial. Urogynecology (Phila) 2024 Feb 1;30(2):138-146. [doi: 10.1097/SPV.0000000000001401] [Medline: 37556387]
36. Dilimulati D, Shao X, Wang L, et al. Efficacy of WeChat-based digital intervention versus metformin in women with polycystic ovary syndrome: randomized controlled trial. J Med Internet Res 2024 Oct 2;26:e55883. [doi: 10.2196/55883] [Medline: 39357046]
37. Gu C, Guo L, Zhu X, et al. Clinical evaluation of the implementation outcomes of the maternity care program for low-risk pregnant women [Article in Chinese]. J Nurs Sci 2024(15):27-32. [doi: 10.3870/j.issn.1001-4152.2024.15.027]
38. Wang LP, Yao LH, Wang D, et al. A WeChat-based nursing intervention program improves the postoperative rehabilitation of breast cancer patients: results from a randomized controlled trial. Gland Surg 2024 Mar 27;13(3):383-394. [doi: 10.21037/gs-24-40] [Medline: 38601277]
39. Zhou K, Wang W, Zhao W, et al. Benefits of a WeChat-based multimodal nursing program on early rehabilitation in postoperative women with breast cancer: a clinical randomized controlled trial. Int J Nurs Stud 2020 Jun;106:103565. [doi: 10.1016/j.ijnurstu.2020.103565] [Medline: 32272281]
40. Han D, Wang D, Yang J, Li X. Effect of multidisciplinary collaborative continuous nursing on the psychological state and quality of life of patients with cervical cancer. Am J Transl Res 2021;13(6):6654-6661. [Medline: 34306409]
41. Li M, Chan CWH, Chan D, et al. The effectiveness of WeChat couple-based psychosocial support for gynaecological cancer: a randomised controlled trial. BJOG 2023 May;130(6):599-609. [doi: 10.1111/1471-0528.17375] [Medline: 36648406]
42. Guan LI, Wang YH. Application of WeChat platform in prenatal care in pregnant women [Article in Chinese]. J China Med Univ 2014 Dec;43(12):1147-1149. [doi: 10.3969/j.issn.0258-4646.2014.12.025]

XSL•FO
RenderX

43. Wang A, Li N, Fan X. PDCA circulation combined with continuing nursing guided by WeChat on improving the nursing value of patients with gynecological inflammatory diseases during out-of-hospital treatment. Food Sci Technol 2021 Jul;41(3):737-742. [doi: 10.1590/fst.19520]
44. Su M, Du C, Sun L, Sun X. Evaluation of health education for female college students of dysmenorrhea based on WeChat platform [Article in Chinese]. Chin J Sch Health 2016;37(12):1804-1807. [doi: 10.16835/j.cnki.1000-9817.2016.12.014]
45. Yang M, Jia G, Sun S, Ye C, Zhang R, Yu X. Effects of an online mindfulness intervention focusing on attention monitoring and acceptance in pregnant women: a randomized controlled trial. J Midwife Womens Health 2019 Jan;64(1):68-77. [doi: 10.1111/jmwh.12944]
46. Lu X, Zhang L, He J. Significance of WeChat continuous management combined with drug therapy in the treatment of persistent infection of high-risk human papillomavirus in postmenopausal women [Article in Chinese]. Chin J Primary Med Pharm 2021;28(9):1383-1387. [doi: 10.3760/cma.j.issn.1008-6706.2021.09.021]
47. Wu Q, Liu Q, Pan S. Effect of WeChat platform based continuity of care on the quality of life and psychosocial competence in postoperative patients with breast cancer. Int J Clin Exp Med 2021(2):1162-1171 [FREE Full text]
48. Chen J, Huang J, Ooi S, et al. Effect of flexible patterns of health education on enhancing the compliance of pregnant women from Tibet, China. Medicine (Baltimore) 2020;99(1):e18447. [doi: 10.1097/MD.0000000000018447]
49. Wang Q, Zhang K, Zhang X, et al. WeChat mini-program, a preliminary applied study of the gestational blood glucose management model for pregnant women with gestational diabetes mellitus. Diabetes Res Clin Pract 2025 Jan;219:111943. [doi: 10.1016/j.diabres.2024.111943] [Medline: 39615794]
50. Yu S, Guo C. Influence of health education intervention based on WeChat platform on postpartum women's pelvic floor function exercise cognition and pelvic floor function [Article in Chinese]. Chin J Health Psychol 2023;31(10):1463-1468. [doi: 10.13342/j.cnki.cjhp.2023.10.005]
51. Wang L, Zou L, Yi H, et al. The implementation of online and offline hybrid weight management approach for pregnant women based on the Fogg behavior model in Hainan, China: a pilot randomized controlled trial. BMC Pregnancy Childbirth 2024;24(1). [doi: 10.1186/s12884-024-06699-2]
52. Xu YF, Fei. Xu X, Song K, et al. Effects of extended care based on the WeChat platform on self-efficacy and quality of life of postoperative breast cancer patients. Indian J Pharm Sci 2021;83:23-30. [doi: 10.36468/pharmaceutical-sciences.spl.165]
53. Zou Q, Zhang G, Liu Y. Health education using telephone and WeChat in treatment of symptomatic uterine myoma with high-intensity focused ultrasound. Med Sci Monit Basic Res 2018 Sep 6;24:127-133. [doi: 10.12659/MSMBR.911040] [Medline: 30185767]
54. Guo Y, Zhou L, Sun B, Wang C, Zhang J. Application of online-offline integrated medical care management in patients with gestational diabetes. Ginekol Pol 2021;92(10):720-725. [doi: 10.5603/GP.a2021.0054] [Medline: 33914304]
55. Yu XR, Xie WP, Liu JF, Dai WS, Cao H, Chen Q. Effect of WeChat follow-up education on breastfeeding efficiency of infants after congenital cardiac surgery and parents' satisfaction: a short-term follow-up study. J Paediatr Child Health 2021 Dec;57(12):1899-1904. [doi: 10.1111/jpc.15605] [Medline: 34085341]
56. Sang M, Wu Q, Tao Y, et al. Usage of mobile health interventions among overweight/obese PCOS patients undergoing assisted reproductive technology treatment during the COVID-19 pandemic. Gynecol Endocrinol 2022 Sep;38(9):776-780. [doi: 10.1080/09513590.2022.2112170] [Medline: 35993296]
57. Zhang X, Li Y, Wang J, et al. Effectiveness of digital guided self-help mindfulness training during pregnancy on maternal psychological distress and infant neuropsychological development: randomized controlled trial. J Med Internet Res 2023 Feb 10;25:e41298. [doi: 10.2196/41298] [Medline: 36763452]
58. Li C, Tang L, Yang M, et al. A study to evaluate the efficacy of different interventions for improving quality of maternal health care service in China. Telemed J E Health 2020 Oct;26(10):1291-1300. [doi: 10.1089/tmj.2019.0230] [Medline: 31928505]
59. Li XQ, Sun CF, Guo M. Benefits of nursing care service in the assisted reproduction clinic to self-cycle-management and self-efficiency of infertility patients. Zhonghua Nan Ke Xue 2017 Jun;23(6):536-539. [Medline: 29722947]
60. Inhae C, Jiwon K. Effects of mHealth intervention on sexual and reproductive health in emerging adulthood: a systematic review and meta-analysis of randomized controlled trials. Int J Nurs Stud 2021 Jul;119:103949. [doi: 10.1016/j.ijnurstu.2021.103949] [Medline: 34004469]
61. van den Heuvel JF, Groenhof TK, Veerbeek JH, et al. eHealth as the next-generation perinatal care: an overview of the literature. J Med Internet Res 2018 Jun 5;20(6):e202. [doi: 10.2196/jmir.9262] [Medline: 29871855]
62. Watterson JL, Walsh J, Madeka I. Using mHealth to improve usage of antenatal care, postnatal care, and immunization: a systematic review of the literature. Biomed Res Int 2015;2015:153402. [doi: 10.1155/2015/153402] [Medline: 26380263]
63. Chen X, Zhou X, Li H, Li J, Jiang H. The value of WeChat application in chronic diseases management in China. Comput Methods Programs Biomed 2020 Nov;196:105710. [doi: 10.1016/j.cmpb.2020.105710]
64. Ge Y, Yao D, Ung COL, et al. Digital medical information services delivered by pharmaceutical companies via WeChat: qualitative analytical study. J Med Internet Res 2023 Nov 17;25:e43812. [doi: 10.2196/43812] [Medline: 37976079]

XSL•FO
RenderX

65. Tang H, Zhu Z, Ying J, et al. The prognostic role of WeChat-based instant doctor-patient communication in patients with advanced pancreatic cancer. J Cancer Surviv 2024 Dec;18(6):1903-1911. [doi: 10.1007/s11764-023-01429-7] [Medline: 37460857]
66. Li T, Ding W, Li X, Lin A. Mobile health technology (WeChat) for the hierarchical management of community hypertension: protocol for a cluster randomized controlled trial. Patient Prefer Adherence 2019;13:1339-1352. [doi: 10.2147/PPA.S215719] [Medline: 31496664]
67. Sun W, Yu L, Liu S, et al. Comparison of maternal and neonatal outcomes for patients with placenta accreta spectrum between online-to-offline management model with standard care model. Eur J Obstet Gynecol Reprod Biol 2018 Mar;222:161-165. [doi: 10.1016/j.ejogrb.2018.01.035] [Medline: 29408749]
68. Li Y, Xiao QL, Li M, et al. Community-based intervention via WeChat official account to improve parental health literacy among primary caregivers of children aged 0 to 3 years: protocol for a cluster randomized controlled trial. Front Public Health 2022;10:1039394. [doi: 10.3389/fpubh.2022.1039394] [Medline: 36684867]
69. Yang P, Lo W, He ZL, Xiao XM. Medical nutrition treatment of women with gestational diabetes mellitus by a telemedicine system based on smartphones. J Obstet Gynaecol Res 2018 Jul;44(7):1228-1234. [doi: 10.1111/jog.13669] [Medline: 29797375]
70. Brown Will SE. Perioperative nursing care in obstetrics and gynecology. AORN J 2019 Aug;110(2):130-133. [doi: 10.1002/aorn.12771] [Medline: 31355432]
71. Bahçaci U, Uysal SA, Namal E. Relaxation training via tele-rehabilitation program in patients with breast cancer receiving chemotherapy during COVID-19. Digit Health 2024;10. [doi: 10.1177/20552076241261909] [Medline: 39610824]
72. Bayat A, Amiri-Farahani L, Soleimani M, Eshraghi N, Haghani S. Effect of short-term psychological intervention on anxiety of pregnant women with positive screening results for chromosomal disorders: a randomized controlled trial. BMC Pregnancy Childbirth 2021 Nov 9;21(1):757. [doi: 10.1186/s12884-021-04206-5] [Medline: 34753431]
73. Lumbiganon P, Martis R, Laopaiboon M, Festin MR, Ho JJ, Hakimi M. Antenatal breastfeeding education for increasing breastfeeding duration. Cochrane Database Syst Rev 2012 Sep 12;12(9):CD006425. [doi: 10.1002/14651858.CD006425.pub3] [Medline: 22972092]
74. Wang Y, Guo F, Chen X, Yu R, Qin D, Hua F. Selective outcome reporting among randomized controlled trials published in leading dental journals: a research-on-research study. J Dent 2024 Dec;151:105448. [doi: 10.1016/j.jdent.2024.105448] [Medline: 39489327]
75. Lancee M, Lemmens CMC, Kahn RS, Vinkers CH, Luykx JJ. Outcome reporting bias in randomized-controlled trials investigating antipsychotic drugs. Transl Psychiatry 2017 Sep 12;7(9):e1232. [doi: 10.1038/tp.2017.203] [Medline: 28895941]
76. Krauss A. Why all randomised controlled trials produce biased results. Ann Med 2018 Jun;50(4):312-322. [doi: 10.1080/07853890.2018.1453233] [Medline: 29616838]
77. Sultana M, Al-Jefri M, Lee J. Using machine learning and smartphone and smartwatch data to detect emotional states and transitions: exploratory study. JMIR Mhealth Uhealth 2020 Sep 29;8(9):e17818. [doi: 10.2196/17818]
78. Shafiei SB, Lone Z, Elsayed AS, Hussein AA, Guru KA. Identifying mental health status using deep neural network trained by visual metrics. Transl Psychiatry 2020 Dec 14;10(1):430. [doi: 10.1038/s41398-020-01117-5] [Medline: 33318471]
79. Barrera M, Castro FG, Strycker LA, Toobert DJ. Cultural adaptations of behavioral health interventions: a progress report. J Consult Clin Psychol 2013 Apr;81(2):196-205. [doi: 10.1037/a0027085] [Medline: 22289132]
80. Maar MA, Yeates K, Perkins N, et al. A Framework for the study of complex mHealth interventions in diverse cultural settings. JMIR Mhealth Uhealth 2017 Apr 20;5(4):e47. [doi: 10.2196/mhealth.7044] [Medline: 28428165]
81. Mönninghoff A, Kramer JN, Hess AJ, et al. Long-term effectiveness of mHealth physical activity interventions: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2021 Apr 30;23(4):e26699. [doi: 10.2196/26699] [Medline: 33811021]
82. Guo Y, Xu Z, Qiao J, et al. Development and feasibility testing of an mHealth (text message and WeChat) intervention to improve the medication adherence and quality of life of people living with HIV in China: pilot randomized controlled trial. JMIR Mhealth Uhealth 2018 Sep 4;6(9):e10274. [doi: 10.2196/10274] [Medline: 30181109]
83. Marzano L, Bardill A, Fields B, et al. The application of mHealth to mental health: opportunities and challenges. Lancet Psychiatry 2015 Oct;2(10):942-948. [doi: 10.1016/S2215-0366(15)00268-0] [Medline: 26462228]

Abbreviations

CNKI: China National Knowledge Infrastructure
mHealth: mobile health
PICO: population, intervention, comparison, and outcomes framework
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
ROBINS-I: Risk of Bias in Nonrandomized Studies of Interventions

XSL•FO
RenderX

Edited by J Sarvestan; submitted 20.Aug.2025; peer-reviewed by GLP Bodagala, S Wei; revised version received 29.May.2026; accepted 03.Jun.2026; published 06.Jul.2026.

Please cite as:
Xiao J, Li R, Wu Y, Wu T
WeChat-Based Nursing Interventions in Women's Mobile Health: Systematic Review
JMIR Nursing 2026;9:e82679
URL: https://nursing.jmir.org/2026/1/e82679
doi:10.2196/82679

© Jun Xiao, Rui Li, Yangyang Wu, Tong Wu. Originally published in JMIR Nursing (https://nursing.jmir.org), 6.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

360° Contextual Simulation Videos for Undergraduate Nursing Students: Electroencephalography-Based Quasi-Experimental Study

Yi-Chang Chen[1*], PhD; Chen-Jung Chen[2*], PhD

[1]Institute of Long-Term Care, MacKay Medical University, New Taipei City, Taiwan

[2]Department of Nursing, MacKay Medical University, No. 46, Sec. 3, Zhongzheng Rd., Sanzhi Dist., New Taipei City, Taiwan

[*]all authors contributed equally

Corresponding Author:
Chen-Jung Chen, PhD
Department of Nursing, MacKay Medical University, No. 46, Sec. 3, Zhongzheng Rd., Sanzhi Dist., New Taipei City, Taiwan

Abstract

Background: Screen-based 360° contextual simulation videos are increasingly being used in nursing education; however, evidence regarding their effects on learning outcomes and neural electroencephalography (EEG) patterns remains limited.

Objective: This study aims to evaluate whether integrating 360° contextual simulation videos with traditional classroom instruction influences nursing students' perceptions of simulation design and educational practices, their levels of satisfaction and self-confidence in learning, and EEG-based neural correlates of action observation.

Methods: In a quasi-experiment, 55 third-year nursing students (experimental n=28; control n=27) were assigned to 360° contextual simulation video learning (laptop; no head-mounted display) or traditional classroom-based instruction (didactic instruction, guided discussion, and 2D instructional videos). The primary outcomes were Simulation Design Scale (SDS) and Educational Practices Questionnaire (EPQ) scores. Secondary outcomes included Student Satisfaction and Self-Confidence in Learning Scale (SCLS) scores and EEG indices (power at C3, Cz, and C4; μ-rhythm suppression scores). Group differences were tested with 2-tailed *t* tests and analysis of covariance adjusted for baseline values.

Results: Between-group analyses (analysis of covariance adjusted for baseline values) revealed higher posttest scores on the SDS ($F_{1,35}$=7.80, η^2=.182; *P*=.008), EPQ ($F_{1,35}$=6.00, η^2=.146; *P*=.02), and SCLS ($F_{1,35}$=7.01, η^2=.167; *P*=.01) in the experimental group compared to the control group. Within-group analyses (paired *t* tests) indicated significant increases in SDS and EPQ scores for the experimental group (*P*<.001 and *P*=.005), whereas the control group showed a decrease in the SDS score (*P*=.001) and no significant change in the EPQ score (*P*=.11). The SCLS score did not significantly change in the experimental group (*P*=.21) but increased in the control group (*P*=.04). EEG analyses revealed significant postexposure reductions in power at C3, Cz, and C4 for the experimental group (all *P*<.001), with between-group effects at Cz ($F_{1,35}$=425.29, η^2=.926; *P*<.001) and C4 ($F_{1,35}$=5.45, η^2=.138; *P*=.03). μ-Suppression, which is an indirect EEG correlate that is often reported during action observation, was greater in the experimental group at Cz (*P*=.03) and C4 (*P*=.03) but not at C3 (*P*=.13). No adverse events were reported.

Conclusions: Integrating 360° contextual simulation videos into undergraduate nursing education was associated with higher SDS and EPQ scores and improved self-reported satisfaction and self-confidence, in addition to EEG-based findings that may reflect neural EEG patterns associated with action observation. Larger multisite randomized trials are warranted to confirm efficacy and cost-effectiveness.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2500102861; https://www.chictr.org.cn/showproj.html?proj=232700

(JMIR Nursing 2026;9:e84720) doi:10.2196/84720

KEYWORDS

computer; education; electroencephalography; EEG; learning; nursing; satisfaction; simulation; training; video

Introduction

Problem and Rationale

Simulation-based learning has emerged as an essential pedagogical approach in undergraduate nursing education, especially for skill-based disciplines, such as nursing [1,2]. In particular, simulation-based learning has significant value in resource-limited settings where clinical or practical training opportunities may be restricted [1,3]. In undergraduate nursing education, repeated supervised practice is often constrained by limited clinical placements and patient access, limited faculty supervision time, and the limited availability of high-fidelity

XSL•FO
RenderX

simulation facilities. Therefore, simulation-based learning can provide a standardized and safe environment for repeated observation, practice preparation, and feedback to support foundational skill acquisition and clinical decision-making [1-4]. Contextual simulations replicate real-world scenarios in controlled learning environments, thus allowing students to develop critical skills in a safe and structured manner [2]. This approach has been widely recognized for its ability to improve learners' knowledge, skills, and behavioral performance and to support clinical decision-making abilities in health professions education [1,4]. The effectiveness of simulation-based education has been well documented in nursing education and broader health professions education [1,4].

What Is Known About Screen-Based 360° Contextual Simulation Videos

Screen-based 360° contextual simulation videos (laptop, no head-mounted display) can provide standardized exposure to nursing procedures and clinical contexts while allowing learners to control their viewing direction during action observation, which is particularly relevant for early-stage skill acquisition and preparation before supervised practice [5,6]. With recent advancements in more immersive panoramic viewing experience learning technologies, screen-based 360° contextual simulation videos have been identified as innovative alternatives to conventional instructional methods across undergraduate nursing education contexts [5,6]. Recent systematic reviews and meta-analyses have also supported the effectiveness of immersive virtual reality in nursing education for improving learners' knowledge, skills, and problem-solving abilities [7]. Unlike standard fixed-view videos, our screen-based 360° contextual simulation video system allowed learners to actively control their viewing direction while observing context-rich clinical procedures on a laptop screen without a head-mounted display [5,6].

Knowledge Gap

Despite the increasingly widespread adoption of screen-based 360° contextual simulation videos in undergraduate nursing skills training, empirical evidence remains limited, particularly in resource-limited educational settings where supervised practice opportunities are constrained by limitations with respect to clinical patient access, faculty time, and access to high-fidelity simulation facilities [3]. In this context, evidence of the efficacy of screen-based 360° contextual simulation videos—especially regarding observational learning during procedural skill observation and associated electroencephalography (EEG)-based indices during action observation (eg, μ-rhythm suppression)—remains sparse [6,8]. Existing studies have rarely examined both (1) learners' perceptions of simulation quality and educational practices and (2) learners' satisfaction and self-confidence alongside EEG-based indices during action observation [6]. Clarifying these outcomes together may help educators understand not only how students perceive screen-based 360° contextual simulation videos but also how such exposure relates to EEG patterns that are commonly reported during procedural observation. However, our understanding of the direct impact of screen-based 360° contextual simulation videos on cognitive and neural mechanisms, particularly with respect to skill acquisition, learning retention, and neurophysiological engagement, remains limited [6,9]. In addition, the use of EEG to analyze neural correlates in contexts involving screen-based 360° contextual simulation video learning remains relatively underexplored [9]. Accordingly, this study addresses this gap using a 2-group comparison to examine students' perceived simulation design and educational practices, satisfaction and self-confidence in learning, and sensorimotor EEG indices during procedural action observation in response to a screen-based 360° contextual simulation video intervention. Furthermore, investigations into how these simulations influence neural engagement, particularly through observational learning processes, have been limited.

Study Justification and Aim

Mirror neuron theory proposes that observing others' actions can engage the neural systems involved in action execution, supporting observational learning [10]. In EEG research, μ-rhythm suppression over sensorimotor regions is commonly reported during action observation and is often interpreted as an indirect neural correlate that is consistent with mirror neuron system engagement; however, it is not a direct measure of mirror neuron EEG patterns [8,11]. This framework offers a plausible neurophysiological explanation, suggesting that observing tasks in contextual settings stimulates neural circuits that are associated with action execution and cognitive processing and may support motor skill acquisition and memory retention [8,10]. Specifically, EEG studies have identified μ-rhythm suppression as an indirect marker that is consistent with mirror neuron system engagement in the context of observational learning within immersive environments [8,11]. The Kolb experiential learning framework was used as a guiding lens through which to interpret how video-based exposure and facilitated reflection may support early phases of skill learning in undergraduate nursing education [12]. Additionally, recent neuroeducation research suggests that observational learning in immersive media may engage the mirror neuron system, as indexed by EEG μ-rhythm suppression [8,11]. Furthermore, cognitive load theory provides a useful design lens for screen-based simulation by emphasizing strategies that reduce extraneous load and support meaningful processing during procedural learning [13].

To our knowledge, few studies have directly integrated EEG-based learning analytics with screen-based 360° contextual simulation videos in undergraduate education. As multimodal learner data—including biometric signals—are increasingly leveraged in artificial intelligence–enhanced learning analytics to enable real-time feedback and personalization, understanding how immersive simulations influence neural engagement is important for informing adaptive educational models [14]. The aim of this study was to evaluate the effects of screen-based 360° contextual simulation videos on undergraduate nursing students' (1) perceived simulation design and educational practices, (2) satisfaction and self-confidence in learning, and (3) sensorimotor EEG indices during procedural action observation (sensorimotor power and μ-rhythm suppression) by comparing an experimental group with a control group using questionnaires and EEG measures. The screen-based 360° contextual simulation videos were delivered on a laptop screen

XSL•FO
RenderX

without a head-mounted display; therefore, the intervention is distinct from headset-based virtual reality simulation.

Methods

Design

This study had a 2-group, pretest-posttest quasi-experimental design with 2 assessment time points. The participants were assigned to either an experimental group or a control group using a computer-generated allocation sequence. However, because recruitment was convenience-based within a single university cohort and blinding was not implemented, the study is reported as quasi-experimental. The reporting of this study was guided by the CONSORT (Consolidated Standards of Reporting Trials) 2010 checklist, where applicable (Checklist 1).

Setting and Participants

Participants were recruited through announcements made in undergraduate courses and departmental emails within a university setting, and a nursing cohort was used as the case sample. An a priori power analysis was conducted using G*Power to estimate the minimum sample size planned for an analysis of covariance (ANCOVA) framework (2 groups, 1 covariate), assuming a medium effect size (f=0.25) in the absence of robust preliminary data, with α=.05 and power=0.80, yielding a minimum required sample of 34. The inclusion criteria for this study were as follows: (1) being a third-year undergraduate student in a nursing program, (2) being able to converse in Taiwanese and Mandarin, and (3) being willing to participate voluntarily in the research. The exclusion criteria were deafness or cognitive impairment according to various sources (such as school health information and parental notification). The participants were assigned to either the experimental group (screen-based 360° contextual simulation video learning) or the control group (dose-matched screen-based 2D instructional video learning), with the video learning conducted during after-class skills practice sessions.

Procedures

The study was conducted over 1 week in July 2024 and included 2 assessment time points: baseline (T0) and immediately postintervention (T1). On day 1, all participants completed baseline questionnaires and underwent baseline EEG recording. Pretest EEG recording was conducted in the eyes-open resting condition and did not involve viewing any simulation content or videos. After the baseline assessment, the participants were allocated to either the experimental or control group. Both groups completed 4 learning modules within the week. Each module followed the same overall structure in both groups: a brief orientation, video viewing on a laptop (approximately 10 - 12 min), and a 20-minute instructor-facilitated discussion using structured prompts; each module was designed to be completed within approximately 60 minutes. Immediately after the final module, posttest questionnaires were administered and collected under standardized proctoring by faculty members and teaching assistants, and the participants underwent posttest EEG recording. This procedure was identical for the control group; the posttest EEG data were recorded immediately after the fourth control module in the same eyes-open resting condition.

Theoretical Framework

As noted in the *Introduction* section, the Kolb experiential learning cycle informed the design and interpretation of the intervention; the operationalization of each stage in this study is described below. Concrete experience was operationalized as students' exposure to screen-based 360° procedural scenarios. Reflective observation was supported through structured, instructor-guided debriefing immediately after viewing. Abstract conceptualization was supported by linking the observed actions to key procedural principles and checklists during the discussion and through assigned review materials. Active experimentation was operationalized as students' subsequent application of the observed steps during routine skills practice (outside the 360° viewing session) using the same procedural checklist and course materials [12].

Interventions

The study included 2 intervention conditions delivered via the university's institutional eLearning platform and accessed through a standard web browser on laptop computers. The 4 modules covered (1) the collection of urine specimens from a urine drainage bag, (2) the measurement and documentation of vital signs, (3) infant and pediatric cardiopulmonary resuscitation, and (4) chest physiotherapy. The modules were matched by topic and intended duration across groups and were completed during after-class skills practice sessions within the 1-week study period.

Experimental Condition: Screen-Based 360° Contextual Simulation Videos

The experimental intervention consisted of screen-based 360° contextual simulation video learning delivered on laptop computers without a head-mounted display. In this study, "immersion" refers to panoramic 360° first-person viewing of realistic clinical procedures with context-rich audiovisual cues, and "interactivity" refers to learners' ability to actively control their viewing direction (through panning and rotating) during observation. Because the videos were presented on a laptop screen, the level of immersion and interaction was lower than that of fully immersive headset-based virtual reality. All screen-based 360° contextual simulation videos were self-developed by the nursing faculty and information engineering professionals through interdisciplinary discussions and were delivered via the institutional eLearning platform. Before each module, the students received a brief orientation that included learning objectives, instructions on how to navigate the 360° video (eg, how to change viewing direction), and expectations for participation. Each module included viewing of the corresponding screen-based 360° contextual simulation video (approximately 10 - 12 min), followed by a 20-minute facilitated contextual debriefing led by faculty members using structured prompts focused on key steps, safety considerations, and clinical decision-making. Each module was designed to be completed within approximately 60 minutes and focused on a single, discrete nursing skill. The participants had no prior

XSL•FO
RenderX

exposure to the 4 study-specific 360° modules before baseline assessment.

Control Condition: Dose-Matched 2D Instructional Videos

The control group covered the same 4 skill topics and used dose-matched, screen-based 2D instructional videos and the same overall module structure and scheduling. The 2D instructional videos were developed by the same teaching team and delivered through the same institutional eLearning platform accessed on a laptop via a web browser. Before each module, the students received a brief orientation that included learning objectives, instructions on how to access and view the 2D video (fixed-view, with no panning or rotation control), and expectations for participation. The educational content included (1) didactic instruction and demonstration of key procedural steps, (2) viewing of the corresponding 2D instructional video (approximately 10‐12 min), and (3) a 20-minute instructor-facilitated discussion using the same structured prompts as the experimental condition (key steps, safety considerations, and clinical decision-making). Each module was designed to be completed within approximately 60 minutes and was matched to the experimental condition by topic and duration. The control activities did not involve high-fidelity mannequin simulation, standardized-patient scenarios, or 360° video viewing.

Instructional Design and Standardization (Cognitive Load-Informed Features)

The intervention was designed with cognitive load theory–informed considerations to minimize extraneous demands and support learning efficiency. Across both conditions, each video was kept short (approximately 10‐12 min) and focused on a single skill topic. Key procedural steps were standardized using step-by-step procedural checklists reinforced during the instructor-guided discussion, and the learning environment and delivery platform were kept consistent across groups. We did not directly measure cognitive load; therefore, these elements are reported as design features rather than demonstrated mechanisms. Accordingly, unmeasured cognitive load and related factors (eg, attention or fatigue) may have influenced EEG indices and learning outcomes, which should be interpreted cautiously.

Questionnaire Measures

Quantitative data were collected via 1 demographic questionnaire and 3 additional questionnaires, and perceptions of simulation design and educational practices were assessed using the Simulation Design Scale (SDS) and the Educational Practices Questionnaire (EPQ), respectively. Satisfaction and self-confidence in learning were assessed using the Student Satisfaction and Self-Confidence in Learning Scale (SCLS). All of these instruments are contextual simulation research questionnaires developed by the National League for Nursing [15]. The reliability of the questionnaires was evaluated according to Cronbach α coefficient. Experts have recommended the use of these measurement questionnaires when implementing contextual simulations in education programs [15].

The SDS is composed of 20 items and is used to assess students' views concerning design elements pertaining to simulation and their importance. The questions address course themes and knowledge, support, problem-solving, feedback, and simulation contexts. In terms of reliability, the Cronbach α coefficient for the overall SDS was 0.96. The α coefficients for the knowledge of course themes, support, problem-solving, feedback, and simulation context scales were 0.92, 0.92, 0.86, 0.90, and 0.87, respectively. In terms of validity, the root mean square error of approximation (RMSEA), normed fit index (NFI), weighted root mean square residual (WRMR), and chi-square significance tests indicated a poor fit, whereas the comparative fit index (CFI), Tucker-Lewis index (TLI), and adjusted goodness-of-fit index (AGFI) values were greater than the commonly accepted thresholds. The conceptual model accounted for 17 out of 20 SDS items (85%). The correlations between different pairs of theoretical factors ranged between 0.67 and 0.89 [16].

The EPQ comprises 16 questions and is used to evaluate students' views concerning the importance of using simulations in virtual environments in educational practice. The responses are scored on a 5-point Likert-type scale. These 16 items cover various subscales, including active learning, cooperation, varied learning styles, and high expectations. In terms of reliability, the Cronbach α coefficient for the overall EPQ was 0.95. The α coefficients for the active learning, cooperation, diversified learning style, and high expectation subscales were 0.93, 0.90, 0.88, and 0.88, respectively. In terms of validity, the RMSEA, WRMR, and chi-square significance tests indicated that the fit was not good, whereas the CFI, TLI, NFI, and AGFI values were greater than the commonly acceptable thresholds. The conceptual model accounted for 13 out of 16 EPQ items (approximately 80%). The correlations between different pairs of conceptual factors ranged between 0.77 and 0.86 [16].

The SCLS comprises 14 items and is used to measure participants' satisfaction with and confidence in the current learning protocol on a 5-point Likert-type scale. In terms of reliability, the Cronbach α coefficient for the entire SCLS was 0.92, and the α coefficients for the satisfaction and confidence subscales were 0.92 and 0.83, respectively. In terms of validity, the RMSEA, WRMR, and chi-square significance tests indicated that the degree of fit was not good, whereas the CFI, TLI, NFI, and AGFI values were greater than the commonly accepted thresholds. The conceptual model accounted for 11 out of 14 (approximately 76%) SCLS items. The correlation coefficient between the satisfaction and confidence factors was 0.78 [16].

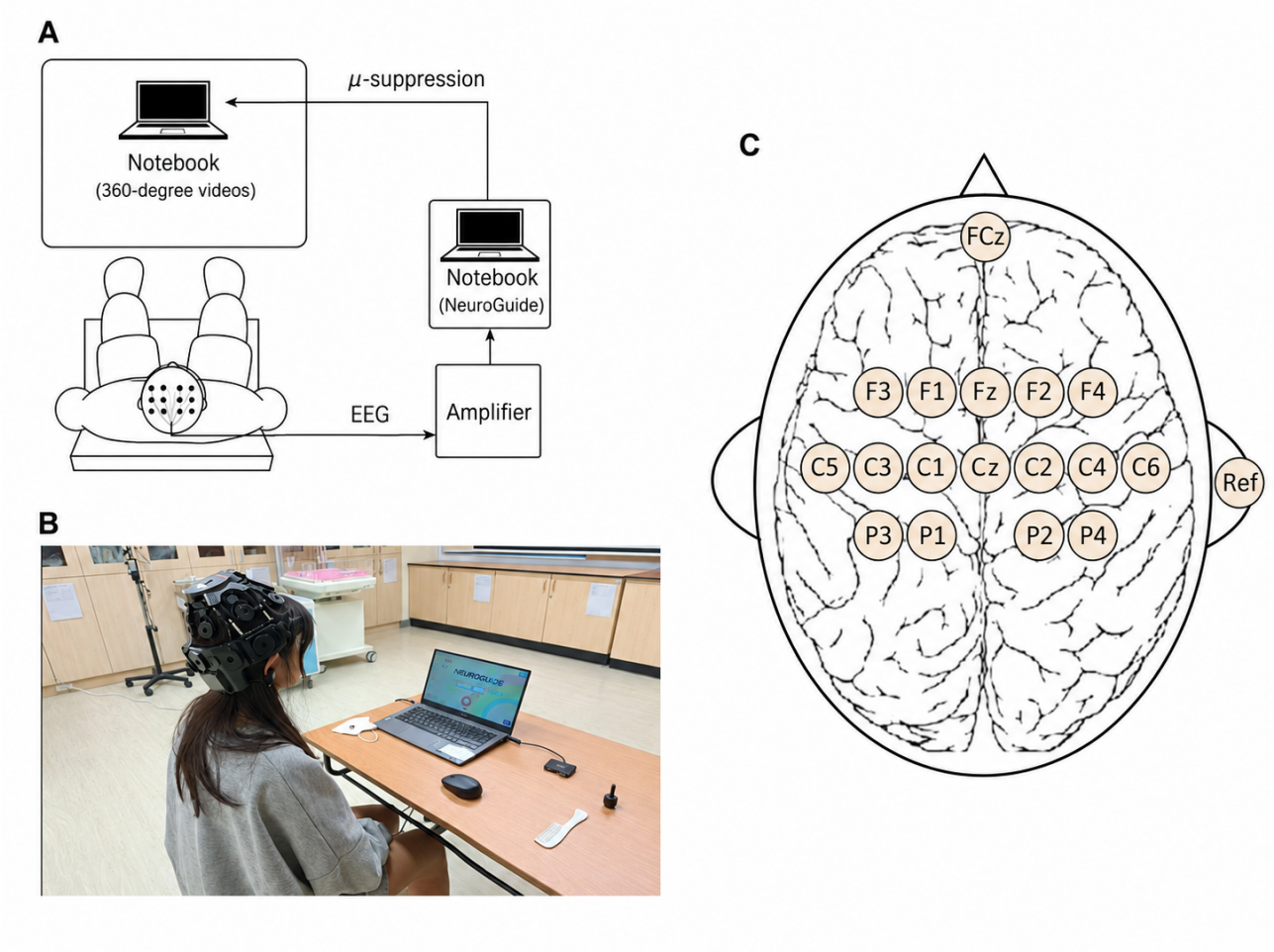
EEG Measures

EEG recordings were obtained in a quiet room with the participants seated comfortably in front of a laptop screen. Electrode impedances were checked before each recording and maintained below the manufacturer-recommended threshold throughout the session. The baseline EEG (T0) data were recorded under standardized resting conditions prior to any module exposure, and the postintervention EEG (T1) data were recorded immediately after the completion of the module under the same recording setup. The same acquisition settings and recording environment were used for both groups to ensure comparability across time points and conditions. After the EEG

XSL•FO
RenderX

recordings were obtained, the EEG data were analyzed by a professional pediatric neurologist based on a standardized protocol. The EEG system consisted of both hardware (DSI-24 EEG headset; Wearable Sensing) and software (NeuroGuide version 3.4.0; Applied Neuroscience, Inc.). The EEG data were collected with the assistance of a Neuroscan SynAmps system with a 0.1 - to 30-Hz bandpass filter and a sampling rate of 500 Hz. For analysis, we extracted the first 160 seconds of eyes-open EEG data from each recording. The participants did not perform an active task during this segment (Figure 1). The μ waves were defined as the oscillations measured in the sensorimotor cortex; therefore, only the data obtained from C3, Cz, and C4 were analyzed. The EEG data were analyzed with the assistance of MATLAB and EEGLAB. The fast Fourier transform was applied with a Hanning window, and the power spectral density was computed for each 2-second epoch [17] (details are provided in Multimedia Appendix 1).

Figure 1. Schematic representation of the experimental setup for electroencephalography (EEG)-based analysis in pediatric nursing training (**A**), a participant using the screen-based 360° contextual simulation video system with EEG recording (**B**), and electrode placement for EEG data acquisition (**C**).



Statistical Analysis

Missing questionnaire data were handled using a prespecified rule. The SDS, EPQ, and SCLS comprise 50 (20+16+14) items in total. Questionnaire records with ≥10 missing items at a given assessment time point were excluded from the questionnaire-based analyses because this level represents ≥20% missingness, a level at which scale scores may be unstable. All questionnaire- and EEG-derived outcomes were analyzed by an analyst who was not involved in intervention delivery or questionnaire administration [18] using IBM SPSS Statistics (version 25). Descriptive statistics are reported as the means (SDs). All statistical tests were 2-tailed, and statistical significance was set at $P<.05$. For the questionnaire outcomes, Pearson correlation coefficients were calculated to examine the associations between the pretest and posttest scores within each group. The analyses of the SDS and EPQ scores were treated as primary, whereas those of the SCLS scores and EEG indices were treated as secondary. Within-group pretest-to-posttest changes were evaluated using 2-tailed paired-samples *t* tests for each questionnaire outcome. Between-group differences in questionnaire outcomes were examined using ANCOVA, with the posttest score as the dependent variable, the group as the fixed factor, and the baseline (pretest) score as the covariate. Assumptions were assessed using Levene tests for homogeneity of variances and by testing the interaction between the group and baseline scores to evaluate the homogeneity of regression slopes. Effect sizes are reported as partial η^2 [19]. For EEG power measures at C3, Cz, and C4, Pearson correlation coefficients were calculated to examine the associations between the pretest and posttest values within each group. Within-group pretest-to-posttest changes in EEG power were evaluated using

XSL•FO
RenderX

2-tailed paired-samples *t* tests at each electrode site. Between-group differences in EEG power were assessed using Levene tests for homogeneity of variances and tests of the group × baseline interaction to evaluate the homogeneity of regression slopes. When the assumption of homogeneity of regression slopes was met, ANCOVA was conducted with the posttest EEG value as the dependent variable, the group as the fixed factor, and the baseline EEG value as the covariate. When the assumption was violated, ANCOVA was not performed for that electrode. Effect sizes are reported as partial η^2 [19]. Between-group differences in the μ-suppression scores at C3, Cz, and C4 were examined using 2-tailed independent-samples *t* tests. Levene tests were used to assess the homogeneity of variances prior to each *t* test.

Ethical Considerations

The study protocol was conducted in accordance with institutional and national research ethics standards and the Declaration of Helsinki and approved by the Institutional Review Board of the MacKay Memorial Hospital, Taiwan (No. 21MMHIS438e). Informed consent was obtained from all participants. The participants were provided with a consent form on day 1 before baseline (T0), and consent was obtained immediately prior to administering the baseline questionnaires and conducting EEG recording. No participants withdrew after providing their consent. Participant privacy and confidentiality were protected throughout the study. All questionnaire and EEG data were coded and deidentified before analysis, and no names, student identification numbers, or other directly identifying information were included in the analytic dataset. Research data were stored in password-protected files and were accessible only to authorized members of the research team. The findings are reported in aggregate form only. Participants each received NT $300 (US $9.5) as compensation for their time. A detailed study protocol is provided in Multimedia Appendix 2.

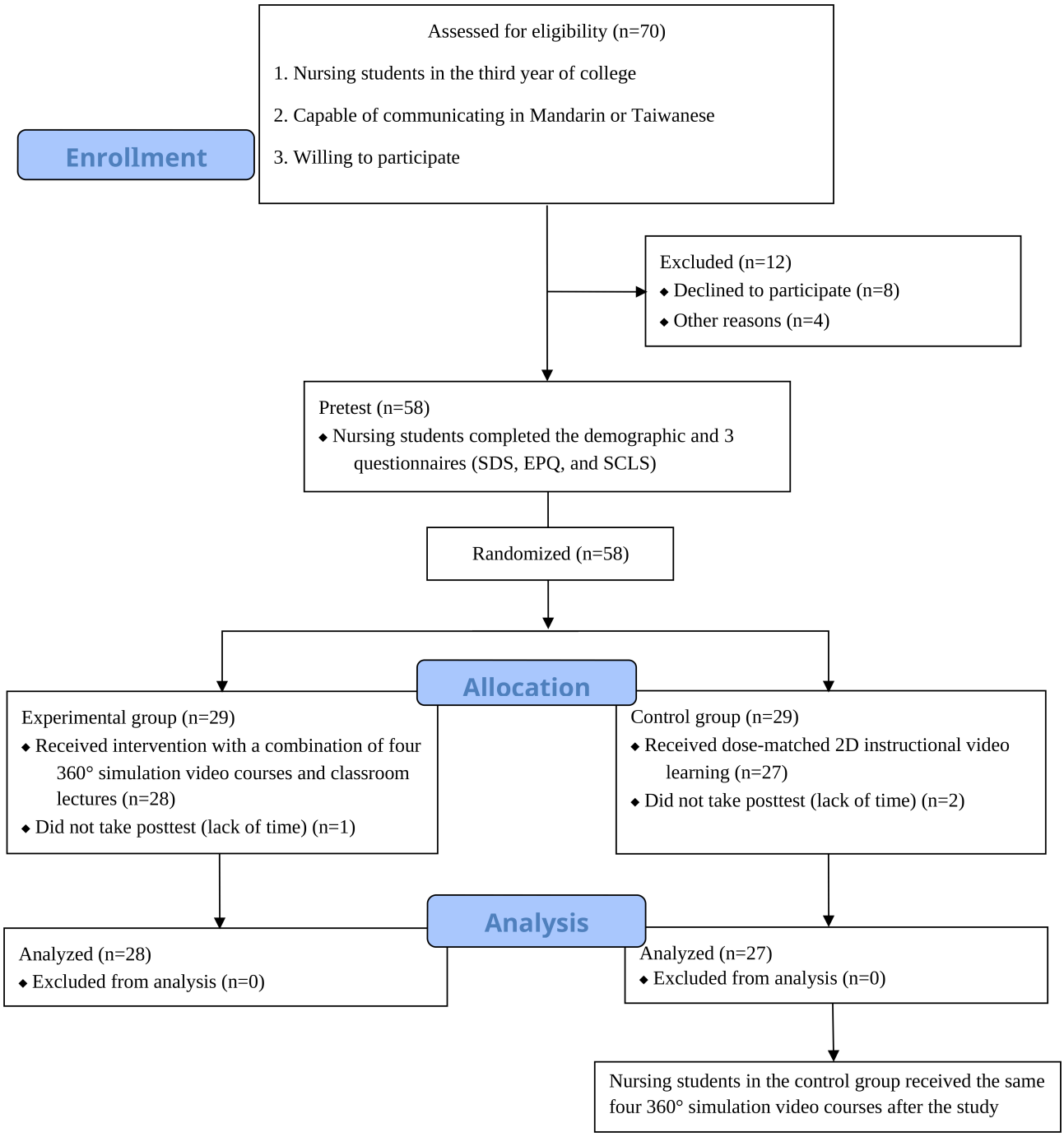
Results

Participants

No questionnaire records were excluded at baseline (T0). At posttest (T1), 3 questionnaire records were excluded because of ≥10 missing items (1 in the experimental group and 2 in the control group). A total of 55 out of 70 students (79%), including 28 students in the experimental group and 27 students in the control group, completed both the pretest and posttest assessments. The CONSORT flow diagram, which outlines the participant recruitment and retention process used in this study, is shown in Figure 2. Baseline demographic characteristics and pretest outcome measures were comparable between the experimental and control groups, as shown in Table 1.

XSL•FO
RenderX

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow diagram of participant recruitment, allocation, intervention, and analysis. EEG: electroencephalography; EPQ: Educational Practices Questionnaire; SCLS: Student Satisfaction and Self-Confidence in Learning Scale; SDS: Simulation Design Scale.



XSL•FO
RenderX

Table . Descriptive characteristics and group comparisons of the SDS[a], EPQ[b], SCLS[c] scores, and EEG[d] power measures.

Characteristic	Experimental group (n=28)	Control group (n=27)	Total (n=55)	Pearson χ^2 (*df*)	*t* test (*df*)	*P* value	95% CI
Gender, n (%)				0.189 (1)	—[e]	.66	—
Male	7 (25)	8 (30)	15 (27)				
Female	21 (75)	19 (70)	40 (73)				
Academic performance (GPA[f]), n (%)				1.413 (3)	—	.70	—
2.6 - 4.0	10 (36)	9 (33)	19 (35)				
1.6 - 2.5	4 (14)	8 (30)	12 (22)				
1.1 - 1.5	9 (32)	5 (19)	14 (25)				
0 - 1.0	5 (18)	5 (19)	10 (18)				
Age (y), mean (SD)	16.85 (0.69)	17.00 (0.83)	16.95 (0.78)	—	0.577 (53)	.57	−0.39 to 0.69
Number of hours spent on the internet weekly, mean (SD)	37.15 (10.14)	37.76 (20.49)	37.56 (17.65)	—	0.125 (53)	.90	−9.20 to 10.41
SDS score, mean (SD)	77.54 (8.66)	75.26 (6.51)	76.00 (7.24)	—	−0.931 (53)	.36	−7.24 to 2.68
EPQ score, mean (SD)	64.46 (8.34)	59.81 (7.66)	61.33 (8.09)	—	−1.756 (53)	.09	−10.03 to 0.74
SCLS score, mean (SD)	61.77 (10.72)	58.04 (6.40)	59.25 (8.11)	—	−1.379 (53)	.18	−9.21 to 1.75
EEG power, mean (SD)							
C3	11.07 (6.9)	13.47 (11.85)	11.87 (8.80)	—	−0.797 (53)	.43	−8.48 to 3.69
Cz	40.60 (133.13)	71.20 (187.27)	50.80 (151.49)	—	−0.590 (53)	.56	−135.77 to 74.57
C4	14.24 (11.88)	15.84 (17.76)	14.78 (13.90)	—	−0.335 (53)	.74	−11.27 to 8.08

[a]SDS: Simulation Design Scale.

[b]EPQ: Educational Practices Questionnaire.

[c]SCLS: Student Satisfaction and Self-Confidence in Learning Scale.

[d]EEG: electroencephalography.

[e]Not applicable.

[f]GPA: grade point average, used as an indicator of academic performance [20].

Effect of Screen-Based 360° Contextual Simulation Videos on Questionnaire Measures

Within-Group Analyses

In the experimental group, the correlation coefficients between the pretest and posttest scores for the SDS, EPQ, and SCLS were 0.258, −0.006, and −0.145, respectively. Two-tailed paired-sample *t* tests revealed statistically significant increases in the SDS (t_{27}=−6.44; *P*<.001) and EPQ (t_{27}=−3.47; *P*=.005) scores, whereas the changes observed in the SCLS scores were not statistically significant (t_{27}=−1.32; *P*=.21; Table 2).

XSL•FO
RenderX

Table . Effects of the screen-based 360° contextual simulation videos on the SDS[a], EPQ[b], SCLS[c] scores, and EEG[d] power measures in the experimental and control groups (n=55).

Variable	Experimental group (n=28)			95% CI	Control group (n=27)			95% CI	ANCOVA term	*P* value$_3$	η^2
	Mean (SD)	*t* test (*df*)	*P* value$_1$		Mean (SD)	*t* test (*df*)	*P* value$_2$				
SDS		–6.44 (27)	<.001[e]	–22.34 to –11.04		–5.50 (26)	.001[e]	–16.38 to –7.47			
Pretest	77.54 (8.66)				75.26 (6.51)				Covariate	.79	.002
Posttest	94.23 (6.41)				87.19 (8.03)				Group	.008[e]	.182
EPQ		–3.47 (27)	.005[e]	–16.90 to –3.87		–1.67 (26)	.11	–9.82 to 1.01			
Pretest	64.46 (8.34)				59.81 (7.66)				Covariate	.047[e]	.108
Posttest	74.85 (6.78)				64.22 (15.15)				Group	.02[e]	.146
SCLS		–1.32 (27)	.21	–11.44 to 2.81		–2.13 (26)	.04[e]	–5.90 to –0.10			
Pretest	61.77 (10.72)				58.04 (6.40)				Covariate	.45	.016
Posttest	66.08 (3.57)				61.04 (6.20)				Group	.01[e]	.167
C3[f]		3.93 (27)	<.001[e]	2.14 to 6.82		1.79 (26)	.10	–0.43 to 4.36			
Pretest	11.07 (6.95)				13.47 (11.85)				Covariate	—[g]	—
Posttest	6.60 (3.58)				11.50 (9.59)				Group	—	—
Cz[f]		5.41 (27)	<.001[e]	4.93 to 10.99		2.05 (26)	.06	–0.31 to 10.16			
Pretest	40.60 (133.13)				71.20 (187.27)				Covariate	<.001[e]	.998
Posttest	32.65 (131.11)				66.28 (183.41)				Group	<.001[e]	.926
C4[f]		4.82 (27)	<.001[e]	3.14 to 7.84		0.847 (26)	.41	–3.99 to 9.07			
Pretest	14.24 (11.88)				15.84 (17.76)				Covariate	<.001[e]	.698
Posttest	8.75 (10.05)				13.39 (12.08)				Group	.03[e]	.138

[a]SDS: Simulation Design Scale.

[b]EPQ: Educational Practices Questionnaire.

[c]SCLS: Student Satisfaction and Self-Confidence in Learning Scale.

[d]EEG: electroencephalography.

[e]$P<.05$. P_1 was calculated by performing 2-tailed paired *t* tests within the experimental group, and P_2 was calculated by performing 2-tailed paired *t* tests within the control group. P_3 and η^2 were calculated by analysis of covariance to compare the posttest scores between the groups, and adjustments were made for baseline differences.

[f]C3, Cz, and C4: EEG electrode placement locations based on the international 10-20 system.

[g]Not applicable.

In the control group, the correlation coefficients between the pretest and posttest scores for the SDS, EPQ, and SCLS were −0.192, 0.435, and 0.323, respectively. Two-tailed paired-samples *t* tests indicated a statistically significant

XSL•FO
RenderX

decrease in the SDS score (t_{26}=−5.50; *P*=.001) and a statistically significant increase in the SCLS score (t_{26}=−2.13; *P*=.04), whereas the change in the EPQ score (t_{26}=−1.67; *P*=.11) was not statistically significant (Table 2).

Between-Group Analyses

The assumptions for ANCOVA were checked (homogeneity of variances and homogeneity of regression slopes) and were met for the SDS, EPQ, and SCLS scores (details are provided in Multimedia Appendix 3). ANCOVA revealed significant between-group effects for the SDS, EPQ, and SCLS scores (Table 2), indicating that exposure to the screen-based 360° contextual simulation videos significantly influenced these questionnaire outcomes. Effect sizes (partial η^2) are reported in Table 2. Taken together, the results indicate that, relative to the dose-matched control condition, the 360° video condition was associated with more favorable posttest student-reported perceptions of simulation design and educational practices (SDS and EPQ) and higher satisfaction or self-confidence (SCLS) after adjusting for baseline scores.

Effects of Screen-Based 360° Contextual Simulation Videos on EEG Measures

Within-Group Analyses

In the experimental group, the correlation coefficients between the pretest and posttest values at electrode sites C3, Cz, and C4 were 0.550, 0.999, and 0.873, respectively. Two-tailed paired-sample *t* tests revealed statistically significant reductions in EEG power frequencies at all 3 electrode sites: C3 (t_{27}=3.93; *P*<.001), Cz (t_{27}=5.41; *P*<.001), and C4 (t_{27}=4.82; *P*<.001). These findings indicate a significant decrease in EEG power frequencies after exposure to the screen-based 360° contextual simulation videos.

In the control group, the mean (SD) pretest and posttest EEG power frequencies at C3, Cz, and C4 were 13.47 (SD 11.85) Hz versus 11.50 (SD 9.59) Hz, 71.20 (SD 187.27) Hz versus 66.28 (SD 183.41) Hz, and 15.84 (SD 17.76) Hz versus 13.39 (SD 12.08) Hz, respectively. The correlation coefficients between the pretest and posttest values at C3, Cz, and C4 were 0.953, 0.999, and 0.803, respectively. Two-tailed paired-sample *t* tests for the control group yielded the following results: C3 (t_{26}=1.79; *P*=.10), Cz (t_{26}=2.05; *P*=.06), and C4 (t_{26}=0.85; *P*=.41; Table 2). None of these *P* values reached statistical significance, thus indicating no significant changes in EEG power frequencies at any of the 3 electrode sites among participants who did not engage with the screen-based 360° contextual simulation videos.

Between-Group Analyses

The assumptions for between-group comparisons were assessed (details are provided in Multimedia Appendix 3). For the EEG power outcomes, the assumption of homogeneity of regression slopes was violated for C3; therefore, ANCOVA was not performed for C3. The ANCOVA results for Cz and C4 indicated significant between-group differences in posttest eyes-open resting-state EEG power after adjusting for baseline values (Table 2). Two-tailed independent-samples *t* tests revealed significant between-group differences in μSC at Cz and C4 (Table 3). In this study, μSC reflects eyes-open resting-state μ-band modulation rather than task-evoked suppression. These results indicate between-group differences in posttest resting-state μ-band indices at selected sites.

Table . Effects of the screen-based 360° contextual simulation videos on the μ-suppression scores between groups (n=55).

Variable μSC[a]	*t* test (df)	*P* value	Mean difference	SE difference	95% CI
C3	1.57 (53)	.13	16.89	10.77	−4.93 to 38.72
Cz	2.29 (53)	.03[b]	22.45	9.80	2.59 to 42.30
C4	2.21 (53)	.03[b]	23.69	10.71	1.20 to 45.38

[a]μSC: μ-band indices derived from eyes-open resting electroencephalography.

[b]*P*<.05. *P* values were calculated by performing 2-tailed independent samples *t* tests to examine the differences between groups.

Discussion

Principal Findings

This study revealed that exposure to screen-based 360° contextual simulation videos was associated with significant changes in both questionnaire outcomes and resting-state EEG indices. In the within-group analyses, the experimental group showed significant pretest-to-posttest improvements in SDS and EPQ scores and significant reductions in EEG power at C3, Cz, and C4, whereas the control group showed limited or nonsignificant changes in most measures. In between-group analyses, ANCOVA indicated that significant group effects for SDS, EPQ, and SCLS and resting-state EEG outcomes also differed between groups at selected electrode sites (Cz and C4), with additional between-group differences observed in the μ-rhythm indices at Cz and C4.

Key Considerations

The findings indicate that the students positively responded to the use of screen-based 360° contextual simulation videos, reporting high levels of satisfaction with instructional design and teaching methodologies. These findings suggest that, compared with the control condition, exposure to screen-based 360° contextual simulation videos was associated with more favorable student-reported perceptions of simulation design and educational practices (SDS and EPQ scores) and higher self-reported satisfaction and self-confidence (SCLS scores). Additionally, the resting-state EEG indices (eyes-open) showed pre-post changes in the experimental condition, and group differences were observed at selected sensorimotor sites (eg, Cz and C4; Table 2). Because EEG data were collected in the eyes-open resting condition rather than during video viewing, these findings should be interpreted as overall pre-post

XSL•FO
RenderX

sensorimotor rhythm modulation following learning exposure, not as direct video format-evoked action-observation responses. From an educational perspective, the use of screen-based 360° contextual simulation videos may represent a scalable approach to support procedural action observation in health professions training [5,6]. However, motor learning and cognitive engagement were not directly measured in this study and should be examined using objective performance outcomes and broader cognitive measures in future work.

EEG analyses indicated significant pre-post modulation of sensorimotor rhythm power at C3, Cz, and C4 in the eyes-open resting condition, with group differences at selected sites (eg, Cz and C4). Because EEG data were recorded in the eyes-open resting state at T0 and T1, these results reflect overall pre-post changes following exposure to the respective learning conditions rather than neural responses elicited during video viewing. Because the EEG analyses in this study focused on central sensorimotor sites (C3, Cz, and C4), we cannot attribute the observed effects to prefrontal regions or hemispheric dominance. In addition, μ-rhythm suppression is an indirect and nonspecific marker that can be influenced by attention, visual processing demands, and cognitive load. Therefore, these findings should be interpreted as EEG correlates consistent with sensorimotor rhythm modulation rather than as direct evidence of mirror neuron activation, cognitive engagement, or motor learning [8,11]. In future studies, EEG data should be recorded during video viewing under both conditions to examine whether the video format is associated with task-evoked sensorimotor EEG differences during procedural observation.

Taken together, the combination of higher posttest questionnaire scores (SDS, EPQ, and SCLS) and distinct sensorimotor EEG modulation (Cz, C4 power changes, and μ-rhythm indices) suggests that the screen-based 360° contextual simulation video condition was associated with more favorable learner perceptions and with pre-post modulation of resting sensorimotor EEG indices at selected sites. However, EEG data have limited spatial resolution, and μ-suppression is nonspecific [8]; due to the design of this study, specific neural generators could not be isolated, and the potential contributions of attention, visual processing, or unmeasured cognitive load could not be disentangled. Objective technical performance and longer-term retention were not assessed in this study and should be considered in future trials. Beyond EEG outcomes, students in the 360° video condition reported more favorable perceptions of simulation design and educational practices (SDS and EPQ scores) and higher satisfaction or self-confidence (SCLS scores) than did students in the control condition. These self-reported findings support the acceptability of screen-based 360° contextual simulation video learning as a supplement to undergraduate nursing skills education.

In interpreting these findings, Kolb [12] experiential learning cycle provides a useful lens, but the present intervention should be viewed as partially aligned with the model. The screen-based 360° contextual simulation videos primarily supported concrete experience (exposure to realistic procedural scenarios) and reflective observation (facilitated debriefing), and the linked procedural checklists may have supported abstract conceptualization. Consistent with evidence that observing clinical procedures can support skill retention [21], the intervention may help scaffold the early phases of learning. However, active experimentation (hands-on practice with feedback) was not implemented or objectively measured within the study sessions; thus, a complete experiential learning cycle cannot be claimed. In future work, hands-on performance should be integrated and assessed to operationalize active experimentation and retention more directly [22].

Research Implications and Future Directions

From an educational perspective, screen-based 360° contextual simulation videos may offer a scalable supplement for undergraduate nursing skills education, particularly when supervised practice opportunities are constrained. In this study, the 360° video condition was associated with more favorable student-reported perceptions of simulation design and educational practices (SDS and EPQ scores) and higher satisfaction or self-confidence (SCLS scores), along with pre-post modulation of resting sensorimotor EEG indices at selected sites. These EEG findings should be interpreted as indirect, nonspecific neural correlates that are consistent with sensorimotor rhythm modulation rather than as direct evidence of mirror neuron activation, cognitive engagement, motor learning, or clinical competence [8,11]. In future studies, objective performance outcomes (eg, skills checklists), longer-term follow-up, and direct measures of cognitive load should be evaluated, and blended designs that integrate video-based observation with structured hands-on practice and feedback should be incorporated [23].

Limitations

This study was conducted within an academic setting, which may have introduced an inherent power dynamic between the researchers and the participants. The presence of course instructors during data collection could have unintentionally influenced the students' responses to the questionnaires, thereby introducing potential response bias and affecting the reliability of the self-reported measures. Additionally, the relatively small sample size and single-institution design may have limited the statistical power and generalizability of the findings, particularly in the context of diverse higher education environments. Because EEG data were collected in the eyes-open resting condition at T0 and T1 (and not during video viewing), the study could not test whether different video formats directly elicit different action-observation EEG responses. Because all participants completed the 4 modules in a fixed order and EEG data were collected only at baseline and immediately after the final (fourth) module, the EEG findings reflect an overall pre-post change across the full sequence rather than scenario-specific effects. Potential cumulative effects (eg, learning, fatigue, or changes in cognitive load) across the 4 modules cannot be disentangled. In addition, cognitive load was not directly measured, which limits the interpretability of the EEG indices. Although standardized proctoring procedures were used and confidentiality was emphasized, the presence of faculty members and teaching assistants during questionnaire completion may still have introduced social desirability or response bias, which may have affected the validity of the self-reported outcomes. Excluding questionnaire records with substantial missingness

XSL•FO
RenderX

may have introduced response bias if missingness was not random (eg, systematic skipping among less engaged students). Although the number of excluded posttest records was small (3 records at T1), the questionnaire findings should be interpreted with caution in light of potential nonrandom missingness.

Conclusions

This study suggests that integrating screen-based 360° contextual simulation videos into undergraduate nursing education is associated with improved learner-reported outcomes (SDS, EPQ, and SCLS scores) and modulation of eyes-open resting-state sensorimotor EEG indices. Future research should include analyses of objective performance outcomes and task-evoked EEG during video viewing to clarify relevant format-specific mechanisms.

Acknowledgments

The authors sincerely thank all the participants for contributing their time and effort to this study. During manuscript revision and copyediting, the authors made limited use of ChatGPT by OpenAI for minor editorial assistance, including formatting responses to copyeditor queries, and checking consistency with journal style requirements. The tool was not used to generate study data, conduct statistical analyses, interpret results, or create references. All artificial intelligence–assisted content was reviewed, revised, and verified by the authors, who take full responsibility for the accuracy, originality, and integrity of the manuscript.

Funding

This work was supported by the Ministry of Science and Technology, Taiwan (MOST 111 - 2635-H-715-001; NSTC 112 - 2410-H-715-005). The funder had no role in the study design; data collection, analysis, or interpretation; manuscript preparation; or the decision to submit the work for publication. This study was registered at the Chinese Clinical Trial Registry (ChiCTR) (number ChiCTR2500102861; registered May 21, 2025).

Data Availability

Owing to institutional data-sharing policies and participant confidentiality agreements, the datasets generated and analyzed during the study are not publicly available but are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: YCC, CJC
Data curation: YCC
Formal analysis: YCC
Funding acquisition: CJC
Investigation: YCC
Methodology: YCC, CJC
Project administration: YCC
Resources: CJC
Supervision: CJC
Validation: CJC
Visualization: YCC
Writing – original draft: YCC
Writing – review & editing: YCC, CJC

Multimedia Appendix 1
Electroencephalography (EEG) data.
[DOCX File, 46 KB - nursing_v9i1e84720_app1.docx]

Multimedia Appendix 2
Study protocol.
[DOCX File, 51 KB - nursing_v9i1e84720_app2.docx]

Multimedia Appendix 3
Detailed assumption checks.
[DOCX File, 47 KB - nursing_v9i1e84720_app3.docx]

XSL•FO
RenderX

Checklist 1
CONSORT checklist.
[PDF File, 223 KB - nursing_v9i1e84720_app4.pdf]

References

1. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011 Sep 7;306(9):978-988. [doi: 10.1001/jama.2011.1234] [Medline: 21900138]
2. Koukourikos K, Tsaloglidou A, Kourkouta L, et al. Simulation in clinical nursing education. Acta Inform Med 2021 Mar;29(1):15-20. [doi: 10.5455/aim.2021.29.15-20] [Medline: 34012208]
3. Mondal H, Mondal S. Adopting augmented reality and virtual reality in medical education in resource-limited settings: constraints and the way forward. Adv Physiol Educ 2025 Jun 1;49(2):503-507. [doi: 10.1152/advan.00027.2025] [Medline: 40136005]
4. Görücü S, Türk G, Karaçam Z. The effect of simulation-based learning on nursing students' clinical decision-making skills: systematic review and meta-analysis. Nurse Educ Today 2024 Sep;140:106270. [doi: 10.1016/j.nedt.2024.106270] [Medline: 38924975]
5. Rosendahl P, Wagner I. 360° videos in education—a systematic literature review on application areas and future potentials. Educ Inf Technol 2024 Feb;29(2):1319-1355. [doi: 10.1007/s10639-022-11549-9]
6. Blair C, Walsh C, Best P. Immersive 360° videos in health and social care education: a scoping review. BMC Med Educ 2021 Nov 24;21(1):590. [doi: 10.1186/s12909-021-03013-y] [Medline: 34819063]
7. Hsieh JY, Lin PC, Sun WN, Lin TR, Kuo CC, Hsu HT. Effectiveness of immersive virtual reality in nursing education for nursing students and nursing staffs: a systematic review and meta-analysis. Nurse Educ Today 2025 Aug;151:106725. [doi: 10.1016/j.nedt.2025.106725] [Medline: 40215711]
8. Muthukumaraswamy SD, Johnson BW. Primary motor cortex activation during action observation revealed by wavelet analysis of the EEG. Clin Neurophysiol 2004 Aug;115(8):1760-1766. [doi: 10.1016/j.clinph.2004.03.004] [Medline: 15261854]
9. Jallad ST, Işık B. The effectiveness of virtual reality simulation as learning strategy in the acquisition of medical skills in nursing education: a systematic review. Ir J Med Sci 2022 Jun;191(3):1407-1426. [doi: 10.1007/s11845-021-02695-z] [Medline: 34227032]
10. Buccino G, Binkofski F, Riggio L. The mirror neuron system and action recognition. Brain Lang 2004 May;89(2):370-376. [doi: 10.1016/S0093-934X(03)00356-0] [Medline: 15068920]
11. Lee JH, Lee SE, Kwon YS. Exploring empathic engagement in immersive media: an EEG study on mu rhythm suppression in VR. PLoS One 2024;19(5):e0303553. [doi: 10.1371/journal.pone.0303553] [Medline: 38758939]
12. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development, 2nd edition: Pearson FT Press; 2014.
13. Sweller J. The development of cognitive load theory: replication crises and incorporation of other theories can lead to theory expansion. Educ Psychol Rev 2023;35(4):95. [doi: 10.1007/s10648-023-09817-2]
14. Mohammadi M, Tajik E, Martinez-Maldonado R, Sadiq S, Tomaszewski W, Khosravi H. Artificial intelligence in multimodal learning analytics: a systematic literature review. Comput Educ Artif Intell 2025 Jun;8:100426. [doi: 10.1016/j.caeai.2025.100426]
15. Use of NLN surveys and research instruments. National League for Nursing. 2021. URL: https://www.nln.org/professional-development-programs/research/tools-and-instruments [accessed 2026-01-18]
16. Franklin AE, Burns P, Lee CS. Psychometric testing on the NLN student satisfaction and self-confidence in learning, simulation design scale, and educational practices questionnaire using a sample of pre-licensure novice nurses. Nurse Educ Today 2014 Oct;34(10):1298-1304. [doi: 10.1016/j.nedt.2014.06.011] [Medline: 25066650]
17. Rader C, Brenner N. A new principle for fast Fourier transformation. IEEE Trans Acoust Speech Signal Process 1976 Jun;24(3):264-266. [doi: 10.1109/TASSP.1976.1162805]
18. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research: SAGE Publications; 2017.
19. Cohen J. Statistical Power Analysis for the Behavioral Sciences: Academic Press; 1988.
20. Chen CJ, Sung HC, Lee MS, Chang CY. The effects of Chinese five-element music therapy on nursing students with depressed mood. Int J Nurs Pract 2015 Apr;21(2):192-199. [doi: 10.1111/ijn.12236] [Medline: 24593291]
21. Hober C, Bonnel W. Student perceptions of the observer role in high-fidelity simulation. Clin Simul Nurs 2014 Oct;10(10):507-514. [doi: 10.1016/j.ecns.2014.07.008]
22. Thidemann IJ, Söderhamn O. High-fidelity simulation among bachelor students in simulation groups and use of different roles. Nurse Educ Today 2013 Dec;33(12):1599-1604. [doi: 10.1016/j.nedt.2012.12.004] [Medline: 23302256]
23. Padilha JM, Machado PP, Ribeiro A, Ramos J, Costa P. Clinical virtual simulation in nursing education: randomized controlled trial. J Med Internet Res 2019 Mar 18;21(3):e11529. [doi: 10.2196/11529] [Medline: 30882355]

XSL•FO
RenderX

Abbreviations

AGFI: adjusted goodness-of-fit index
ANCOVA: analysis of covariance
CFI: comparative fit index
CONSORT: Consolidated Standards of Reporting Trials
EEG: electroencephalography
EPQ: Educational Practices Questionnaire
NFI: normed fit index
RMSEA: root mean square error of approximation
SCLS: Student Satisfaction and Self-Confidence in Learning Scale
SDS: Simulation Design Scale
TLI: Tucker-Lewis index
WRMR: weighted root mean square residual

Edited by E Borycki; submitted 24.Sep.2025; peer-reviewed by F Alshehri, M Luctkar-Flude; revised version received 03.Mar.2026; accepted 07.Apr.2026; published 30.Jun.2026.

Please cite as:
Chen YC, Chen CJ
360° Contextual Simulation Videos for Undergraduate Nursing Students: Electroencephalography-Based Quasi-Experimental Study
JMIR Nursing 2026;9:e84720
URL: https://nursing.jmir.org/2026/1/e84720
doi:10.2196/84720

© Yi-Chang Chen, Chen-Jung Chen. Originally published in JMIR Nursing (https://nursing.jmir.org), 30.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Associations of Emotional Intelligence, AI Self-Efficacy, and AI Literacy Among Nursing Undergraduates Under the NUR.S.E.S. Framework: Network Analysis

XiaoHui Fan[1*], MD; Jingjing Sun[2*], MD; LingHui Zhang[2], MD; Jinrong Wang, PhD, PH; BinHao Dong[2], MD; XiaoHong Gao[2], MD; Ruihua Jin[2], MD; Panpan Huai[2], MD, PhD

[1]Biomedical Research Center of Medical college, Xijing University, Xi'An, Shaanxi Province, China

[2]School of Nursing, Shanxi Medical University, 56 Xinjiannan Road, Yingze District, Taiyuan City, Shanxi Province, Taiyuan, Shanxi, China

[*]these authors contributed equally

Corresponding Author:
Panpan Huai, MD, PhD
School of Nursing, Shanxi Medical University, 56 Xinjiannan Road, Yingze District, Taiyuan City, Shanxi Province, Taiyuan, Shanxi, China

Abstract

Background: With the rapid development of generative artificial intelligence (AI) and its deep integration into nursing education, nursing students' AI literacy (AS) has become a critical competency for their professional development. However, the patterns of associations among emotional intelligence (EI), AI self-efficacy (AILS), and AS in relation to comprehensive AS remain unclear.

Objective: Based on the NUR.S.E.S. framework and using network analysis methods, this study systematically mapped the complex relational network among EI, AILS, and AS among undergraduate nursing students. It identified nodes with high centrality and bridging strength within this network, offering preliminary insights that may inform future educational interventions.

Methods: A cross-sectional survey design was used, with 982 undergraduate nursing students from a university conveniently sampled in September 2025 as research participants. Assessments were conducted using the EI Scale, the AILS Scale, and the AS Scale. Using R (version 4.5.1; R Core Team), we constructed a Gaussian graph model, calculated centrality metrics such as node and bridge strength, and assessed network stability using the bootstrap method.

Results: Network analysis showed that emotion regulation (strength centrality=1.355) and evaluative ability (strength centrality=1.323) showed the highest strength centrality, indicating their prominent positions within the network. Emotional perception (bridge strength=0.427) and comfort with AI (bridge strength=0.242) are the most critical bridge nodes, appearing to connect EI with AI technology systems. Simultaneously, the network architecture suggests that AILS may play a bridging role, effectively linking EI (particularly emotional perception as a bridging factor) with higher levels of AS.

Conclusions: Cultivating AS among undergraduate nursing students is a system that deeply integrates emotional, cognitive, and technical confidence. EI was closely associated with AS, and AILS appeared to occupy a bridging position in the network. Educational interventions might consider enhancing emotional perception and comfort with AI, pending validation through longitudinal or experimental designs.

(JMIR Nursing 2026;9:e90253) doi:10.2196/90253

KEYWORDS

artificial intelligence literacy; emotional intelligence; AI self-efficacy; network analysis; nursing education

Introduction

The rapid advancement of digital health care technologies, particularly the rise of generative artificial intelligence (GenAI) represented by large language models, is profoundly reshaping the practice paradigms and knowledge boundaries within health care [1]. For nursing students who will become the backbone of the future health care system, this represents both an unprecedented opportunity and a formidable challenge to their capabilities [2]. GenAI can be broadly defined as a subset of artificial intelligence (AI) that focuses on creating new, original, human-like content based on the data it is trained on [3]. A 2024 survey by the Digital Education Council reveals that as many as 85% of students frequently use GenAI in their academic work, with ChatGPT (OpenAI) being the most widely adopted platform among them [4]. Another global survey conducted in 2025 across 109 countries, which collected 20,000 responses, similarly indicated that students are extensively using GenAI

XSL•FO
RenderX

tools, with ChatGPT usage exceeding 70% [5]. In the face of the widespread adoption of GenAI among students, education must go beyond introducing technological tools. More crucially, we must ensure that future nursing professionals can critically understand, prudently apply, and rationally evaluate AI, thereby transforming technology into a genuine professional enabler [6]. This lays the foundation for nursing students to apply AI tools and technologies appropriately and proficiently.

The American Nurses Association (ANA) asserts that, against the backdrop of emerging AI technologies, nurses applying AI in health care practice require not only ethical guidance but also uphold the core values of care and compassion inherent to nursing [7]. AI literacy (AS) serves as the core competency and vital bridge for integrating ethical guidance and humanistic care into AI application practices. It is regarded as a key competency in nursing education [8]. These capabilities extend beyond merely mastering and applying AI technologies (such as clinical decision support systems and predictive analytics) to encompass the higher-order thinking skills required to independently critically evaluate and assess the outcomes they generate [9]. As the backbone of the future health care system, nursing students are key participants in shaping AI applications in nursing. Their level of AS directly determines whether they can lead the intelligent transformation in this field, thereby profoundly influencing the breadth and depth of their personal career development [10]. Therefore, cultivating AS among nursing students is of paramount importance [11].

However, the cultivation of AS is not an isolated process; it is deeply rooted in the complex psychological, cognitive, and emotional capacity system of individuals [12]. Research indicates that an increasing number of people are turning to GenAI for guidance on interpersonal relationships, self-awareness, and emotional regulation [13]. Emotional intelligence (EI) refers to the ability to recognize, understand, and manage one's own emotions, while also understanding and empathizing with the emotions of others [14]. Higher EI levels are significantly associated with nursing students' stress management, clinical decision-making, and subjective well-being [15]. Research indicates that the use of GenAI can effectively alleviate stress and improve mental health outcomes, while simultaneously enhancing engagement and satisfaction in learning environments [16]. Specifically in the field of nursing education, Ching and Ho [17] found that including AI virtual humans in gamified teaching methods can enhance nursing students' EI. Derakhshan [18] indicates that AI-assisted teaching can enhance students' emotional engagement and goal orientation. Deng and Chen [19] also found that AI-driven environments, by providing personalized feedback and adaptive pathways, enhance their emotional resilience by reducing frustration and boosting students' confidence. These findings collectively reveal that EI not only provides the psychological foundation for students to engage in effective and healthy interactions with AI, but it may also profoundly influence the development and enhancement of their AS by affecting their confidence and initiative when confronting new technologies.

EI provides the intrinsic emotional foundation for nursing students to effectively use AI, while confidence in their mastery of AI technology relies on AI self-efficacy (AILS). AILS is defined as an individual's belief in their ability to perform specific tasks, which directly influences nurses' willingness to adopt AI and their actual use of it [20]. Unlike general self-efficacy, AILS in nursing AI applications manifests as nursing students' confidence in mastering AI tools, overcoming technical challenges, and using them to solve clinical or learning problems [21]. This confidence in specific technical domains serves as the intrinsic motivation driving health care professionals to actively engage with, deeply explore, and persistently use AI tools [22]. Research indicates that enhancing nursing students' self-efficacy is crucial for alleviating psychological distress, ultimately improving their professional well-being and the quality of their future care [23]. In AI-enabled learning environments, Wang et al [24] indicate that self-efficacy and university support significantly enhance students' acceptance of AI-driven blended learning. Gao et al [25] further confirm that integrating AI chatbots into writing instruction can enhance students' AI self-efficacy. Lyu and Salam [26] also indicate that AI-driven personalized learning can enhance students' self-efficacy, motivation, and digital literacy. Additionally, high levels of AILS significantly predict nursing students' willingness to adopt AI technologies and the depth of their usage, thereby enhancing professional growth and adaptability in the rapidly evolving health care field [8].

Thus, EI and AILS jointly constitute the 2 core psychological factors influencing the development of AS among nursing students. EI ensures nursing students maintain essential emotional insight and empathy when interacting with AI and applying it in patient care, thereby providing the necessary human touch to technological implementation. Meanwhile, AILS provides the psychological drive to explore and use AI technologies, motivating students to transform potential technological advantages into realistic nursing capabilities [17]. To systematically guide and cultivate this complex psychological capability system, this study introduces an emerging theoretical framework—the NUR.S.E.S. framework [17]. This framework, recently proposed by Hoelscher and Pugh [27], aims to establish a structured pathway for developing AS within the nursing profession. Compared to generic technology adoption models or general educational technology frameworks, its novelty lies in deeply integrating the development of AS with the humanistic care, ethical responsibility, and clinical leadership inherent to the nursing profession, thereby better aligning with the contextual needs of nursing education. It is an acronym representing 6 core components:

- Navigate AI basics: promote understanding of AI fundamentals as a leadership responsibility critical to patient safety.
- Use AI strategically: purposefully apply AI tools to enhance care quality, safety, and outcomes.
- Recognize AI pitfalls and guide balanced approaches to innovation and risks such as bias, overreliance, and inaccuracy.
- Skills support: advocating for skill enhancement and professional development to prepare for AI integration in nursing.
- Ethics in action: guiding ethical AI use in nursing by promoting transparency, fairness, and accountability.

XSL•FO
RenderX

- Shape the future: directing AI implementation toward outcomes that align with nursing values and advance care equity.

This framework provides nursing educators with a forward-looking theoretical foundation rooted in nursing values, guiding students to develop their humanistic care, critical thinking, and ethical judgment capabilities simultaneously while mastering technical skills. This study used the NUR.S.E.S. framework, marking its first application in nursing education both domestically and internationally. Its internal structure clearly corresponds to core constructs such as EI (corresponding to the E and S dimensions), AILS (corresponding to the S dimension), and AS (corresponding to the N, U, R, and E dimensions), providing a structurally coherent and contextually relevant theoretical foundation for systematically analyzing the complex system formed by these variables.

As a cutting-edge, data-driven approach, network analysis transcends traditional linear assumptions to reveal intricate interactions among variables and precisely pinpoint their relative importance within a system [21]. This method provides a unique, systematic perspective for visualizing and quantifying the dynamic relational structure among EI, AILS, and AS [28]. By analyzing the interrelationships among variables, network analysis not only maps each measurement indicator to the theoretical dimensions of the NUR.S.E.S. framework but also tests whether the framework's predefined dimensional divisions align with the actual patterns of variable associations observed in the data. Specifically, it examines whether variables within the same framework dimension are closely interconnected in the relational network and whether there are significant, unexpected associations between different dimensions, thereby providing empirical support for the framework's validity [29].

Therefore, the objective of this study is to systematically analyze the complex network relationships and underlying mechanisms among EI, AILS, and AS based on the NUR.S.E.S. framework. By constructing a psychological network model, this study aims to identify factors that are associated with nursing students' AS, thereby providing precise targets for educational interventions. We plan to conduct a survey of nursing students at a Chinese university to gain an in-depth understanding of their practical application of AI, including EI, AILS, and AS. By analyzing these findings, we aim to identify existing challenges and provide crucial theoretical foundations and practical guidance for cultivating future nursing professionals who can master intelligent technologies while upholding humanistic care.

Methods

Objects and Methods

A total of 982 full-time undergraduate nursing students from a particular university were selected as research participants. Among them, 117 were male, and 865 were female, with ages ranging from 18 to 24 years (mean age 21.18, SD 0.51 years). The inclusion criteria were (1) full-time undergraduate nursing students enrolled at the university; (2) students who had encountered or used AI-related technologies or tools in their daily studies or lives; and (3) voluntary participation in the experiment with signed informed consent and confidentiality agreements. Exclusion criteria applied to students on formal leave or extended sick leave during data collection. Participants could withdraw freely without affecting their studies or academic performance.

Ethical Considerations

The research procedures have been approved by the Ethics Review Committee of Xijing University (Approval number: XJU-202502). Participants' privacy has been protected. All participants provided written informed consent. No harm was inflicted upon participants. All nursing students had the right to choose whether to participate in this study. Data were securely stored and encrypted.

Instruments

General Demographic Questionnaire

The research team independently developed a questionnaire to collect basic information. This survey aims to systematically examine key social-demographic factors that may influence research outcomes, primarily covering age, gender, home address, prior experience as a student leader, academic year, and average frequency of use of electronic devices and AI tools.

EI Scale

The EI Scale developed by Schutte et al [30] was used to assess students' EI levels. The Chinese version was revised by the Chinese psychologist Caikang [31]. This version has demonstrated good reliability and validity among Chinese students. This scale comprises 4 dimensions, including emotional perception (12 items), self-emotional management (8 items), managing others' emotions (6 items), and emotional application (7 items), totaling 33 items. A 5-point Likert scale was used to score, ranging from "Strongly Disagree" to "Strongly Agree," with values from 1 to 5. Higher scores indicate a higher level of EI among students. The Cronbach α coefficient for the full scale was 0.90, and the test-retest reliability Cronbach α coefficient was 0.78. In this study, the Cronbach α coefficient for the scale was 0.959.

AILS Scale

The AILS scale, developed by Wang et al [32], is used to assess users' self-efficacy when using AI technologies and products. The scale comprises 4 dimensions, including assistance (7 items), anthropomorphic interaction (5 items), comfort with AI (6 items), and technical skills (4 items), totaling 22 items. Scoring uses a 7-point Likert scale from "Strongly Disagree" to "Strongly Agree," with values ranging from 1 to 7. Higher scores indicate greater levels of AILS among users. Confirmatory factor analysis resulted in $\chi^2_{196}=1.984$; comparative fit index (CFI)=0.941; Tucker-Lewis index (TLI)=0.930; root-mean-square error of approximation (RMSEA)=0.079; and standardized root-mean-square residual (SRMR)=0.071. Cronbach α coefficient in this study was 0.953.

AS Scale

The AS Scale was developed by Wang and Chuang [33]. The scale comprises 4 dimensions, including awareness (3 items), usage (3 items), evaluation (3 items), and ethics (3 items). It is

XSL•FO
RenderX

based on a 7-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree," with values ranging from 1 to 7, respectively. Higher scores indicate greater AS among individuals. Confirmatory factor analysis yielded the following indices: CFI=0.99, TLI=0.99, GFI=0.98, RMSEA=0.01, SRMR=0.03. The Cronbach α coefficient for the full scale was 0.830, while the Cronbach α coefficient in this study was 0.864.

Correspondence Table Between the NUR.S.E.S. Framework and Questionnaire Measurement Dimensions

Guided by the theoretical framework of NUR.S.E.S., this study establishes conceptual mappings between the 6 core competency dimensions of the framework (navigate, use, recognize, skills support, ethics in action, and shape the future) and their corresponding measurement variables. This mapping aims to structure the AS of nursing students as defined by the framework into measurable psychological and competency indicators, thereby providing a theoretical basis for subsequent network analysis and testing the intrinsic interrelationships among the framework's components.

Theoretical Rationale for the Mapping Between NUR.S.E.S. Components and Measured Variables

Overview

The mapping shown in Table 1 follows from the conceptual definitions of each NUR.S.E.S. component as originally articulated by Hoelscher and Pugh [27], combined with established theoretical and empirical literature in nursing education, EI, and self-efficacy research.

XSL•FO
RenderX

Table . Correspondence table between the NUR.S.E.S. framework and questionnaire measurement dimensions.

Framework letters	Representative of core variables	Dimensions corresponding to the questionnaire	Dimensional implications	Mapping dimensions
N	Navigate AI[a] basics	Awareness	• Basic understanding of fundamental AI concepts, application scenarios, and its impact on the nursing field.	"Navigate AI basics" emphasizes initial exploration and foundational understanding of the field of AI. Only by acquiring this foundational understanding can students freely explore the AI knowledge system; therefore, the "awareness" dimension has been selected to measure this ability.
U	Use AI strategically	Use	• The ability to operate and apply specific AI tools in practical nursing education, research, or simulated scenarios.	"Use AI strategically" focuses on practical application. This dimension directly assesses students' ability to apply AI as a tool in specific scenarios, reflecting the practical aspect of "use."
R	Recognize AI pitfalls	Evaluation	• Critically examine and evaluate the accuracy and reliability of AI-generated information.	"Recognizing AI pitfalls" requires critical thinking to identify the limitations and risks of AI. The "evaluation" dimension is precisely the core competency involved in assessing and discerning information, and it aligns closely with the requirement to identify flaws.
S	Skills support	Technical Skills Assistance	• Technical skills: confidence in one's ability to learn and master AI technology itself. • Assistance: believe that AI can serve as an effective tool to assist in completing learning or work tasks.	• "Skills support" comprises 2 aspects: first, confidence in mastering AI skills themselves (technical skills); and second, trust in AI as a supportive tool (assistance). These 2 dimensions together underpin students' self-efficacy at the skills level.

XSL•FO
RenderX

Framework letters	Representative of core variables	Dimensions corresponding to the questionnaire	Dimensional implications	Mapping dimensions
E	Ethics in action Emotional Intelligence (EI)	Ethics, Emotional Perception, Self-emotion Management, Managing Others' Emotions, Emotional Application	• Ethics: understand and adhere to ethical standards such as privacy and fairness involved in the use of AI in nursing practice. • Emotional perception: accurately recognize your own and others' emotional states. • Self-emotion management: the ability to regulate and control one's own emotions. • Managing others' emotions: the ability to influence and regulate others' emotions. • Emotional application: use emotional information to aid in thinking and problem-solving.	"Ethics in action" emphasizes the application of ethics in real-world nursing settings. EI as a whole construct is a foundational capacity for ethical practice, including perceiving, managing, and utilizing emotions. The 4 subdimensions of EI (EI1-EI4) are collectively used to measure EI, not to be split into other letters of the NURSES framework. The network analysis treats these 4 subdimensions as nodes belonging to the same EI community, which supports their theoretical coherence. Therefore, mapping all 4 EI subdimensions to the E dimension reflects the role of EI as a holistic supporter of ethics in action, not a forced segmentation.
S	Shape the future	Comfort with AI, Anthropomorphic Interaction	• Comfort with AI: feeling relaxed and at ease when interacting with AI, without anxiety or a sense of threat. • Anthropomorphic interaction: tends to engage in natural communication with AI in a manner similar to human interaction.	"Shape the future" embodies students' open and proactive approach to leading the development of AI in the field of nursing. Comfort with AI reflects their level of acceptance, while anthropomorphic interactions demonstrate their willingness to collaborate with machines; together, these 2 factors form the psychological foundation for actively shaping the future of nursing.

[a]AI: artificial intelligence.

Navigate AI Basics (N)

Hoelscher and Pugh [27] define this component as promoting understanding of AI fundamentals as a leadership responsibility critical to patient safety. In operational terms, foundational knowledge of AI concepts and applications constitutes the prerequisite for any further engagement with AI technologies [34]. Accordingly, the awareness dimension of the AS Scale [33], which assesses students‘ understanding of basic AI concepts, application scenarios, and potential impacts, serves as a direct measure of this component.

Use AI Strategically (U)

This component concerns the purposeful application of AI tools to enhance care quality, safety, and outcomes [27]. Strategic use implies not only familiarity with AI but the ability to operate and apply specific AI tools in practical nursing education, research, or clinical scenarios. The Use dimension of the AS Scale [33], which captures students' self-reported ability to work with AI tools in task-specific contexts, directly corresponds to this component.

Recognize AI Pitfalls (R)

Hoelscher and Pugh [27] emphasize that balanced approaches to innovation must include recognition of risks such as bias, overreliance, and inaccuracy. This critical evaluative competency is captured by the evaluation dimension of the AS Scale [33], which assesses students' ability to critically examine and judge the accuracy and reliability of AI-generated information.

Skills Support (First S)

The skills support component advocates for skill enhancement and professional development to prepare for AI integration in nursing [27]. This component encompasses both the confidence in one's ability to learn and master AI technology itself and the belief that AI can serve as an effective tool to assist in completing learning or work tasks. The AILS Scale [32] operationalizes these 2 facets through its technical skills and assistance dimensions, respectively, drawing on Bandura's broader theory of self-efficacy as domain-specific confidence

XSL•FO
RenderX

in one's ability to execute courses of action required to achieve desired outcomes [20].

Ethics in Action (E)

This component guides ethical AI use in nursing by promoting transparency, fairness, and accountability [27]. Importantly, ethical action in nursing requires not only knowledge of ethical principles but also the emotional capacities that underpin moral sensitivity and compassionate care. A substantial body of nursing literature demonstrates that EI—the ability to perceive, understand, manage, and use emotions—is positively associated with ethical sensitivity and moral reasoning in both nursing students and practicing nurses [8,15,31,35]. For example, Liu et al [36] found that EI mediates the relationship between empathy and moral sensitivity in Chinese student nurses, suggesting that emotional competence is not separate from but foundational to ethical practice [1]. The Schutte EI Scale [30], which measures 4 dimensions of EI (perception of emotion, self-emotion management, managing others' emotions, and emotional application), was therefore selected to capture the emotional competence component of ethical action. These 4 subdimensions are treated as a single conceptual community in the network analysis, consistent with the holistic contribution of emotional competence to ethical practice.

Shape the Future (Second S)

The final component directs AI implementation toward outcomes that align with nursing values and advance care equity [27]. Proactive, leadership-oriented engagement with AI presupposes that nurses feel psychologically comfortable with AI technologies and are willing to engage with them in a natural, quasisocial manner. Positive affective attitudes toward AI, including low anxiety and high comfort, have been identified as psychological prerequisites for adoption and sustained use in health care and educational contexts [21,37,38]. The AILS Scale captures these prerequisites through its comfort with AI and anthropomorphic interaction dimensions [32], which assess relaxed, nonanxious engagement with AI and the tendency to interact with AI in a human-like, conversational manner. These dimensions do not directly measure leadership behaviors but rather the psychological foundation without which proactive shaping of AI's future in nursing is unlikely to occur.

This mapping provides a theoretically grounded operationalization of the NURSES framework into measurable psychological and competency variables, enabling empirical examination of the interrelationships among its 6 components through network analysis. The mapping is not intended as an exhaustive or exclusive operationalization but as a starting point grounded in existing validated instruments and established theoretical links.

Data Collection

Researchers (XF and JW) contacted class administrators at target institutions via social media (WeChat; Tencent Holdings Limited). After explaining the study objectives and procedures and obtaining permission, they joined the corresponding class groups. Study information was disseminated through these class groups. Potential participants voluntarily confirmed their participation online after reading the informed consent form. The research team screened applicants based on predetermined inclusion and exclusion criteria. Students meeting the criteria were invited to participate in the survey. To facilitate organization and enhance response efficiency, researchers distributed electronic questionnaire links uniformly in classrooms during after-school hours. Data collection occurred from September 25 to September 28, 2025, using the online survey platform "WJX" for questionnaire distribution and retrieval. Participants independently completed the questionnaire after clicking the link and could exit at any time during the process. Completing the full questionnaire took an average of 10 - 15 minutes. A total of 986 questionnaires were distributed, with 982 returned. Platform settings ensured only fully completed questionnaires were recorded as valid submissions, eliminating incomplete submissions from being counted. Verification confirmed that returned questionnaires were valid. Data from 982 participants were ultimately analyzed, yielding a valid response rate of 99.59%.

Data Analysis

Descriptive statistics were calculated using SPSS (version 25.0; IBM Corp), with means and SDs for continuous variables and frequencies (percentages) for categorical variables. All continuous variables underwent normality testing in this study, and the results indicated that their distributions were generally normal [39].

Network analysis was performed using R (version 4.3.0; R Core Team). In 2-tailed tests, $P<.05$ was considered statistically significant. A Gaussian graph model was selected to construct the network structure using the R package "*qgraph*" [37]. To reduce spurious correlations and obtain a simplified network, the graph minimum absolute contraction and selection operator was used. The graph minimum absolute contraction and selection operator implements regularization techniques to constrain small partial correlations to zero, effectively eliminating potential false positive edges and yielding sparse, interpretable network structures [40]. To avoid overfitting and determine optimal tuning parameters, an extended Bayesian information criterion model is used to select the best-fitting model. This approach ensures that only the most robust and meaningful connections are retained in the final network.

In a network, nodes represent variables, and edges between 2 nodes represent the unique relationship between them, controlling for the influence of other nodes. Red and blue edges denote negative and positive correlations, respectively. The thickness of the edges corresponds to the strength of the association; thicker edges indicate stronger correlations [41]. The R package "*qgraph*" is used to calculate the strength, proximity, and betweenness centrality of nodes in a network [37]. Strength is defined as the sum of the absolute values of the edge weights connecting a node to all other nodes. Previous studies have shown that strength is the most stable and interpretable centrality measure in networks [42]. Compared to other centrality measures such as closeness and betweenness centrality, strength is more suitable for evaluating the importance of nodes [43]. Therefore, this study adopts strength as the centrality metric, with nodes exhibiting the highest strength

XSL•FO
RenderX

values being regarded as the most influential nodes within the network.

Additionally, the R package "*networktools*" was used to calculate bridging strength and identify potential bridging nodes [44]. Bridging strength refers to the sum of the absolute values of the edge weights connecting a node to other community nodes, indicating the node's importance to other communities. Nodes with the highest bridging strength values are considered bridging nodes within the network. Bridging nodes connect multiple domains, meaning interventions targeting these nodes can simultaneously impact multiple domains, which maximizes the effectiveness of the intervention.

Use the R package "*mgm*" to calculate the predictability for each node [45]. Predictability is a metric that reflects the extent to which the variance of a specific node can be explained by all its neighbors in the network. Nodes with high predictability are susceptible to the influence of their neighboring nodes.

The R package "*bootnet*" assesses the accuracy and stability of networks [46]. First, the accuracy of edge weights (the number of bootstrap samples is 1000) was examined by calculating their 95% CIs using nonparametric bootstrapping. Second, the stability of the network model is assessed using the case deletion procedure and the correlation stability (CS) coefficient, which indicates the proportion of samples that can be excluded from the analysis while maintaining at least 0.7 correlation with the original sample. According to network analysis guidelines, the CS coefficient should not fall below 0.25, and a value above 0.50 indicates sufficient stability [47]. Finally, a self-assessment method was used to conduct a difference test, examining variations in edge weights, node strengths, and node bridge strengths.

Results

General Information for Undergraduate Nursing Students

A total of 982 graduate nursing students were included in the statistical analysis, with their characteristics listed in Table 2. The mean age of participants was 21.7 years, and the vast majority (865/982, 88.1%) of participants were female. Among participants, 511 out of 982 (52.0%) resided in rural areas. The number of students in the first, second, third, and fourth years was 246 out of 982 (25.1%) participants, 242 out of 982 (24.6%) participants, 251 out of 982 (25.6%) participants, and 243 out of 982 (24.7%) participants, respectively. Among the 982 undergraduate nursing students, daily active use of GenAI tools was reported as follows: 22 out of 982 (2.2%) participants rarely used them, 608 out of 982 (61.9%) participants used them several times weekly, 297 out of 982 (30.2%) participants used them 3 - 5 times daily, and 55 out of 982 (5.6%) participants used them more than 5 times daily.

XSL•FO
RenderX

Table . General information on nursing undergraduates.

Statistical variable	Frequency, n (%)
Sex	
Male	117 (11.9)
Female	865 (88.8)
Age (years)	
<18	42 (4.3)
18-20	733 (74.6)
20-22	204 (20.8)
≥22	3 (0.3%)
Family residence	
Townships	471 (48.0)
Rural	511 (52.0)
Have you ever served as a student leader?	
Yes	479 (48.8)
No	503 (51.2)
Grade level	
Freshman year	246 (25.1)
Sophomore year	242 (24.6)
Junior year	251 (25.6)
Senior year	243 (24.7)
The average daily time spent using electronic devices (such as smartphones, computers, tablets, etc) is approximately (hours)	
<2	42 (4.3)
2-4	291 (29.6)
4-6	439 (44.7)
6-8	161 (16.4)
>8	49 (5.0)
On average, how often do you actively use AI[a] tools (such as ChatGPT [OpenAI], Wenxin Yiyan [Baidu], Deepseek [DeepSeek Artificial Intelligence Co, Ltd], etc) each day?	
Hardly ever used	22 (2.2)
Several times a week	608 (61.9)
3-5 times daily	297 (30.2)
>5 times a day	55 (5.6)
What is your primary purpose for using AI tools?	
Learning (such as researching information, assisting with homework, etc)	929 (94.6)
Entertainment and Leisure (such as smart chatbots, image/music generation, and video/audio recommendations)	468 (47.7)
Daily life (such as planning itineraries and obtaining health advice)	615 (62.6)
Work efficiency (eg, report writing, PowerPoint [Microsoft Corp] creation, data analysis)	595 (60.6)
Social needs (such as helping edit Moments captions and recommending chat topics)	336 (34.2)
Others	12 (1.2)

XSL•FO
RenderX

Statistical variable	Frequency, n (%)
How familiar do you consider yourself to be with AI technology?	
I have absolutely no idea.	16 (1.6)
I've heard of it but don't know much about it.	117 (11.9)
General understanding	569 (57.9)
Fairly familiar	262 (26.7)
Very familiar	18 (1.8)

[a]AI: artificial intelligence.

Descriptive Statistics of Variables (List in Table 3)

Descriptive statistics for nursing undergraduates' EI, AILS, and AS are presented in Table 3. Normality tests confirmed that all variables were approximately normally distributed. Data were described using mean (SD). SDs fell within normal ranges, with no outliers. Notably, "Emotional Perception (EI1)" demonstrated the highest predictive power and served as the critical component of the entire network.

Table . Descriptive statistics of emotional intelligence, AI self-efficacy, and AI literacy variables among nursing undergraduates (n=982).

Variable	Name	Mean (SD)	Expected influence	Bridge expected influence	Predictability
Emotional intelligence	EI[a]	3.57 (0.55)			
Emotional perception	EI1	3.41 (0.54)	1.248	0.427	0.832
Self-emotion management	EI2	3.51 (0.56)	1.016	0.118	0.788
Managing others' emotions	EI3	3.67 (0.61)	1.099	−0.006	0.801
Emotional application	EI4	3.68 (0.62)	1.355	−0.024	0.840
AI[b] self-efficacy	AILS[c]	4.68 (0.78)			
Assistance	AILS1	4.55 (0.84)	1.075	0.129	0.627
Anthropomorphic interaction	AILS2	4.62 (0.83)	1.091	0.069	0.679
Comfort with AI	AILS3	4.96 (0.98)	0.991	0.242	0.686
Technical skills	AILS4	4.59 (0.92)	0.854	0.077	0.496
AI literacy	AS[d]	4.60 (0.59)			
Awareness	AS1	5.19 (0.95)	0.862	0.015	0.597
Usage	AS2	4.43 (1.12)	0.838	−0.030	0.545
Evaluation	AS3	4.83 (0.94)	1.323	−0.011	0.726
Ethics	AS4	3.94 (1.19)	0.837	−0.033	0.287

[a]EI: emotional intelligence.

[b]AI: artificial intelligence.

[c]AILS: AI self-efficacy.

[d]AS: AI literacy.

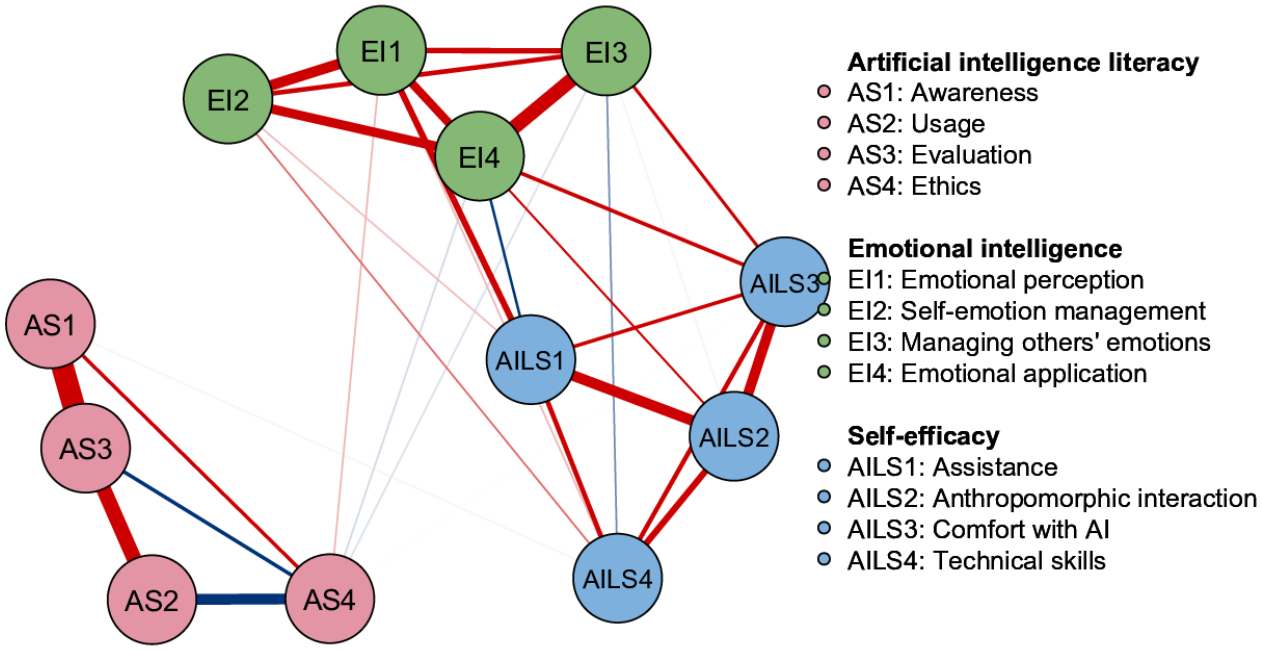
Network Analysis of EI, AILS, and AS Among Nursing Undergraduates

Figure 1 illustrates the network structure of EI, AILS, and AS among undergraduate nursing students. Node colors represent theoretical constructs: blue=AILS, green=EI, and red=AS. Letters on nodes denote dimension codes. Red indicates positive correlations; blue indicates negative correlations. Thicker lines denote stronger connections. Of the 66 possible connections, 54 (81.82%) were nonzero, and the network density was 0.82, indicating that the variables were generally closely interconnected within this sample. Among these, 41 were positive correlations and 13 were negative correlations. From the perspective of community structure, the nodes are primarily clustered into 3 distinct communities based on theoretical constructs, including emotional intelligence (EI1-EI4), AI self-efficacy (AILS1-AILS4), and AI competence (AS1-AS4). Connections within each community are relatively dense, while cross-community connections are relatively sparse, which is largely consistent with the dimensional divisions of the NURSES framework. Stronger connections primarily occurred within dimensions rather than between them. The 3 strongest intradimensional connections were EI4 (emotional use)-EI3 (managing others' emotions), AS3 (assessment)-AS1 (awareness), and AILS2 (anthropomorphic interaction)-AILS1

XSL•FO
RenderX

(assistance). The strongest interdimensional connection was between EI1 (emotional perception) and AILS1 (assistance). All edge weights (association strengths) are listed in Table 4.

Figure 1. Network analysis of emotional intelligence (EI), AI self-efficacy (AILS), and AI literacy (AS) among nursing undergraduates. Node colors represent theoretical constructs: blue=AILS, green=EI, and red=AS. Letters on nodes denote dimension codes. Red indicates positive correlations; blue indicates negative correlations. Thicker lines denote stronger connections. AILS: artificial intelligence self-efficacy; AS: AI literacy; EI: emotional intelligence.



XSL•FO
RenderX

Table . Marginal weights of emotional intelligence (EI), artificial intelligence (AI) self-efficacy (AILS), and AI literacy (AS) among nursing undergraduates.

	EI1[a]	EI2[b]	EI3[c]	EI4[d]	AILS1[e]	AILS2[f]	AILS3[g]	AILS4[h]	AS1[i]	AS2[j]	AS3[k]	AS4[l]
EI1	0.00	0.34	0.20	0.28	0.22	0.11	0.00	0.05	0.00	–0.00	0.00	0.05
EI2	0.34	0.00	0.18	0.31	0.05	0.02	–0.02	0.07	–0.01	0.00	–0.01	0.02
EI3	0.20	0.18	0.00	0.44	0.00	–0.03	0.14	–0.07	0.00	0.00	0.00	–0.04
EI4	0.28	0.31	0.44	0.00	–0.12	–0.02	0.15	0.18	–0.01	0.00	0.00	–0.04
AILS1	0.22	0.05	0.00	–0.12	0.00	0.34	0.15	0.18	–0.01	0.00	–0.01	0.00
AILS2	0.11	0.02	–0.03	–0.02	0.34	0.00	0.33	0.22	0.00	0.00	–0.01	0.01
AILS3	0.00	–0.02	0.14	0.15	0.15	0.33	0.00	0.17	0.00	0.00	0.00	–0.03
AILS4	0.05	0.07	–0.07	0.18	0.18	0.22	0.17	0.00	0.03	–0.03	0.02	0.00
AS1	0.00	–0.01	0.00	–0.01	–0.01	0.00	0.00	0.03	0.00	0.00	0.66	0.15
AS2	–0.00	0.00	0.00	0.00	0.00	0.00	0.00	–0.03	0.00	0.00	0.46	–0.35
AS3	0.00	–0.01	0.00	0.00	–0.01	–0.01	0.00	0.02	0.66	0.46	0.00	–0.15
AS4	0.05	0.02	–0.04	–0.04	0.00	0.01	–0.03	0.00	0.15	–0.35	–0.15	0.00

[a]EI1: emotional perception.

[b]EI2: self-emotion management.

[c]EI3: managing others' emotions.

[d]EI4: emotional application.

[e]AISL1: assistance.

[f]AISL2: anthropomorphic interaction.

[g]AISL3: comfort with artificial intelligence.

[h]AISL4: technical skills.

[i]AS1: awareness.

[j]AS2: usage.

[k]AS3: evaluation.

[l]AS4: ethics.

Node Centrality Measures

The standardized estimates of centrality indicators for factors influencing EI, AILS, and AS among nursing undergraduates are shown in Figure 2. Emotional use (EI4) and assessment (AS3) exhibit the highest strength centrality, indicating that these 2 variables showed the highest strength centrality, indicating they are highly connected within the network. Emotional perception (EI1) and ethics (AS4) exhibited the strongest mediating centrality, signifying their pivotal bridging role in network information transmission with potent moderating capacity; emotional perception (EI1) and awareness (AS1) also demonstrated the most prominent proximity centrality, reflecting their highest information transmission efficiency and accessibility within the network. Stability tests for network centrality metrics revealed correlation stability coefficients of 0.749, 0.749, and 0.361 for node strength, intermediary centrality, and closeness centrality, respectively (see Figure 3). This indicates robust stability for these centrality measures, with node strength demonstrating the highest stability.

XSL•FO
RenderX

Figure 2. Standardized estimates of centrality measures for emotional intelligence, artificial intelligence (AI) self-efficacy, and AI literacy among nursing undergraduates. AILS1: assistance; AILS2: anthropomorphic interaction; AILS3: comfort with AI; AILS4: technical skills; AS1: awareness; AS2: usage; AS3: evaluation; AS4: ethics; EI1: emotional perception; EI2: self-emotion management; EI3: managing others' emotions; EI4: emotional application.

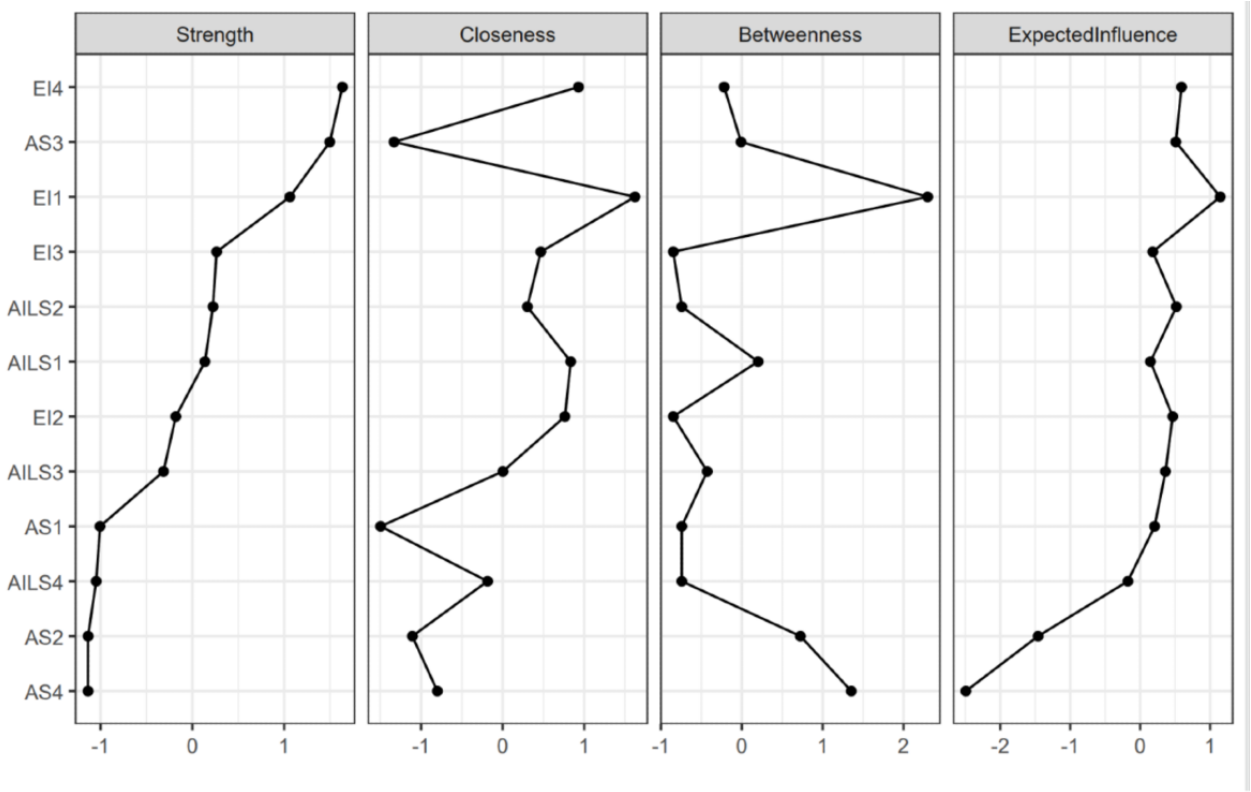
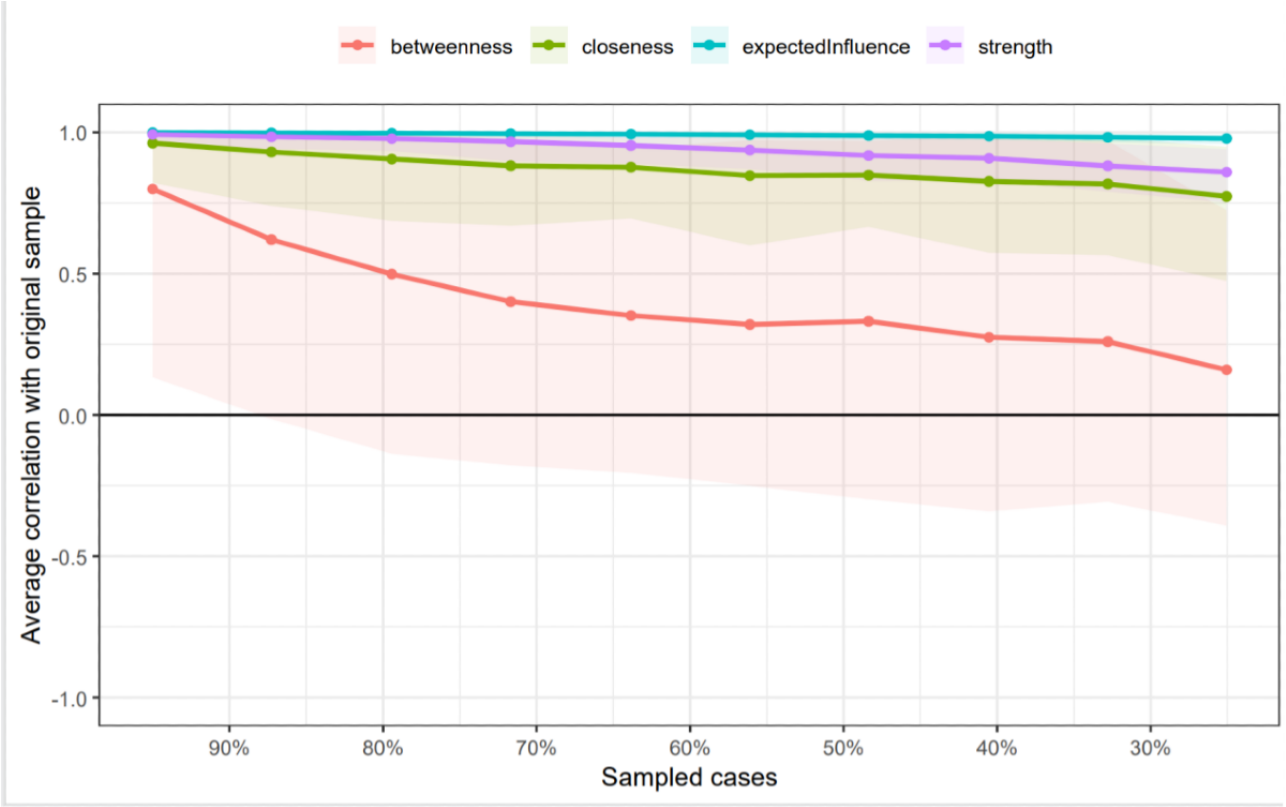


Figure 3. Correlation stability coefficients for node strength, betweenness, and closeness.

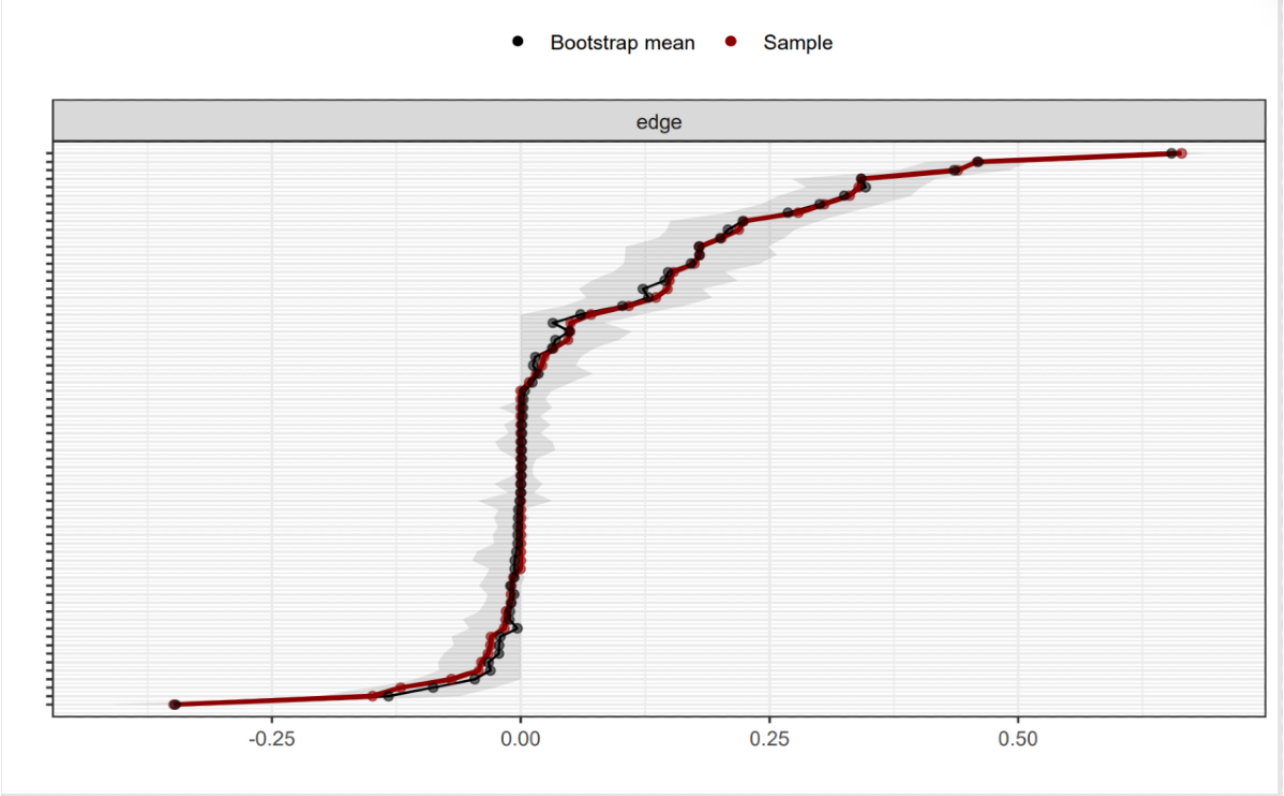


XSL•FO
RenderX

Edge Accuracy

The accuracy of edge weights was estimated using nonparametric bootstrapping. Results indicate that the 95% CIs for edge weights in the network are narrow, suggesting that the estimated edge weights in this study are highly precise (Figure 4).

Figure 4. Guiding edge weights for emotional intelligence, artificial intelligence (AI) self-efficacy, and AI literacy among nursing undergraduates with 95% CIs.

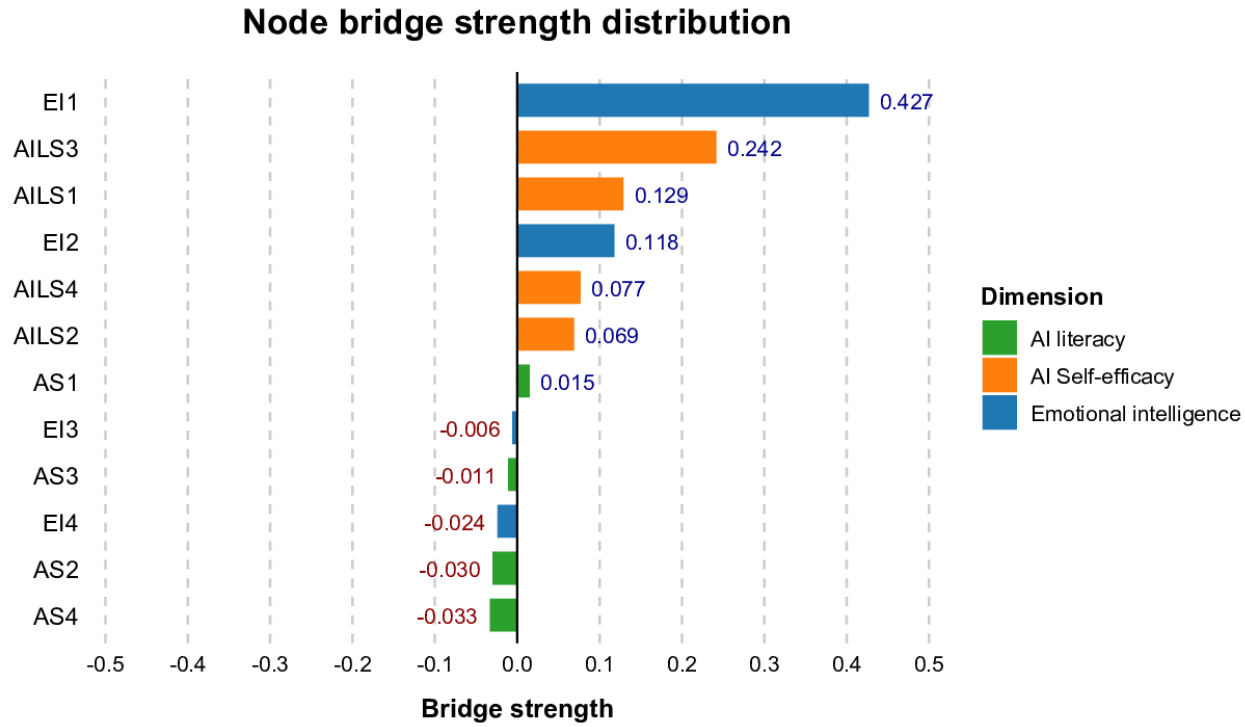


Node Bridge Strength

Figure 5 displays the bridge strength for each node. EI1 “Emotion Awareness” (bridge strength=0.427) exhibits the highest bridge strength, followed by ALS3 “Comfort with AI” (bridge strength=0.242), ALS1 “Assistance” (bridge strength=0.129), and EI2 “Self-Emotion Management” (bridge strength=0.118), suggesting they may serve as bridging nodes in this network.

XSL•FO
RenderX

Figure 5. Depicts the bridge strength (z-score) of each node in the network. AILS1: assistance; AILS2: anthropomorphic interaction; AILS3: comfort with AI; AILS4: technical skills; AS1: awareness; AS2: usage; AS3: evaluation; AS4: ethics; EI1: emotional perception; EI2: self-emotion management; EI3: managing others' emotions; EI4: emotional application.



Discussion

Principal Findings

This study is the first to use network analysis methods within the NUR.S.E.S. framework to examine the associations among EI, AILS, and AS among undergraduate nursing students. The network analysis identified that emotion use (EI4) and assessment ability (AS3) showed the highest strength centrality, indicating their prominent positions within the network. Emotion perception (EI1) had the highest bridge strength, suggesting it may serve as a connecting node between the EI domain and the AILS and AS domains. These findings provide insights into the psychological mechanisms underlying the development of AS among nursing students, offer a concrete explanation of how the 6 dimensions of the NURSES framework interact at the psychological level, and provide a critical foundation for the future development of targeted, comprehensive interventions that address the full spectrum from emotional foundations to technical confidence.

Overall Characteristics of EI, AILS, and AS

This study conducted a descriptive analysis of nursing students' EI, AILS, and AS. Results indicate that undergraduate nursing students' AILS (mean 4.68, SD 0.78) and AS (mean 4.60, SD 0.59) were generally at a moderately high level, while EI (mean 3.57, SD 0.55) scored relatively low. This situation highlights a key phenomenon: nursing students possess strong confidence in their ability to learn and master AI technologies and have established a foundational level of AS. However, the development of their EI has not kept pace, consistent with the findings of Bewersdorff et al [38]. From a specific-dimension perspective, within AILS, the "Comfort Level with Artificial Intelligence (AILS3)" dimension scored the highest (mean 4.96, SD 0.98), indicating that nursing students experience relatively positive feelings when interacting with AI. This finding aligns with the research results of Volpato et al [48]. In terms of EI, the "Emotional perception (EI1)" dimension had the lowest mean score 3.41 (SD 0.54), suggesting that nursing students have significant room for improvement in their ability to accurately identify and understand their own and others' emotions. This weakness may directly impact nursing students' empathetic communication and emotional regulation in high-pressure clinical settings, posing a potential risk to patient safety; failure to promptly identify patients' anxiety or distress may delay nursing decisions.

Emotion Perception: the Psychological Bridge Connecting the Emotional World and the AI World

The network structure constructed in this study showed that emotional perception (EI1), defined as the ability to recognize one's own and others' emotional states, was strongly connected to AILS and AS nodes. In this sample, emotional perception had the highest bridge strength, indicating that it was relatively well-connected to nodes in other communities. This pattern is consistent with Fu's [49] finding that emotional competence is associated with mental health and with deeper engagement with AI in educational contexts. The high bridge strength suggests that the ability to recognize emotions may be associated with trust in AI technology, which aligns with previous research showing that individuals with higher EI tend to monitor emotional cues and manage stress effectively [50]. Specifically, the positive connection between emotional perception and the assistance dimension (AILS1) indicated that students who reported greater awareness of their own confusion, anxiety, or sense of accomplishment during technology learning were also

XSL•FO
RenderX

more likely to view AI as a helpful tool [51]. This pattern is consistent with the observation by Mosleh et al [35]. that AI integration in teaching creates opportunities for emotional recognition and regulation, which may be easier for students with stronger emotional perception.

Emotional Application and Evaluation: the Engine Driving the Development of AS

The centrality measures showed that emotional use (EI4) and assessment (AS3) had the highest strength centrality, indicating that they were highly connected nodes within the network [52]. This pattern suggests that AS is associated not only with cognitive factors but also with emotional factors. Emotional use, defined as the ability to use emotional information for planning, creativity, and problem-solving, was positively connected with critical evaluation of AI outputs [13]. This finding is consistent with the observation that some individuals turn to GenAI for emotional guidance [50] and that emotional use may help nursing students engage in critical thinking when faced with complex AI outputs [53]. Evaluation ability, which concerns the judgment of AI-generated content [33], was closely connected with emotional use in the network. This pattern aligns with the finding by Deng and Chen [19] that AI-driven environments can support cognitive flexibility and independent judgment, possibly through emotional resilience. These exploratory observations invite further investigation using longitudinal or experimental designs to examine whether and how emotional use and evaluation skills might be related to AS [54].

AILS: a Potentially Important Bridging Node Between Emotion and AS

Network analysis results showed that EI, represented by emotional perception (EI1), had a positive connection with AILS, represented by assistance (AILS1) and comfort with AI (AILS3). For example, the edge weight between EI1 and AILS1 was 0.22. AILS also showed positive connections with AS, represented by usage (AS2) and evaluation (AS3). This pathway pattern, in which EI connected to AILS and AILS in turn connected to AS, is consistent with a mediating structure [55]. However, due to the cross-sectional design, this pattern should be interpreted as exploratory. It suggests that AILS may occupy a bridging position, but formal mediation analysis using longitudinal data would be required to test this hypothesis [49]. Prior research has shown that EI is associated with self-efficacy and with cognitive engagement with AI [49]. The connection pattern observed here is compatible with the idea that confidence in technical skills (AILS4) is associated with more frequent use of AI tools (AS2) [56]; that comfort with AI (AILS3) and anthropomorphic interaction (AILS2) are associated with lower anxiety and more exploration [25]; and that the belief that AI serves as an effective assistant (AILS1) is associated with reduced operational anxiety [53]. These observations are hypothesis-generating and require replication in independent samples with stronger designs. The significance of this approach for clinical practice lies in the fact that only by simultaneously strengthening both the emotional foundation and technical confidence can knowledge be translated into safe and effective clinical practice.

The Deep Mechanisms of AS Revealed by Negative Correlations

The most significant negative correlation was found between “Ethics (AS4)” and “Usage (AS2)” in the network. This indicates that at the current stage, nursing students who are more concerned about AI ethical issues (such as data privacy) may be more cautious in their use of AI tools, even exhibiting lower usage frequency [57]. Although this contradicts findings showing positive correlations between EI and reasons for adopting AI services or adoption/usage intentions [58], it reflects the development of critical thinking. When nursing students recognize the potential risks of technology, they may transition from blind tool users to critical thinkers, adopting more rational, selective use behaviors [59]. This finding resonates with the perspective of Ricon [60] that premature or excessive enthusiasm in medical AI may obscure ethical concerns, while measured caution is a prerequisite for responsible innovation. Participants in one study expressed concerns about potential privacy breaches and the risk of personal data theft [61]. Therefore, educators should not pursue unrestricted usage frequency but instead focus on cultivating students’ high-level application skills grounded in strong ethical vigilance.

Second is the negative correlation between “Emotional Intelligence in Others (EI3)” and “Anthropomorphic Interaction (AILS2).” This negative correlation may reflect multiple mechanisms. First, individuals with high EI are able to clearly distinguish the fundamental differences between human emotional connections and AI-simulated interactions. The ability to manage others’ emotions manifests as the capacity to perceive, influence, and regulate others’ emotions in interpersonal interactions, whereas the tendency toward anthropomorphic interaction is a psychological inclination to ascribe human characteristics to AI and interact with it in a social manner. Nursing students with high EI have a deeper understanding that, although AI can simulate empathetic responses, it lacks genuine emotional intent and reciprocity. Consequently, they consciously avoid viewing AI as an emotional substitute to prevent the development of inappropriate emotional dependence, which is consistent with Dakakni and Safa [62]. Second, cultural factors may play a role. In the context of Eastern culture, which emphasizes interpersonal harmony and emotional connection, nursing education places a high value on “authentic interpersonal interaction.” Nursing students with high EI may have internalized these professional values, leading them to adopt a more cautious or even resistant attitude toward anthropomorphizing AI. Third, from the perspective of clinical practice, this negative correlation carries positive ethical implications. The core of nursing practice lies in establishing trust within therapeutic interpersonal relationships. If nursing students are overly inclined to anthropomorphize AI and blur the boundaries between humans and machines, it may diminish their sensitivity to the emotional needs of real patients. Therefore, the lower tendency toward anthropomorphic interaction exhibited by nursing students with high EI can be viewed as an adaptive strategy for upholding the core values of the nursing profession [63].

XSL•FO
RenderX

System Integration Under the NUR.S.E.S. Framework: Synergistic Development of EI, AILS, and AS

The findings of this study provide empirical support for the NUR.S.E.S. framework, revealing the underlying psychological mechanisms behind its 6 dimensions. Specifically, emotional perception (EI1), as the strongest bridge node, empirically demonstrates the central role of E (EI) and S (social-emotional comfort) as the emotional cornerstones within the framework [64]; emotional application (EI4) and assessment (AS3), as core driving nodes, correspond to E and R (defect identification), respectively, indicating that emotional competence and technical critical thinking share an intrinsic synergistic mechanism; the various dimensions of AILS occupy key positions in the network that connect emotion and cognition, validating the theoretical premise within the framework that S (AILS) serves as the core driving force [65]. Ultimately, the synergy of all these elements achieves the framework's ultimate goal—S (shape the future)—that is, shaping the future [27].

In addition, this study provides an initial exploratory mapping of the interrelated factors associated with AS among nursing undergraduates, and these findings may offer insights for future research. From a practical perspective, nodes with high centrality, such as emotional use and evaluation, could exert influence on the overall network configuration through their connections with other nodes. Bridge nodes, such as emotional perception and comfort with AI, may transmit influences across different domains, making them potential targets for hypothesis-driven interventions. Accordingly, nursing educators might consider several tentative strategies, fostering students' emotional perception skills, which appear to link EI with AILS; creating low-anxiety AI learning environments to enhance comfort with AI; helping students set personalized learning goals and reflect on their emotional responses to strengthen emotional use and critical evaluation; and introducing case-based learning that combines clinical scenarios with AI outputs to ground technical critique and ethical decision-making in empathy for patients' emotional needs. Such a multidimensional approach, addressing emotional, cognitive, and skill components simultaneously, may support the development of nursing students as future professionals who can shape AI integration in nursing practice. However, all suggestions require empirical validation through longitudinal or experimental designs.

Research Significance, Limitations, and Future Directions

This study is the first to visualize, through network analysis, the patterns of association among EI, AILS, and AS in nursing students. It identified nodes that showed high centrality and bridging strength (emotional perception, emotional application, evaluation, comfort with AI). The observed association pattern suggests that emotional competence may be closely linked with technological confidence, which in turn may be linked with AS. These exploratory findings offer preliminary clues for future hypothesis-driven research. However, they do not establish causality, nor do they directly identify effective intervention targets. Any educational implications drawn from these results are speculative and require empirical validation [56].

This study has several limitations. First, the cross-sectional design cannot establish causal relationships between variables. Although the network pattern is consistent with some causal pathways, a reverse pathway (eg, higher AS leading to greater AILS) is equally plausible. Future research should use longitudinal or intervention designs to test dynamic relationships. Second, the sample originated from a single university, limiting generalizability. Cross-regional and cross-cultural comparisons are needed. Third, self-report scales may introduce bias; future studies could incorporate behavioral data and qualitative interviews.

Despite these limitations, this study offers profound theoretical insights and precise practical guidance for cultivating a new generation of nursing professionals who are "technically proficient, confident, and emotionally intelligent" in the era of AI. Future education and research should focus on developing integrated intervention programs that combine emotional education with technical training, empowering nursing students to embrace the challenges of the intelligent health care era fully.

Acknowledgments

The authors confirm that no generative artificial intelligence tools were used at any stage of this manuscript. Ruihua Jin and Panpan Huai share corresponding authorship.

Funding

This work was supported by the GenAI-Empowered Project Case Teaching Application Research – Taking "Nursing Psychology" as an Example (JGYB2543).

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Xiaohui Fan: Conceptualization, Methodology, Writing review & editing, Visualization. Jingjing Sun: Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing original draft, Writing review & editing, Visualization. Linghui Zhang & Jinrong Wang : Conceptualization, Methodology, Software, Formal analysis, Data curation. BinHao Dong & Xiaohong

XSL•FO
RenderX

Gao: Investigation. Ruihua Jin: Conceptualization, Formal analysis, Resources, Supervision. Panpan Huai: Conceptualization, Formal analysis, Resources, Supervision.
All authors have reviewed and approved the final manuscript and consent to its publication.

Conflicts of Interest

None declared.

References

1. Sharma N, Jindal N. Emerging artificial intelligence applications: metaverse, IoT, cybersecurity, healthcare - an overview. Multimed Tools Appl 2024;83(19):57317-57345. [doi: 10.1007/s11042-023-17890-6]
2. Labrague LJ, Aguilar-Rosales R, Yboa BC, Sabio JB. Factors influencing student nurses' readiness to adopt artificial intelligence (AI) in their studies and their perceived barriers to accessing AI technology: a cross-sectional study. Nurse Educ Today 2023 Nov;130:105945. [doi: 10.1016/j.nedt.2023.105945] [Medline: 37625351]
3. García Peñalvo FJ, Vázquez Ingelmo A. What do we mean by GenAI? A systematic mapping of the evolution, trends, and techniques involved in generative AI. Int J Interact Multimed Artif Intell 2023;8(4):7-16. [doi: 10.9781/ijimai.2023.07.006]
4. Labrague L. A five-tier framework for guiding responsible AI use in nursing students' coursework: a faculty guide. Teach Learn Nurs 2026 Jan;21(1):62-66. [doi: 10.1016/j.teln.2025.09.012]
5. Amoah A, Asiama RK, Kwablah E. ChatGPT early usage among students: a global evidence of determinants. Dev Sustain Econ Finance 2025 Sep;7:100065. [doi: 10.1016/j.dsef.2025.100065]
6. Postrano LMC. Nursing students' AI literacy and ethical understanding of AI in nursing education: reflections and implications. Teaching and Learning in Nursing 2026 Jan;21(1):109. [doi: 10.1016/j.teln.2025.09.026]
7. ANA Center for Ethics and Human Rights. The ethical use of artificial intelligence in nursing practice. : American Nurses Association; 2022 URL: https://www.nursingworld.org/~48f653/globalassets/practiceandpolicy/nursing-excellence/ana-position-statements/the-ethical-use-of-artificial-intelligence-in-nursing-practice_bod-approved-12_20_22.pdf [accessed 2026-06-14]
8. El-Sayed BKM, El-Sayed AAI, Alsenany SA, Asal MGR. The role of artificial intelligence literacy and innovation mindset in shaping nursing students' career and talent self-efficacy. Nurse Educ Pract 2025 Jan;82:104208. [doi: 10.1016/j.nepr.2024.104208] [Medline: 39637623]
9. Agustin A, Bhatti S, Bourne J, et al. The history of technology and innovation in nursing. Can J Nurs Inform 2020;15(2) [FREE Full text]
10. McGrow K. Artificial intelligence: essentials for nursing. Nursing 2019 Sep;49(9):46-49. [doi: 10.1097/01.NURSE.0000577716.57052.8d] [Medline: 31365455]
11. Le Lagadec D, Jackson D, Cleary M. Artificial intelligence in nursing education: prospects and pitfalls. J Adv Nurs 2024 Oct;80(10):3883-3885. [doi: 10.1111/jan.16276] [Medline: 38864369]
12. Al-Olaimat K, Salameh B, Abdulhalim Alqadi R, et al. Readiness and acceptance of nursing students regarding AI-based health care technology on the training of nursing skills in Saudi Arabia: cross-sectional study. JMIR Nurs 2025 Jul 30;8:e71653. [doi: 10.2196/71653] [Medline: 40737490]
13. Vowels LM. Are chatbots the new relationship experts? Insights from three studies. Comput Hum Behav Artif Hum 2024 Aug;2(2):100077. [doi: 10.1016/j.chbah.2024.100077]
14. Lin H, Chen Q. Artificial intelligence (AI) -integrated educational applications and college students' creativity and academic emotions: students and teachers' perceptions and attitudes. BMC Psychol 2024 Sep 16;12(1):487. [doi: 10.1186/s40359-024-01979-0] [Medline: 39285268]
15. Por J, Barriball L, Fitzpatrick J, Roberts J. Emotional intelligence: its relationship to stress, coping, well-being and professional performance in nursing students. Nurse Educ Today 2011 Nov;31(8):855-860. [doi: 10.1016/j.nedt.2010.12.023] [Medline: 21292360]
16. Sallam M. The utility of chatgpt as an example of large language models in healthcare education, research and practice: systematic review on the future perspectives and potential limitations. MedRxiv. Preprint posted online on Feb 21, 2023. [doi: 10.1101/2023.02.19.23286155]
17. Ching YC, Ho YC. AI virtual human-augmented game-based teaching to enhance emotional intelligence in nursing students: protocol for a single-group pretest-posttest action research study. JMIR Res Protoc 2025 Oct 17;14:e80290. [doi: 10.2196/80290] [Medline: 41106814]
18. Derakhshan A. EFL students' perceptions about the role of generative artificial intelligence (GAI)-mediated instruction in their emotional engagement and goal orientation: a motivational climate theory (MCT) perspective in focus. Learn Motiv 2025 May;90:102114. [doi: 10.1016/j.lmot.2025.102114]
19. Deng K, Chen X. Exploring the impact of AI-driven emotional resilience on academic persistence, motivation, cognitive flexibility, and autonomy in self-regulated learning: a self-determination theory perspective. Learn Motiv 2025 Nov;92:102167. [doi: 10.1016/j.lmot.2025.102167]
20. Locke EA. Self-efficacy: the exercise of control. Pers Psychol 1997;50(3):801-804 [FREE Full text]

XSL•FO
RenderX

21. Zeng Q, Zhu J, Yang Q, Su S, Huang X. AI-driven maternal and child healthcare nursing education: network analysis of self-efficacy and usage demands. Nurse Educ Today 2026 Jan;156:106899. [doi: 10.1016/j.nedt.2025.106899] [Medline: 41100953]
22. Alruwaili MM, Abuadas FH, Alsadi M, et al. Exploring nurses' awareness and attitudes toward artificial intelligence: implications for nursing practice. Digit Health 2024;10:20552076241271803. [doi: 10.1177/20552076241271803] [Medline: 39114115]
23. Shahrour G, Dardas LA. Acute stress disorder, coping self-efficacy and subsequent psychological distress among nurses amid COVID-19. J Nurs Manag 2020 Oct;28(7):1686-1695. [doi: 10.1111/jonm.13124] [Medline: 32767827]
24. Wang S, Sun Z, Wang H, Yang D, Zhang H. Enhancing student acceptance of artificial intelligence-driven hybrid learning in business education: interaction between self-efficacy, playfulness, emotional engagement, and university support. Int J Manage Educ 2025 Jul;23(2):101184. [doi: 10.1016/j.ijme.2025.101184]
25. Gao J, Zhang J, Li Y. Do AI chatbot-integrated writing tasks influence writing self-efficacy and critical thinking ability? An exploratory study. Comput Educ Artif Intell 2025 Dec;9:100472. [doi: 10.1016/j.caeai.2025.100472]
26. Lyu W, Salam ZA. AI-powered personalized learning: enhancing self-efficacy, motivation, and digital literacy in adult education through expectancy-value theory. Learn Motiv 2025 May;90:102129. [doi: 10.1016/j.lmot.2025.102129]
27. Hoelscher SH, Pugh A. N.U.R.S.E.S. embracing artificial intelligence: a guide to artificial intelligence literacy for the nursing profession. Nurs Outlook 2025;73(4):102466. [doi: 10.1016/j.outlook.2025.102466] [Medline: 40561583]
28. Epskamp S, Fried EI. A tutorial on regularized partial correlation networks. Psychol Methods 2018 Dec;23(4):617-634. [doi: 10.1037/met0000167] [Medline: 29595293]
29. Gignac GE, Zajenkowski M. People tend to overestimate their romantic partner's intelligence even more than their own. Intelligence 2019 Mar;73:41-51. [doi: 10.1016/j.intell.2019.01.004]
30. Schutte NS, Malouff JM, Hall LE, et al. Development and validation of a measure of emotional intelligence. Pers Individ Dif 1998 Aug;25(2):167-177. [doi: 10.1016/S0191-8869(98)00001-4]
31. Caikang W. A study on the relationship between emotional intelligence and anxiety, depression, and mood in college students. Chinese J Clin Psychol 2002;10(4):298-299. [doi: 10.3969/j.issn.1005-3611.2002.04.021]
32. Wang B, Rau PLP, Yuan T. Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. Behav Inf Technol 2023 Jul 4;42(9):1324-1337. [doi: 10.1080/0144929X.2022.2072768]
33. Wang YY, Chuang YW. Artificial intelligence self-efficacy: scale development and validation. Educ Inf Technol 2024 Mar;29(4):4785-4808. [doi: 10.1007/s10639-023-12015-w]
34. Bewersdorff A, Nerdel C, Zhai X. How AI literacy correlates with affective, behavioral, cognitive and contextual variables: a systematic review. Comput Edu Artif Intell 2025 Dec;9:100493. [doi: 10.1016/j.caeai.2025.100493]
35. Mosleh SM, Alsaadi FA, Alnaqbi FK, Alkhzaimi MA, Alnaqbi SW, Alsereidi WM. Examining the association between emotional intelligence and chatbot utilization in education: a cross-sectional examination of undergraduate students in the UAE. Heliyon 2024 Jun 15;10(11):e31952. [doi: 10.1016/j.heliyon.2024.e31952] [Medline: 38868023]
36. Liu F, Zhou H, Yuan L, Cai Y. Effect of empathy competence on moral sensitivity in Chinese student nurses: the mediating role of emotional intelligence. BMC Nurs 2023 Dec 19;22(1):483. [doi: 10.1186/s12912-023-01650-w] [Medline: 38114985]
37. Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D. qgraph: network visualizations of relationships in psychometric data. J Stat Soft 2012 May;48(4):18. [doi: 10.18637/jss.v048.i04]
38. Bewersdorff A, Hornberger M, Nerdel C, Schiff DS. AI advocates and cautious critics: how AI attitudes, AI interest, use of AI, and AI literacy build university students' AI self-efficacy. Comput Educ Artif Intell 2025 Jun;8:100340. [doi: 10.1016/j.caeai.2024.100340]
39. Wang M, Yang M, Li X, et al. The interplay between perceived stress, academic control, achievement motivation, and procrastination among nursing graduate students: a network analysis. Nurse Educ Today 2025 Nov;154:106850. [doi: 10.1016/j.nedt.2025.106850] [Medline: 40779877]
40. Chen J, Guo L, Chen TZ, et al. Prediction and explanation of the increase in suicide risk of emerging adults: a comprehensive approach combining logistic regression, glasso network analysis, and Bayesian networks. J Affect Disord 2025 Aug;383:469-479. [doi: 10.1016/j.jad.2025.04.171]
41. Qian W, Tian X, Zhang J, Li X, Li J. The buffering structure of moral injury in nursing internship students: a network analysis based on moral courage, self-compassion, and professional values. Nurse Educ Today 2026 Jan;156:106902. [doi: 10.1016/j.nedt.2025.106902] [Medline: 41130036]
42. McNally RJ. Can network analysis transform psychopathology? Behav Res Ther 2016 Nov;86:95-104. [doi: 10.1016/j.brat.2016.06.006] [Medline: 27424882]
43. Su L, Jiang L, Ma Y, Wang Z, Wang X, Lin Y. Central and bridging roles of social relationships within the multilayer health ecology model in motoric cognitive risk syndrome: a network analysis. J Am Med Dir Assoc 2025 Sep;26(9):105771. [doi: 10.1016/j.jamda.2025.105771] [Medline: 40651497]
44. Zhuang S, Song Y, Jin S, et al. Exploring the influence of parenting styles and childhood abuse on impulsivity in young and middle-aged patients with new drug dependency: a network analysis. Addict Behav 2025 Nov;170:108423. [doi: 10.1016/j.addbeh.2025.108423] [Medline: 40618445]

XSL•FO
RenderX

45. Li P, Zhang H, Ma J, et al. Associations between lipid parameters and depression with the moderating role of inflammation: network analyses of 87,636 UK Biobank participants. Psychoneuroendocrinology 2025 Oct;180:107543. [doi: 10.1016/j.psyneuen.2025.107543] [Medline: 40652641]
46. Yao X, Yang Y, Lu D, et al. Central symptoms of depression and their associations with health-related quality of life among migrant older with children: a network analysis. Geriatric Nursing 2025 May;63:415-421. [doi: 10.1016/j.gerinurse.2025.03.039]
47. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: a tutorial paper. Behav Res Methods 2018 Feb;50(1):195-212. [doi: 10.3758/s13428-017-0862-1] [Medline: 28342071]
48. Volpato R, DeBruine L, Stumpf S. Trusting emotional support from generative artificial intelligence: a conceptual review. Comput Hum Behav Artif Hum 2025 Aug;5:100195. [doi: 10.1016/j.chbah.2025.100195]
49. Fu L. The role of STEM teachers' emotional intelligence and psychological well-being in predicting their artificial intelligence literacy. Acta Psychol (Amst) 2025 Mar;253:104708. [doi: 10.1016/j.actpsy.2025.104708] [Medline: 39813946]
50. Hussien RM, Khalil T, Almegewly WH, Khedr MA. Cultivating compassion: how standardized patient simulation-based training enhances nurses' emotional intelligence, empathy, and perception of the quality of patient care. Nurse Educ Pract 2025 Oct;88:104546. [doi: 10.1016/j.nepr.2025.104546] [Medline: 40961634]
51. Zhang P, Li CZ, Zhao YN, et al. The mediating role of emotional intelligence between negative life events and psychological distress among nursing students: a cross-sectional study. Nurse Educ Today 2016 Sep;44:121-126. [doi: 10.1016/j.nedt.2016.05.025] [Medline: 27429340]
52. Qiu SS, Zhang L, You F, Zhao X. Unpacking media channel effects on AI perception: a network analysis of AI information exposure across channels, overload, literacy, and anxiety among Chinese users. Comput Human Behav 2025 Dec;173:108790. [doi: 10.1016/j.chb.2025.108790]
53. Li H, Zhang Y, Chen M, Zhao T, Jou M. Creative personal identity in the age of generative AI: a social-cognitive pathway of AI literacy, self-efficacy, and mindset. Comput Human Behav 2026 Feb;175:108838. [doi: 10.1016/j.chb.2025.108838]
54. Quan Q, Gao Y, Wang Q. The impact of teacher emotional support on students' engagement in AI-mediated English learning environments: the mediating role of resilience and self-efficacy. Acta Psychol (Amst) 2025 Oct;260:105766. [doi: 10.1016/j.actpsy.2025.105766] [Medline: 41086745]
55. He T, Huang J, Li Y, et al. The mediation effect of AI self-efficacy between AI literacy and learning engagement in college nursing students: a cross-sectional study. Nurse Educ Pract 2025 Aug;87:104499. [doi: 10.1016/j.nepr.2025.104499] [Medline: 40784134]
56. Ding N, Chen M, Hu L. The impact of personality traits and AI literacy on the adoption intentions of AI among design faculty in Chinese higher education. Acta Psychol (Amst) 2025 Oct;260:105709. [doi: 10.1016/j.actpsy.2025.105709] [Medline: 41072293]
57. Tian J, Zhang R. Learners' AI dependence and critical thinking: the psychological mechanism of fatigue and the social buffering role of AI literacy. Acta Psychol (Amst) 2025 Oct;260:105725. [doi: 10.1016/j.actpsy.2025.105725] [Medline: 41076923]
58. Rasheed HMW, Chen Y, Khizar HMU, Safeer AA. Understanding the factors affecting AI services adoption in hospitality: the role of behavioral reasons and emotional intelligence. Heliyon 2023 Jun;9(6):e16968. [doi: 10.1016/j.heliyon.2023.e16968] [Medline: 37484303]
59. Nunes JR, Rita P, Costa Pinto D, Gonzalez-Jimenez H, Akdim K, Wagner RL. Moral appraisals of generative AI: how mindset framing shapes moral judgments, emotional appraisals and privacy behavior. Int J Inf Manage 2025 Oct;84:102940. [doi: 10.1016/j.ijinfomgt.2025.102940]
60. Ricon T. AI and adolescent relationships: bridging emotional intelligence and practical guidance. Comput Hum Behav Rep 2025 Aug;19:100752. [doi: 10.1016/j.chbr.2025.100752]
61. Alanezi F. Assessing the effectiveness of ChatGPT in delivering mental health support: a qualitative study. J Multidiscip Healthc 2024;17:461-471. [doi: 10.2147/JMDH.S447368] [Medline: 38314011]
62. Dakakni D, Safa N. Artificial intelligence in the L2 classroom: implications and challenges on ethics and equity in higher education: a 21st century Pandora's box. Comput Educ Artif Intell 2023;5:100179. [doi: 10.1016/j.caeai.2023.100179]
63. Song I, Pendse SR, Kumar N, De Choudhury M. The typing cure: experiences with large language model chatbots for mental health support. Proc ACM Hum-Comput Interact 2025 Oct 18;9(7):1-29. [doi: 10.1145/3757430]
64. Kong SC, Man-Yin Cheung W, Zhang G. Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds. Comput Educ Artif Intell 2021;2:100026. [doi: 10.1016/j.caeai.2021.100026]
65. Long D, Magerko B. What is AI literacy? competencies and design considerations. 2020 Apr 21 Presented at: CHI '20; Apr 25-30, 2020; Honolulu, HI p. 1-16 URL: https://dl.acm.org/doi/proceedings/10.1145/3313831 [accessed 2026-06-06] [doi: 10.1145/3313831.3376727]

Abbreviations

AI: artificial intelligence
AILS: artificial intelligence self-efficacy

XSL•FO
RenderX

ANA: American Nurses Association
AS: AI literacy
CFI: comparative fit index
CS: correlation stability
EI: emotional perception
GenAI: generative artificial intelligence
RMSEA: root-mean-square error of approximation
SRMR: standardized root-mean-square residual
TLI: Tucker-Lewis index

Edited by J Sarvestan; submitted 23.Dec.2025; peer-reviewed by D Zheng, P Lu; revised version received 17.May.2026; accepted 26.May.2026; published 24.Jun.2026.

Please cite as:
Fan X, Sun J, Zhang L, Wang J, Dong B, Gao X, Jin R, Huai P
The Associations of Emotional Intelligence, AI Self-Efficacy, and AI Literacy Among Nursing Undergraduates Under the NUR.S.E.S. Framework: Network Analysis
JMIR Nursing 2026;9:e90253
URL: https://nursing.jmir.org/2026/1/e90253
doi:10.2196/90253

© XiaoHui Fan, Jingjing Sun, LingHui Zhang, Jinrong Wang, BinHao Dong, XiaoHong Gao, Ruihua Jin, Panpan Huai. Originally published in JMIR Nursing (https://nursing.jmir.org), 24.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Health Literacy and Attitudes Toward Telehealth Use in Practice Among Nursing Students in Saudi Arabia: Cross-Sectional Study

Muna Alharbi[1], RN, PhD; Ahmed S Alharthi[2], RN, BSN, MSc; Abdulrahman A Azab[3], RN, BSN; Renad Alharthi[4], RN, BSN

[1]Community Nursing and Healthcare Department, Faculty of Nursing, Umm Al-Qura University, Al-Abdiyyah Campus, Taif Road, Makkah, Saudi Arabia

[2]East Hawiyah Primary Health Care Centre, Taif Health Cluster, Taif, Saudi Arabia

[3]Hera General Hospital, Makkah Health Cluster, Makkah, Saudi Arabia

[4]Maternity and Children Hospital, Makkah Health Cluster, Makkah, Saudi Arabia

Corresponding Author:
Muna Alharbi, RN, PhD
Community Nursing and Healthcare Department, Faculty of Nursing, Umm Al-Qura University, Al-Abdiyyah Campus, Taif Road, Makkah, Saudi Arabia

Abstract

Background: Nursing students are the future workforce, and their readiness to use digital health tools is important. Previous studies have focused on knowledge and attitudes; however, they have not examined the wide range of digital health literacy levels that may influence nursing students' attitudes toward using telehealth in clinical settings.

Objective: This study aimed to determine the relationship between nursing students' digital health literacy and their attitudes toward telehealth use in practice.

Methods: A cross-sectional design was used. The sample consisted of undergraduate nursing students enrolled in a Bachelor of Nursing program at a selected Saudi Arabian university. The online survey used 2 scales: the Digital Health Care Literacy Scale and the Nurses' Attitudes Toward Use of a Telehealth Scale.

Results: A total of 273 students participated (mean age 21.3, SD 1.9 years). Most of the nursing students demonstrated a high digital health literacy level (n=184, 67.4%; mean Digital Health Care Literacy Scale score 11.9 out of 15, SD 3.1). Digital health literacy was a significant predictor of positive attitudes toward telehealth use in practice (adjusted odds ratio 1.48, 95% CI 1.28-1.71; *P*<.001). Male students were significantly less likely to report positive attitudes than female students (adjusted odds ratio 0.62, 95% CI 0.39-0.97; *P*=.03). However, academic year, telehealth workshops, and informatics courses were not significantly associated with positive attitudes toward telehealth use in practice.

Conclusions: Higher levels of literacy appear to correlate with more positive attitudes toward telehealth use in practice. However, current formal education and workshops had no apparent influence on digital health literacy. This suggests a potential need for strengthening digital training and development in nursing education. This may enhance telehealth readiness and support future digital health care delivery.

(JMIR Nursing 2026;9:e94722) doi:10.2196/94722

KEYWORDS

attitude; digital literacy; nursing education; nursing students; practice; telehealth; telemedicine

Introduction

The World Health Organization defines telehealth as the provision of health care services to individuals at a distance through information and communication technologies. It enables care to be delivered when the provider and the patient are in different physical locations [1]. Telehealth uses communication technologies for diagnosing and treating diseases, supporting research and assessment, and providing continuous professional education [1]. The rapid expansion of telehealth has transformed the health care system nationally and internationally. This expansion has increased the demand for professional competencies in digital health technologies.

Digital health literacy refers to the ability to access, analyze, and apply health information via electronic means to address health-related issues [2,3]. It also includes the ability to apply telehealth services and maintain electronic communication with health care professionals [2,3]. As telehealth becomes

XSL•FO
RenderX

increasingly integrated into health care systems, digital health literacy has become an important competency for nursing students and future health care professionals. Students with stronger digital competencies may engage more effectively with telehealth technologies and adapt more easily to digital health care environments [2,4].

Previous studies have reported moderate to high levels of digital health literacy among nursing students internationally. A study among Indonesian nursing students found that participants possessed fundamental digital health skills. These skills were influenced by factors including higher self-perceived internet skills and more frequent use of the internet for health-related purposes [5]. Similarly, a study from Turkey reported good digital health literacy among nursing students, with higher scores observed among older students and those who reported greater internet use [6]. In contrast, a study of nursing students in Ethiopia found that students had lower digital health literacy levels, particularly among first-year students and those from rural areas [7]. These findings suggest that digital health literacy may vary according to educational exposure, demographic factors, and access to digital technologies [5-7].

Studies conducted in Saudi Arabia have also revealed positive attitudes and moderate knowledge and awareness of telehealth among nursing students [8,9]. This highlights the need for educational programs that strengthen students' digital health competencies [9]. Similar findings were reported in Egypt, where nursing students show positive attitudes toward telehealth despite limited formal training [10,11]. In South Korea, nursing students also had limited direct exposure to telehealth education and clinical telehealth experiences [12]. Nevertheless, telehealth education and self-efficacy were associated with more positive attitudes toward telehealth use in practice [12].

Telehealth has undergone rapid expansion and transformation in Saudi Arabia. In alignment with Vision 2030, the Ministry of Health has launched the E-Health Initiative in Digital Transformation to improve the health care system's effectiveness and efficiency through information technology and digital transformation [13]. According to the Technology Acceptance Model, individual adoption of new technologies is influenced by perceived usefulness and ease of use [14,15]. In the context of telehealth, digital health literacy may contribute to these perceptions by improving individuals' confidence and ability to use digital technologies effectively [2-4].

Although previous studies have investigated nursing students' knowledge and attitudes regarding telehealth, digital health literacy and digital health knowledge are conceptually different and often used interchangeably. Digital health knowledge refers to an individual's understanding of what digital health is; when it is used; and the benefits, limitations, and associated ethical or legal considerations [1,16,17]. Digital health literacy, however, goes beyond awareness to encompass the practical ability to access, use, and assess telehealth services effectively in actual practice [2-4]. Whereas knowledge appears primarily cognitive and content-based, literacy is skill-based and reflects the capacity to apply that knowledge in real-world telehealth interactions [2,3]. Thus, a person may possess knowledge about telehealth without necessarily having the literacy required to use it competently.

Despite increasing integration of telehealth into health care education and practice, there is a lack of evidence regarding the relationship between nursing students' digital health literacy and their attitudes toward telehealth use in practice. Therefore, this study aimed to examine the relationship between digital health literacy and attitudes toward telehealth use in practice among undergraduate nursing students in Saudi Arabia. The study also explored demographic and educational factors associated with positive attitudes toward telehealth use in practice.

Methods

Design

A cross-sectional study was conducted to capture undergraduate nursing students' digital health literacy and attitudes toward telehealth use in practice at a single point in time. This study was reported according to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement guidelines (Checklist 1).

Participants and Setting

This study focused on students enrolled in a Bachelor of Nursing program within a selected university in western Saudi Arabia. In this selected program, telehealth was introduced theoretically to nursing students in many courses; however, limited practical opportunities were available in clinical settings. A convenience sampling method was applied to recruit the participants, which included distribution of invites through social media platforms, emails, and announcements sent through the university's learning management system. To prevent duplicate responses, only 1 submission per device was allowed.

The inclusion criteria for the study were current nursing students in their second year to internship year. First-year students were excluded because they had not yet progressed into the specialized nursing curriculum and had limited exposure to clinical courses, nursing informatics, or telehealth-related learning experiences. Therefore, only second-year, third-year, fourth-year, and internship students were included to ensure that participants had sufficient academic and clinical exposure to meaningfully evaluate digital health literacy and attitudes toward telehealth use in practice. Licensed or registered nurses enrolled in bridging or postlicensure programs were excluded because prior professional experience may influence digital health literacy and attitudes toward telehealth use in practice independently of undergraduate nursing education.

The data were collected from November 2025 to January 2026. During this research period, 440 female and 223 male nursing students were enrolled in the Bachelor of Nursing program. The targeted sample size was 244, which met the required sample size, with a 95% CI and 5% margin of error.

Data Collection Instrument

The survey consisted of 3 sections: part 1 asked about demographic data, while the second and third parts used 2 scales adopted from previously published and publicly accessible

XSL•FO
RenderX

studies [18,19]. The first part of the survey collected demographic data such as sex, age, and academic year (second, third, fourth, or internship). Four questions inquired whether students had attended workshops or lectures related to telehealth, had taken nursing informatics courses, understood the definition of telehealth, and had applied telehealth in clinical placements.

The second part of the survey assessed digital health literacy using the Digital Health Care Literacy Scale (DHLS), with 3 items adopted from a previously published and publicly accessible study [18]. The scale exploratory factor analysis demonstrated good psychometric properties, with Cronbach α of 0.90 and 78% explained variance, supporting the scale's validity and reliability [18]. The responses were measured using a 5-point Likert scale ranging from strongly disagree to strongly agree.

The third part of the survey assessed the students' attitudes toward telehealth in practice using the 19-item Nurses' Attitudes Toward Use of a Telehealth Scale (NATUTS), which was obtained from a previously published and publicly accessible study [19]. NATUTS measured 3 factors: satisfaction, rejection, and development. The instrument demonstrated strong psychometric properties, with Cronbach α values ranging from 0.86 to 0.93 and a total explained variance of 64.4%, supporting its validity and reliability for assessing attitudes toward telehealth [19]. For descriptive interpretations purposes, the items were conceptually grouped under 4 domains, namely perceived benefits, professional integration and reliability, negative emotions and resistance, and engagement and self-development. Responses were recorded using a 5-point Likert scale ranging from strongly disagree to strongly agree.

The survey was administered in English because English is the language of nursing instruction at the study institution, and translation was therefore not needed. The survey was checked for face and content validity to assess its clarity and relevancy to the study objective before distribution. It was sent to 3 experts in nursing education and digital health for feedback. A pilot test was conducted with 10 nursing students to evaluate clarity, readability, and completion time. No modifications were required based on their feedback. Pilot participants were excluded from the final analysis.

The 3 DHLS items were adopted to assess the level of digital health literacy of the students. Each item was rated using the 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The 3 items were rated accordingly, and the total DHLS scores ranged from 3 to 15. The digital health literacy score was classified accordingly: 3 to 7 (low), 8 to 11 (moderate), and 12 to 15 (high). The total NATUTS score was calculated by adding up all 19 items after negatively worded items were reverse-coded to investigate predictors of positive attitudes toward telehealth. The total NATUTS scores ranged from 19 to 95. Higher scores indicated more positive attitudes toward telehealth use in practice. The sample mean score was used as a cutoff to dichotomize the overall attitude score. Mean score dichotomization has been used to facilitate interpretation of logistic regression outcomes when no established cutoff exists. Values below the mean were classified as negative attitudes, and scores equal to or above the mean were classified as positive attitudes. The multivariable logistic regression model was then fitted with the dichotomized variable.

Data Analysis

SPSS (version 28; IBM Corp) was used to perform all statistical analyses. Descriptive statistics were adopted to classify the nursing students' educational background, level of digital health literacy, sociodemographic attributes, and attitude toward telehealth use in practice. Frequencies and percentages denote the categorical variables, and means and SDs denote the continuous variables.

To examine factors associated with positive attitudes toward telehealth use in practice, multivariable logistic regression analysis was performed. Digital health literacy score, age, sex, academic year, exposure to telehealth education, completion of a nursing informatics course, and telehealth use during clinical placement were entered as independent variables in the model. Adjusted odds ratios (AORs) with 95% CIs were calculated. A *P* value of <.05 was considered statistically significant.

Ethical Considerations

The study adhered to ethical research standards and received approval by a relevant institutional review board of Umm Al-Qura University (receipt number HAPO-02-K-021-2025-5-2785). A statement at the beginning of the survey informed participants of the study's objectives, processes, possible benefits, and any related risks. It explained that participants' identities would remain anonymized and protected.

Participants provided consent by completing and submitting the survey anonymously. There were no significant risks, and the only cost to the participants was the time taken to complete the survey. Participation was voluntary, and responses were collected anonymously. No financial compensation or incentives were provided to participants for taking part in the study.

Results

Table 1 presents the sociodemographic and academic characteristics of the participating nursing students (N=273). Most students were aged 21 and 22 years (160/273, 58.6%), followed by 19 and 20 years (82/273, 30%), with a mean age of 21.3 (SD 1.9) years. Most participants were female (209/273, 76.6%). Regarding academic level, most students were in their third year (101/273, 37%), followed by fourth-year students (95/273, 34.8%), nursing interns (45/273, 16.5%), and second-year students (32/273, 11.7%). A high proportion of students reported having attended a telehealth workshop or lecture (198/273, 72.5%) and completing a nursing informatics course (194/273, 71.1%). The vast majority reported knowing the meaning of telehealth (244/273, 89.4%), and more than half reported using telehealth during their clinical placement (148/273, 54.2%).

XSL•FO
RenderX

Table . Sociodemographic and academic characteristics of nursing students (N=273).

Variables	Nursing students
Age (years), n (%)	
19 - 20	82 (30)
21 - 22	160 (58.6)
>23	31 (11.4)
Age (years), mean (SD)	21.3 (1.9)
Sex, n (%)	
Male	64 (23.4)
Female	209 (76.6)
Academic year, n (%)	
Second year	32 (11.7)
Third year	101 (37)
Fourth year	95 (34.8)
Nursing intern	45 (16.5)
Have you attended a workshop or lecture about telehealth?, n (%)	
Yes	198 (72.5)
No	75 (27.5)
Have you taken a nursing informatics course?, n (%)	
Yes	194 (71.1)
No	79 (28.9)
Do you know the meaning of telehealth?, n (%)	
Yes	244 (89.4)
No	29 (10.6)
During your clinical placement, do you use telehealth?, n (%)	
Yes	148 (54.2)
No	125 (45.8)

Table 2 illustrates the distribution of nursing students' responses to the 3-item DHLS. More than half (140/273, 51.3%) strongly agreed that they could use applications or programs such as Zoom independently, while 77 (N=273, 28.2%) agreed. Similarly, 130 (N=273, 47.6%) students strongly agreed and 75 (27.5%) agreed that they could set up video chats without assistance. Regarding problem-solving skills, 91 (N=273, 33.3%) students agreed and 80 (29.3%) strongly agreed that they could resolve basic technical issues on their own, while a smaller proportion (65/273, 23.8%) remained neutral. The overall DHLS score ranged from 3 to 15, with a mean of 11.9 (SD 3.1) out of 15.

XSL•FO
RenderX

Table . Distribution of nursing students' responses to the Digital Health Care Literacy Scale items (N=273).

Digital Health Care Literacy Scale items[a]	Strongly disagree, n (%)	Disagree, n (%)	Neutral, n (%)	Agree, n (%)	Strongly agree, n (%)
I can use applications or programs (such as Zoom) on my cell phone, computer, or another electronic device on my own (without asking for help from someone else).	16 (5.9)	13 (4.8)	27 (9.9)	77 (28.2)	140 (51.3)
I can set up a video chat using my cell phone, computer, or another electronic device on my own (without asking for help from someone else).	15 (5.5)	20 (7.3)	33 (12.1)	75 (27.5)	130 (47.6)
I can solve or figure out how to solve basic technical issues on my own (without asking for help from someone else).	12 (4.4)	25 (9.2)	65 (23.8)	91 (33.3)	80 (29.3)

[a]Range 3-15; mean 11.9, SD 3.1.

The overall distribution of nursing students' digital health literacy levels. Most participants had a high level of digital health literacy (184/273, 67.4%). Exactly 61 (22.3%) students had a moderate level, and a smaller proportion of students (28/273, 10.3%) showed a poor level of digital health literacy.

Table 3 presents the distribution of nursing students' attitudes toward telehealth use in practice. Overall, the findings show a predominantly positive attitude. In the perceived benefits of telehealth domain, most agreed or strongly agreed that telehealth reduces nurses' workload (211/273, 77.3%), is an important part of nursing services (218/273, 79.9%), is essential in nursing (204/273, 74.7%), improves individuals' quality of life (232/273, 85%), simplifies patient-nurse communication (210/273, 76.9%), enables nursing services to be carried out more effectively (209/273, 76.6%), gives professional advantage (202/273, 74%), and eases nurses' workloads (221/273, 80.9%). In the professional integration and reliability domain, 216 (N=273, 79.1%) students agreed or strongly agreed that they would like to include telehealth in professional practice, 173 (63.3%) considered it reliable, and 209 (76.5%) viewed it as essential in nursing services. In the negative emotions and resistance domain, agreement was much lower, indicating limited reluctance: only 61 (22.4%) reported disliking telehealth, 69 (25.3%) felt nervous using it, 67 (24.5%) experienced conflict with its use, 100 (36.7%) used it because they had to, and 71 (26%) would not recommend it to colleagues. Regarding the engagement and self-development domain, over half agreed or strongly agreed that they follow current telehealth developments (135/273, 49.5%), improve themselves regarding telehealth (167/273, 61.2%), and research telehealth (129/273, 47.3%). The overall NATUTS score ranged from 19 to 95, with a mean of 70.93 (SD 11.52) out of 95.

XSL•FO
RenderX

Table . Distribution of nursing students' attitudes toward telehealth use in practice (N=273).

Statements[a]	Strongly disagree and disagree, n (%)	Neutral, n (%)	Strongly agree and agree, n (%)
1. Perceived benefits of telehealth			
I believe that telehealth reduces nurses' workloads.	28 (10.3)	34 (12.5)	211 (77.3)
I believe telehealth is an important part of nursing services.	21 (7.7)	34 (12.5)	218 (79.9)
Telehealth is an essential part of nursing.	27 (9.9)	42 (15.4)	204 (74.7)
I believe that telehealth is effective in improving individuals' quality of life.	19 (7)	22 (8.1)	232 (85)
Telehealth simplifies patient-nurse communication.	22 (8)	41 (15)	210 (76.9)
Telehealth enables nursing services to be carried out more effectively.	23 (8.4)	41 (15)	209 (76.6)
Telehealth gives me an advantage in my profession.	24 (8.8)	47 (17.2)	202 (74)
I believe telehealth eases nurses' workloads.	23 (8.4)	29 (10.6)	221 (81)
2. Professional integration and reliability			
I would like to include telehealth in professional practices.	26 (9.5)	31 (11.4)	216 (79.1)
Telehealth is reliable.	24 (8.8)	76 (27.8)	173 (63.4)
Telehealth is essential in nursing services.	27 (9.9)	37 (13.6)	209 (76.6)
3. Negative emotions and resistance			
I don't like using telehealth.	157 (57.5)	55 (20.1)	61 (22.3)
I get nervous when using telehealth.	138 (50.5)	66 (24.2)	69 (25.3)
I have a conflict with using telehealth.	146 (53.4)	60 (22)	67 (24.5)
I use telehealth because I have to.	78 (28.6)	95 (34.8)	100 (36.6)
I would not recommend telehealth to my colleagues.	156 (57.1)	46 (16.8)	71 (26)
4. Engagement and self-development			
I follow current developments regarding telehealth.	49 (17.9)	89 (32.6)	135 (49.5)
I research telehealth.	57 (20.9)	87 (31.9)	129 (47.3)
I improve myself regarding telehealth.	32 (11.7)	74 (27.1)	167 (61.2)

[a]Mean 70.93, SD 11.52; range 19-95.

Multivariable logistic regression analysis was conducted to determine whether digital health literacy predicted positive attitudes toward telehealth use in practice after controlling for demographic and educational variables. Digital health literacy was a significant predictor of positive attitudes toward telehealth use in practice (AOR 1.48, 95% CI 1.28-1.71; *P*<.001). Male students were significantly less likely to report positive attitudes compared with female students (AOR 0.62, 95% CI 0.39-0.97; *P*=.03). No significant associations were observed for age, academic year, telehealth workshops, informatics courses, or telehealth use during clinical placement (Table 4).

XSL•FO
RenderX

Table . Multivariable logistic regression predicting positive attitudes toward telehealth use in practice (N=273).

Variables	AOR[a] (95% CI)	*P* value[b]
Digital health literacy score	1.48 (1.28 - 1.71)	<.001
Male (vs female)	0.62 (0.39 - 0.97)	.03
Age (years)		
21 and 22 (vs 19 and 20)	0.91 (0.63 - 1.32)	.62
≥23 (vs 19 and 20)	0.74 (0.40 - 1.36)	.33
Academic year		
Third year (vs second year)	1.18 (0.64 - 2.19)	.60
Fourth year (vs second year)	1.32 (0.70 - 2.50)	.39
Intern (vs second year)	1.01 (0.52 - 1.95)	.98
Telehealth workshop (yes vs no)	1.05 (0.66 - 1.67)	.84
Informatics course (yes vs no)	1.12 (0.71 - 1.78)	.62
Telehealth use in clinical placement (yes vs no)	1.09 (0.70 - 1.69)	.69

[a]AOR: adjusted odds ratio.

[b]Significance level: $P<.05$.

Discussion

Principal Findings

The purpose of this study was to examine the relationship between digital health literacy and attitudes toward telehealth use in practice among nursing students. The findings demonstrated that high digital health literacy was significantly associated with positive attitudes toward telehealth use in practice. Students with higher digital health literacy were significantly more likely to perceive telehealth as beneficial, reliable, and relevant for nursing practice.

Most of the nursing students demonstrated a high level of digital health literacy. This result aligns with previous studies that found moderate to high digital health literacy among nursing and health sciences students [20,21]. However, the findings should be interpreted with caution. Although students may possess the ability to access and use digital health information, this does not necessarily indicate preparedness to apply telehealth in complex clinical situations. Previous studies have shown that nursing students often experience difficulties in critically evaluating online health information and in applying digital knowledge in clinical decision-making [20,21]. Thus, possessing digital health literacy skills alone may be insufficient without practical telehealth training and supervised clinical application.

Most of the students stated that telehealth decreases workload, improves the quality of life of patients, and plays a critical role in nursing practice. These results are consistent with those of previous studies demonstrating that nursing students perceive telehealth as beneficial to health care accessibility, efficacy, and communication [19,22]. Positive attitudes toward telehealth use in practice may reflect nursing students' increasing exposure to digital technologies in both educational and health care settings. These attitudes may also be influenced by the rapid expansion of telehealth services following the COVID-19 pandemic [6,20]. Thus, students may perceive telehealth as consistent with contemporary nursing roles which require digital communication and remote patient monitoring [4,22].

Despite the reported positive attitudes, some students exhibited negative feelings such as reluctance and nervousness. These findings are important because they show that positive attitudes can sometimes coexist with discomfort, limited confidence, or perceived external pressure to use digital technologies. This may indicate that some students accept telehealth while still lacking confidence in their practical ability to use it efficiently and effectively. Such findings align with those of previous research, where the nursing students were hesitant due to reliability concerns and insufficient practical exposure [12,23]. These findings suggest that telehealth education should include hands-on simulation and supervised practice experiences to improve confidence and reduce resistance.

The presence of negative emotions toward telehealth use in practice may reflect limitations of current nursing education strategies. Although many students attended telehealth workshops or completed informatics-related courses, these experiences did not lead to more positive attitudes toward telehealth use in practice. This finding suggests that current educational strategies may emphasize theoretical knowledge more than practical competency. Short workshops or isolated informatics courses may increase awareness of telehealth without sufficiently improving students' confidence, communication skills, or ability to integrate telehealth into patient care [23]. Subsequently, students may recognize the importance of telehealth while experiencing anxiety or resistance to its practical implementation.

Digital health literacy independently predicted students' attitudes toward telehealth use in practice even after controlling for demographic and educational variables. This finding suggests that digital health literacy is not only a technical skill but also a determinant of technology acceptance in health care settings.

XSL•FO
RenderX

Students who are confident in using digital tools may perceive telehealth systems as easier to use and thus more useful in clinical care. These results align with the Technology Acceptance Model, which suggests that individuals with higher perceived competence in technology are more likely to adopt digital innovations in practice [14,15]. However, students with lower literacy levels may experience uncertainty and resistance due to difficulties navigating digital health care environments. These findings suggest that digital competencies play a critical role in shaping students' readiness to adopt telehealth technologies in future clinical practice. As telehealth becomes increasingly integrated into health care systems globally, digital health literacy may represent an essential professional competency for future nurses [4]. Therefore, improving digital health literacy may help not only to develop technical competence but also to reduce anxiety and strengthen professional acceptance of telehealth technologies.

The study statistically found that male students were less likely to report positive attitudes toward telehealth use in practice than female students. Although evidence regarding sex differences in attitudes toward telehealth among nursing students remains limited, this result may be related to communication preferences, engagement with patient-centered care, or attitudes toward health care technologies. However, this result should be interpreted cautiously given the lack of literature available and the potential influence of cultural and educational factors. Additional research is needed to better understand the reasons underlying these differences.

Overall, the results imply that telehealth education in nursing programs needs to be further enhanced beyond theoretical instruction. While students generally demonstrated positive attitudes and high digital health literacy, the persistence of nervousness, reluctance, and feelings of obligation indicates gaps between knowledge and practical readiness. Nursing curricula may benefit from integrating experiential telehealth learning activities such as simulations, virtual consultations, case-based learning, and supervised clinical telehealth experiences. These approaches may help students develop greater confidence, reduce resistance, and improve their ability to apply telehealth competently in real health care settings.

To summarize, digital health knowledge reflects understanding, whereas digital health literacy reflects the capacity to apply that understanding in practice. This study contributes to existing evidence that digital health literacy plays an important role in technical competence and is positively associated with students' attitudes toward telehealth use in practice. Nursing education courses need to strengthen digital competencies to reduce resistance and promote more positive perceptions of telehealth within nursing. As telehealth becomes integrated into health care systems worldwide, preparing digitally competent nursing graduates is critical to future workforce readiness.

Limitations

The cross-sectional design and single-educational setting of this study make it difficult to draw conclusions about the causal relationships between students' attitudes toward telehealth use in practice and their level of digital health literacy. It was not possible to ascertain whether exposure to telehealth education results in long-lasting changes in attitudes or abilities because the data were gathered all at once. Additionally, the results are less applicable to other nursing programs with distinct curricula, resources, or student characteristics when the study is limited to a single institution. Consequently, findings should be interpreted cautiously, as they might not be generalizable to the broader population of undergraduate nursing students because of potential self-reported bias.

Conclusions

The findings of this study demonstrate that nursing students have high levels of digital health literacy. In addition, most of the students exhibit positive attitudes toward telehealth use in practice, particularly pertaining to patient care delivery, decreasing workload, and professional integration. The significant finding was the association between higher digital health literacy levels and positive attitudes toward telehealth use in practice. These findings suggest a potential need to strengthen digital health content within nursing curricula to strengthen nurses' digital skills through formal training and continuous professional development.

Educators and clinical leaders can modify training strategies to meet the needs of specific groups. This targeted approach may reduce disparities in digital health and telehealth competencies across the clinical workforce. Offering workshops or informal learning experiences does not guarantee improved competency, so such clinical training should include skills that support telehealth use. Integrating structured, evidence-based digital health literacy curricula into clinical training may ensure that future health care providers develop their competency as technologies evolve.

Acknowledgments

The authors would like to acknowledge the students who provided their time to participate in this research. The authors used generative artificial intelligence (ChatGPT, OpenAI) to assist with language editing and improving the clarity and readability of the manuscript. The final manuscript preparation were performed by the authors, who take full responsibility for the accuracy and integrity of the work.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

XSL•FO
RenderX

Data Availability

The research data will be available upon request from the corresponding author.

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[PDF File, 167 KB - nursing_v9i1e94722_app1.pdf]

References

1. Consolidated telemedicine implementation guide. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240059184 [accessed 2025-11-27]
2. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9. [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
3. Ban S, Kim Y, Seomun G. Digital health literacy: a concept analysis. Digit Health 2024;10:20552076241287894. [doi: 10.1177/20552076241287894] [Medline: 39381807]
4. Longhini J, Rossettini G, Palese A. Digital health competencies among health care professionals: systematic review. J Med Internet Res 2022 Aug 18;24(8):e36414. [doi: 10.2196/36414] [Medline: 35980735]
5. Tarihoran DE, Anggraini D, Juliani E, Ressa R, Fardan I. Indonesian student nurses' e-health literacy skills. Stud Health Technol Inform 2021 Dec 15;284:444-446. [doi: 10.3233/SHTI210767] [Medline: 34920567]
6. Cetinkaya S, Askan F, Gunes NE, Todil T, Yuruk E. E-health literacy of nursing students and investigation of factors affecting e-health literacy during COVID-19 pandemic process: a cross-sectional study. Medicine (Baltimore) 2022 Sep 2;101(35):e30148. [doi: 10.1097/MD.0000000000030148] [Medline: 36107588]
7. Shiferaw KB, Mehari EA, Eshete T. eHealth literacy and internet use among undergraduate nursing students in a resource limited country: a cross-sectional study. Inform Med Unlocked 2020;18:100273. [doi: 10.1016/j.imu.2019.100273]
8. Alshammari A, Alanazi MF, Bahari G. Nursing students' awareness, knowledge, and attitudes regarding telehealth and telenursing use for high-quality healthcare: a cross-sectional study. Nurse Educ Today 2024 Nov;142:106359. [doi: 10.1016/j.nedt.2024.106359] [Medline: 39154592]
9. Bdair IA. Perceptions of pre-licensure nursing students toward telecare and telenursing. Inform Health Soc Care 2024 Jan 2;49(1):42-55. [doi: 10.1080/17538157.2024.2303642] [Medline: 38205799]
10. El- Said Abd Ellatif A, Mohamed Sobhy Elsayed D, Hamido Abosree T. Knowledge and attitude of faculty of nursing students regarding telenursing. J Nurs Sci Benha Univ 2023 Jan;4(1):677-689. [doi: 10.21608/jnsbu.2023.278954]
11. Shehata ZA, Ali SA, Mahmoud SF, Mahmoud DK. Knowledge and attitude of nursing school students regarding telenursing in Sharkia Governorate. Zagazig Nurs J 2024 Jul;20(2):219-235. [doi: 10.21608/znj.2024.380037]
12. Mun M, Choi S, Woo K. Investigating perceptions and attitude toward telenursing among undergraduate nursing students for the future of nursing education: a cross-sectional study. BMC Nurs 2024 Apr 8;23(1):236. [doi: 10.1186/s12912-024-01903-2] [Medline: 38589885]
13. E-health initiative in digital transformation (vision 2030). Ministry of Health, Kingdom of Saudi Arabia. URL: https://www.moh.gov.sa/en/Ministry/nehs/Pages/vision2030.aspx [accessed 2025-11-27]
14. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
15. Venkatesh V, Davis FD. A theoretical extension of the Technology Acceptance Model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
16. Hull SC, Oen-Hsiao JM, Spatz ES. Practical and ethical considerations in telehealth: pitfalls and opportunities. Yale J Biol Med 2022 Sep;95(3):367-370. [Medline: 36187411]
17. Pirtle CJ, Payne KL, Drolet BC. Telehealth: legal and ethical considerations for success. Telehealth Med Today 2019;4:144. [doi: 10.30953/tmt.v4.144]
18. Nelson LA, Pennings JS, Sommer EC, Popescu F, Barkin SL. A 3-item measure of digital health care literacy: development and validation study. JMIR Form Res 2022 Apr 29;6(4):e36043. [doi: 10.2196/36043] [Medline: 35486413]
19. Yılmaz S, Hepokur Yıldırım Ş, Kitiş Y, Kan A. Nurses' attitudes towards telehealth: the development and validation of the nurses' attitudes towards the use of telehealth scale. J Clin Nurs 2025 Jun;34(6):2138-2148. [doi: 10.1111/jocn.17369] [Medline: 39007220]
20. Dadaczynski K, Okan O, Messer M, et al. Digital health literacy and web-based information-seeking behaviors of university students in Germany during the COVID-19 pandemic: cross-sectional survey study. J Med Internet Res 2021 Jan 15;23(1):e24097. [doi: 10.2196/24097] [Medline: 33395396]

XSL•FO
RenderX

21. Sinan O, Ayaz-Alkaya S, Akca A. Predictors of eHealth literacy levels among nursing students: a descriptive and correlational study. Nurse Educ Pract 2023 Mar;68:103592. [doi: 10.1016/j.nepr.2023.103592] [Medline: 36889169]
22. Al Baalharith I, Al Sherim M, Almutairi SH, Albaqami AS. Telehealth and transformation of nursing care in Saudi Arabia: a systematic review. Int J Telemed Appl 2022;2022:8426095. [doi: 10.1155/2022/8426095] [Medline: 36249324]
23. Rettinger L, Putz P, Aichinger L, et al. Telehealth education in allied health care and nursing: web-based cross-sectional survey of students' perceived knowledge, skills, attitudes, and experience. JMIR Med Educ 2024 Mar 21;10:e51112. [doi: 10.2196/51112] [Medline: 38512310]

Abbreviations

AOR: adjusted odds ratio
DHLS: Digital Health Care Literacy Scale
NATUTS: Nurses' Attitudes Toward Use of a Telehealth Scale
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by M Kleib; submitted 05.Mar.2026; peer-reviewed by I Bdair, FA Etindele Sosso; revised version received 20.May.2026; accepted 01.Jun.2026; published 23.Jun.2026.

Please cite as:
Alharbi M, Alharthi AS, Azab AA, Alharthi R
Digital Health Literacy and Attitudes Toward Telehealth Use in Practice Among Nursing Students in Saudi Arabia: Cross-Sectional Study
JMIR Nursing 2026;9:e94722
URL: https://nursing.jmir.org/2026/1/e94722
doi:10.2196/94722

© Muna Alharbi, Ahmed S Alharthi, Abdulrahman A Azab, Renad Alharthi. Originally published in JMIR Nursing (https://nursing.jmir.org), 23.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Leadership Scale for Clinical Nurses: Development and Validation of an Instrument

Ji-Hyeon Lee[1], RN, PhD; Sun-Young Jung[2], RN, PhD

[1]Department of Nursing, Yeungjin University, Daegu, Republic of Korea

[2]Research Institute of Nursing Science, College of Nursing, Daegu Catholic University, 33, Duryugongwon-ro 17-gil, Nam-gu, Daegu, Republic of Korea

Corresponding Author:
Sun-Young Jung, RN, PhD
Research Institute of Nursing Science, College of Nursing, Daegu Catholic University, 33, Duryugongwon-ro 17-gil, Nam-gu, Daegu, Republic of Korea

Abstract

Background: The rapid advancement of digital technologies, combined with the evolving complexity of health care environments, has introduced a new paradigm in nursing practice. Clinical nurses are now required not only to deliver safe and effective patient care but also to demonstrate competencies in digital literacy and innovation. Among these emerging competencies, digital leadership has become a critical attribute—enabling nurses to lead digital transformation, ensure patient safety, enhance care quality, and support system-level change within health care organizations. Despite its increasing relevance, there is a notable absence of validated measurement tools tailored to assess digital leadership in clinical practice.

Objective: This study aimed to develop and psychometrically validate a Digital Leadership Scale for Clinical Nurses (DLS-CN) to systematically evaluate the digital leadership capabilities of nurses working in clinical settings.

Methods: The scale development process followed a rigorous multistep procedure. Initial items were derived from previous qualitative research involving a literature review and in-depth interviews, complemented by an additional literature review conducted in this study. The content validity of 38 preliminary items was evaluated by 9 experts over 2 rounds. A pilot test was conducted with 30 nurses, followed by cognitive interviews with 5 nurses to refine item clarity and relevance. The final set of items was administered to 446 clinical nurses across various health care institutions. Data were randomly split for exploratory factor analysis and confirmatory factor analysis. Additional analyses were conducted to evaluate item discrimination, convergent validity, and internal consistency using IBM SPSS 25.0 and AMOS 23.0.

Results: The finalized DLS-CN consists of 29 items grouped under four domains: (1) ability to use digital technology, (2) digital safety management, (3) digital collaboration mindset, and (4) organizational influence. These 4 factors explained 56.9% of the total variance. The scale showed strong internal consistency (Cronbach α=0.95). Convergent validity was demonstrated through strong positive correlations with the Nursing Informatics Competency Scale (Pearson correlation coefficient r=0.82; P<.001) and the Self-Leadership Scale (Pearson correlation coefficient r=0.83; P<.001).

Conclusions: The DLS-CN is a valid and reliable instrument for measuring digital leadership among clinical nurses. It offers a practical tool for educators, administrators, and researchers to assess and enhance digital leadership capabilities—ultimately supporting the digital transformation of health care systems.

(JMIR Nursing 2026;9:e82101) doi:10.2196/82101

KEYWORDS

nurses; digital technology; leadership; digital health; factor analysis, statistical; nursing informatics; surveys and questionnaires

Introduction

The digital transformation era has accelerated the adoption of digital technologies in the health care sector, particularly after COVID-19, which catalyzed the integration of such technologies into clinical practice [1-3]. Clinical nurses are increasingly using digital tools to ensure safe care environments and to provide evidence-based nursing [4,5]. Common technologies used by nurses include wearable devices, electronic health records, and Barcode Medication Administration systems [6,7], all of which enhance patient safety and efficiency. The adoption of these digital tools in clinical settings has been shown to improve nursing efficiency and patient outcomes [8].

As health care environments become increasingly digitalized, nurses are expected not only to use digital technologies in practice but also to adapt to and support digital change in clinical settings. Nurses serve as coordinators, ensuring effective communication among patients, families, health care

XSL•FO
RenderX

professionals, and nursing peers, which is directly related to patient safety [4,9]. In this context, digital leadership encompasses not only the technical ability to understand and apply digital tools [10] but also the capacity to introduce new communication systems and foster collaboration among team members [11]. Therefore, to successfully implement digitalization and health information technologies within hospital organizations, leadership from nurses who continuously enhance and manage their clinical competencies is essential [12,13]. The level of leadership demonstrated by nurses is closely linked to the quality of nursing services [14].

Digital leadership combines traditional nursing competencies with digital capabilities, thereby improving the quality and efficiency of care delivery [8]. In rapidly evolving health care environments, digital leadership is no longer a personal attribute of individual nurses but a comprehensive competency influencing teams and organizations [15]. Nurses must demonstrate adaptability to digital changes [2], solve problems arising during digital transitions [11,16], and support peer education and motivation [2,17]. Effective digital leadership also includes ethical considerations, such as the protection of patient information in the context of big data usage [18]. Existing tools for measuring digital leadership have primarily been developed for managers or employees in general corporate settings, focusing on behaviors and skills required in digital work environments, including vision sharing and resource support. These tools aim to assess general leadership attributes applicable in digital contexts [19,20]. However, they are limited in that they do not reflect the specific characteristics of nursing practice, which involves unique therapeutic relationships between nurses and patients [21], nor do they account for the evolving nature of contemporary clinical settings due to the introduction of new digital technologies [3]. While digital leadership among nurses—encompassing the ability to adapt to digital transformation and effectively integrate digital technologies with health data—is increasingly essential [2,22,23], there remains a lack of measurement tools that are specifically developed to reflect the unique roles and characteristics of clinical nurses.

The purpose of this study is to develop and validate the Digital Leadership Scale for Clinical Nurses (DLS-CN). This scale is intended to accurately assess the level of digital leadership among clinical nurses and provide a foundation for evaluating and enhancing their leadership capabilities in digital health care environments.

Methods

Study Design

This study is a quantitative methodological research conducted to develop and evaluate a measurement tool for assessing digital leadership among clinical nurses. The development of the DLS-CN was carried out in 2 phases, based on the tool development process proposed by DeVellis and Thorpe [24]. In the first phase, preliminary items were developed, followed by content validity testing, cognitive interviews, and a pilot survey. In the second phase, the tool was administered to clinical nurses to examine its reliability and validity.

Scale Development Process

Phase 1: Preliminary Item Development

Clarification of the Conceptual Definition

This study was based on the hybrid model of Schwartz-Barcott and Kim [25], applying theoretical, fieldwork, and final analytic phases. Based on a previous study analyzing the concept of digital leadership [15], a literature review was conducted, and concept attributes were integrated to establish a conceptual framework and identify core attributes. Through this process, 6 components of digital leadership were identified: digital problem-solving ability, digital ethical sensitivity, digital mindset, self-growth, digital collaboration, and organizational influence.

Item Generation

A total of 104 initial items were generated across 6 components of digital leadership for clinical nurses. Specifically, 21 items measured digital problem-solving ability, 19 items measured digital ethical sensitivity, 16 items measured digital mindset, 9 items measured self-growth, 17 items measured digital collaboration, and 22 items measured organizational influence.

Selection of Measurement Format

Since the items were designed to measure the level of agreement with statements regarding digital leadership among clinical nurses, the widely used Likert scale was deemed appropriate [24]. Therefore, a 5-point Likert scale was selected.

Expert Content Validity Assessment

To evaluate the content validity of the preliminary items, an expert panel of 9 members was formed: 2 clinical nurses with more than 5 years of experience; 2 nurse unit managers; 2 nursing professors; and 1 expert each in digital health care, leadership, and hospital management. After the researchers explained the concept of digital leadership for clinical nurses, the experts rated the appropriateness of the item content and components using a 4-point Likert scale. For item-level evaluation, the item-level content validity index was considered acceptable if it was 0.78 or higher. To assess overall content validity, the scale-level content validity index was calculated by averaging, with 0.90 or higher judged as acceptable [26]. Two rounds of expert validation were conducted. As a result of the first round, 71 items were selected. After the second round, the number was reduced to 46 items. During this process, sentence order and expressions were revised for better clarity and specificity, especially when abstract or ambiguous wording was identified.

Cognitive Interviews

Cognitive interviews were conducted with 5 clinical nurses who met the inclusion criteria of the main study. Each interview lasted 30 to 40 minutes and aimed to understand how participants interpreted and responded to the preliminary items. During the interviews, participants were asked to explain their thought processes while answering each item, including how they understood the wording, whether any terms were unclear, and whether the items were easy or difficult to answer. The interview data were analyzed by reviewing participants' explanations and feedback on item comprehension, wording,

XSL•FO
RenderX

and response processes. Based on these findings, items were revised to improve clarity and reduce potential response errors. When similar or overlapping content was identified, the researchers rechecked whether the same concern was expressed by other interviewees before making revisions. The revised items were then reviewed again with the interviewees to confirm their appropriateness. Some participants questioned the appropriateness of the term "hospital information system," but after reviewing related tools, literature, and expert texts, the term was retained to encompass systems such as electronic medical records and order communication systems. Regarding terms describing digital collaboration tools or systems, alternatives like "intranet" and "groupware" were suggested; however, "digital collaboration system" was found to be the most understandable and suitable. In response to feedback, the term "real-time" was added to better reflect the advantages of system usage. While some general staff nurses noted difficulty adopting new technologies and resources, replacing terms like "utilize" or "apply" was found to weaken alignment with intended attributes, so the original expressions were maintained. After revising the items, the modified 46 items were confirmed for appropriateness through follow-up review with the interviewees.

Pilot Study

To identify issues with the preliminary items, a pilot study was conducted with 30 clinical nurses who met the same criteria as those in the main study. This was conducted via an online survey identical in format to the main study. Respondents rated item comprehensibility on a 5-point scale (1="very difficult" to 5="very easy"), yielding an overall average of 4.14 (SD 0.24), indicating that the items were well understood. The time required to complete the survey ranged from 5 to 25 minutes, with an average of 11 (SD 1.84) minutes. Open-ended questions confirmed that item length, font size, and type were appropriate. Based on the pilot results, items with duplicate meanings or unclear wording were revised. Additionally, a new item was added to the organizational influence component, "I can help organizational members experiencing difficulties in digitalized work environments," to complement the existing item, "I can ask organizational members for help when I face difficulties." All items were revised through linguistic review to ensure correct spacing, terminology, and order. Ultimately, the preliminary item pool was finalized with 47 items across 6 components.

Phase 2: Instrument Evaluation—Reliability and Validity Testing

Participants

Participants were 446 clinical nurses working at tertiary or general hospitals in South Korea. Eligible participants were nurses who were directly involved in patient care and had at least 1 year of clinical experience. Nurses with less than 1 year of experience were excluded because they were considered to be in the novice stage of acquiring the knowledge and skills necessary for nursing practice [27,28]. Nurses in administrative, educational, or nondirect care roles were also excluded. A minimum sample size of 200 is generally recommended for factor analysis, and for confirmatory factor analysis (CFA), a commonly used rule of thumb is that the sample should be at least 10 times the number of estimated parameters [29,30]. Therefore, the total sample size of 446 was considered adequate. To examine the factor structure and cross-validate the model, the total sample was randomly divided into 2 equal groups, with 223 participants assigned to exploratory factor analysis (EFA) and 223 to CFA. This equal split was used to ensure independent and sufficiently large samples for both analyses.

Instruments

To evaluate convergent validity, 2 instruments were selected: the Nursing Informatics Competency Scale [31] and the Self-Leadership Scale [32]. The Nursing Informatics Competency Scale includes 20 items across 5 components: "Basic information and communication technology use" (3 items), "Utilization of nursing information" (5 items), "Professional responsibility and ethics" (5 items), "Use of information and communication technology in nursing" (4 items), and "Attitudes toward nursing informatics" (3 items). Each item is rated on a 4-point Likert scale. The reliability of the original scale was a Cronbach α of 0.91 [31], and in this study, internal consistency reliability was assessed using Cronbach α based on the responses of the study participants, with a Cronbach α coefficient of 0.91, indicating excellent internal consistency. The Self-Leadership Scale consists of 16 items across 4 components: "Collaborative self-dialog strategies" (4 items), "Physical vitality enhancement strategies" (4 items), "Goal-oriented self-training strategies" (4 items), and "Self-respect pursuit strategies" (4 items). This instrument uses a 5-point Likert scale. The original reliability was a Cronbach α of 0.86 [32], and in this study, the Cronbach α was also calculated using the current dataset, yielding a coefficient of 0.91, which indicated excellent internal consistency.

Data Collection

Data were collected from March 10, 2024, to March 30, 2024, through a self-administered online survey using a nonprobability convenience sampling method. Participants were able to complete the survey at their preferred time and location without direct contact with the researchers. To recruit nurses from diverse regions, recruitment notices containing the contact information of the researchers and the survey link (URL) or QR code were posted on online nursing community boards and workplace forums, allowing eligible participants to voluntarily access and complete the survey. All items were set as mandatory, preventing participants from proceeding without responding; consequently, no missing data were observed in the submitted questionnaires. Data collection continued until the predetermined sample size required for item analysis was reached, after which the survey was closed to prevent additional responses. Because the survey link was distributed through open online communities and workplace forums, the total number of individuals who viewed the recruitment notice could not be determined, and thus the response rate could not be calculated.

Data Analysis

Collected data were analyzed using IBM SPSS Statistics version 25.0 and AMOS version 23.0. First, participants' general characteristics were analyzed using frequencies, percentages,

XSL•FO
RenderX

means, and SDs. Second, for the item analysis, means, SDs, skewness, kurtosis, inter-item correlations, item-total correlations, and reliability (if an item was deleted) were calculated. Third, to test the suitability of EFA, the Kaiser-Meyer-Olkin (KMO) measure and the Bartlett test of sphericity were used. Fourth, principal component analysis with varimax rotation was performed to extract factors. Fifth, CFA was conducted to evaluate model fit and test convergent and discriminant validity. Sixth, convergent validity was also examined using the Pearson correlation coefficient. Finally, internal consistency reliability was tested using Cronbach α.

Ethical Considerations

This study was approved by the Institutional Review Board of Daegu Catholic Medical Center (CR-23-148). The online survey introduction clearly explained the study's purpose, informed consent, procedures, voluntary participation, confidentiality, potential benefits and risks, and stated that the data would be used only for research purposes. Participants received a small token of appreciation for their involvement. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki.

Results

General Characteristics of Participants

A total of 446 clinical nurses participated in the survey (Table 1). Among them, 223 participants were included in the EFA, with an average age of 30.71 (SD 5.45) years. Most participants were female (208, 93.3%). Regarding the type of hospital where they worked, 137 participants (61.4%) were employed at tertiary hospitals and 86 (38.6%) at general hospitals. The most common region was Seoul, the capital, with 106 participants (47.5%). The average length of clinical experience was 70.41 (SD 46.68) months. In terms of work departments, 137 participants (61.4%) worked in general wards, 52 (23.3%) in special units (eg, intensive care units, operating rooms, emergency rooms), 17 (7.6%) in outpatient clinics, and 17 (7.6%) in other departments. For the CFA, 223 participants were included, with an average age of 31.07 (SD 4.81) years. The majority were female participants (200, 89.7%). A total of 131 participants (58.7%) worked in tertiary hospitals and 92 (41.3%) in general hospitals. The highest number of participants was also from Seoul (99, 44.4%). The average length of clinical experience was 74.31 (SD 57) months. Regarding work departments, 147 participants (65.9%) worked in general wards, 38 (17.0%) in special units, 25 (11.2%) in outpatient clinics, and 13 (5.8%) in other departments.

XSL•FO
RenderX

Table . General characteristics of participants for factor analysis (N=446).

Characteristics and categories	EFA[a] (n=223)	CFA[b] (n=223)
Age (y), mean (SD)	30.71 (5.45)	31.07 (4.81)
Sex, n (%)		
Female	208 (93.3)	200 (89.7)
Male	15 (6.7)	23 (10.3)
Marital status, n (%)		
Married	161 (72.2)	163 (73.1)
Unmarried	62 (27.8)	60 (26.9)
Educational level, n (%)		
Associate's degree	209 (93.8)	206 (92.4)
Bachelor's degree	13 (5.8)	17 (7.6)
Master's degree	1 (0.4)	0 (0.0)
Hospital type, n (%)		
Tertiary hospital	137 (61.4)	131 (58.7)
General hospital	86 (38.6)	92 (41.3)
Hospital location, n (%)		
	106 (47.5)	99 (44.4)
Capital		
Metropolitan	78 (35.0)	82 (36.8)
Province	39 (17.5)	42 (18.8)
Clinical career (y), mean (SD)	5.87 (3.89)	6.19 (4.75)
Working unit, n (%)		
General ward	137 (61.4)	147 (65.9)
Special department (ICU[c], OR[d], and ER[e])	52 (23.3)	38 (17.0)
OPD[f]	17 (7.6)	25 (11.2)
Other	17 (7.6)	13 (5.8)
Experience related to digital leadership education, n (%)		
Yes	13 (5.8)	7 (3.1)
No	210 (94.2)	216 (96.9)
Necessity of digital leadership education, mean (SD)	4.00 (0.89)	4.01 (0.82)

[a]EFA: exploratory factor analysis.

[b]CFA: confirmatory factor analysis.

[c]ICU: intensive care unit.

[d]OR: operating room.

[e]ER: emergency room.

[f]OPD: outpatient department.

Item Analysis

To examine the distribution of items in the 223 collected responses for factor analysis, the mean, SD, skewness, and kurtosis were calculated for each item. Since none of the items had an absolute skewness value greater than 3.0 or an absolute kurtosis value greater than 8.0, all items met the assumption of normality [33]. Item-total correlation coefficients were reviewed, and items with coefficients below 0.30 or above 0.80 were removed, as these values suggest low contribution to the scale or item redundancy [34]. Seventeen items were removed due to low internal consistency reliability and weak correlation between item pairs. Based on corrected item-total correlation and changes in reliability after item deletion [33], the corrected item-total correlations ranged from 0.43 to 0.72, and the reliability did not significantly change upon item deletion (Table 2). As a result, 30 items were retained for further analysis.

XSL•FO
RenderX

Table . Goodness-of-fit test for the measurement model (n=223).

Goodness-of-fit	*P* value	Chi-square (*df*)	χ^2/df[a]	CFI[b]	RMSEA[c]	SRMR[d]
Measurement model	<.001	661.64 (371)	1.78	0.91	0.06	0.5
Criteria	<.05	__[e]	<3	≥0.90	≤0.06	≤0.08

[a]χ^2/df: normed chi-square.

[b]CFI: comparative fit index.

[c]RMSEA: root mean square error of approximation.

[d]SRMR: standardized root mean squared residual.

[e]Not available.

Exploratory Factor Analysis

EFA using principal component analysis with varimax rotation was conducted on the 30 items of the DLS-CN. Prior to the factor analysis, data adequacy was assessed using the KMO measure and the Bartlett test of sphericity. The KMO value was 0.96, indicating excellent sampling adequacy, and the Bartlett test of sphericity was statistically significant [35], confirming the data's suitability for factor analysis. To determine the number of factors, eigenvalues greater than 1.0 and a cumulative variance explanation rate exceeding 60% were used based on Kaiser rule. Scree plot analysis and the interpretability of the factor structure were also considered. Items were retained based on communalities greater than 0.40, factor loadings above 0.40, and at least 3 items per factor [36]. One item with a factor loading below 0.40 was removed. The final analysis included 29 items, with a KMO of 0.95 and a statistically significant Bartlett test (χ^2_{820}=4799.3; *P*<.001), indicating the appropriateness of factor analysis. The EFA extracted 4 factors with eigenvalues greater than 1.0, explaining a cumulative variance of 56.9%. Communality values ranged from 0.52 to 0.64, and all factor loadings exceeded 0.40, indicating good model fit. The purpose of factor analysis is to understand the pattern of data and clarify the interpretability of factors [37]. The factor loadings and item contents were examined to ensure conceptual coherence. Cross-loading items that were theoretically essential were assigned to conceptually valid factors. The inter-item correlation coefficients within subfactors were as follows: factor 1: 0.43 to 0.64; factor 2: 0.35 to 0.57; factor 3: 0.35 to 0.51; and factor 4: 0.39 to 0.052, indicating that the model estimation was appropriate (Table 3).

XSL•FO
RenderX

Table . Communality and pattern matrix of the final exploratory factor analysis (n=223)[a].

Item number	Communality	Factor loading			
		F1[b]	F2[c]	F3[d]	F4[e]
1	0.58	0.63	0.29	0.30	0.12
46	0.64	0.63	0.05	0.25	0.42
39	0.60	0.62	0.17	0.30	0.31
8	0.61	0.61	0.44	0.16	0.12
9	0.54	0.60	0.32	0.21	0.17
2	0.62	0.60	0.34	0.35	0.18
40	0.63	0.59	0.14	0.24	0.45
45	0.52	0.52	0.33	0.25	0.25
4	0.57	0.49	0.43	0.33	0.21
22	0.66	0.24	0.68	0.20	0.31
14	0.60	0.29	0.68	0.21	0.13
16	0.60	0.31	0.66	0.13	0.24
26	0.58	0.16	0.65	0.29	0.22
30	0.52	0.10	0.55	0.33	0.31
36	0.52	0.30	0.48	0.45	0.06
31	0.62	0.03	0.29	0.62	0.39
33	0.53	0.31	0.20	0.61	0.17
37	0.53	0.40	0.11	0.58	0.14
34	0.54	0.29	0.22	0.58	0.27
20	0.54	0.32	0.34	0.56	0.05
21	0.56	0.20	0.28	0.54	0.39
25	0.54	0.35	0.25	0.49	0.33
19	0.58	0.15	0.27	0.27	0.64
12	0.56	0.26	0.29	0.15	0.63
41	0.59	0.27	0.06	0.42	0.58
18	0.57	0.22	0.46	0.18	0.53
27	0.53	0.20	0.44	0.18	0.52
42	0.52	0.44	0.12	0.28	0.48
38	0.54	0.44	0.36	0.00	0.47
Eigen value	—[f]	13.16	1.26	1.10	1.01
Variance (%)	—	16.31	14.80	13.23	12.60
Cumulative variance (%)	—	16.31	31.11	44.34	56.94

[a]Kaiser-Meyer-Olkin (KMO) value is 0.95, and the Bartlett test of sphericity was significant (χ^2_{406}=3492.66; *P*<.001).

[b]F1: ability to use digital technology.

[c]F2: digital safety management.

[d]F3: digital collaboration mindset.

[e]F4: organizational influence.

[f]Not applicable.

XSL•FO
RenderX

Confirmatory Factor Analysis

CFA was conducted to verify whether the 4 factors and 29 items extracted during the EFA appropriately reflected the constructs of digital leadership among clinical nurses. The model's goodness-of-fit was assessed to confirm the adequacy of the structure. To evaluate the model fit, the chi-square statistic was examined. A model was considered acceptable if the normed chi-square χ^2/df was less than 3, the comparative fit index was above 0.90, the root mean square error of approximation was below 0.06, and the standardized root mean squared residual was below 0.80 [38]. The results of the CFA indicated that the chi-square value was statistically significant, and the normed chi-square χ^2/df was 1.8, satisfying the criterion of being below 3. The comparative fit index was 0.91, exceeding the threshold of 0.90. The root mean square error of approximation was 0.06, which meets the criterion for a good model fit. The standardized root mean squared residual was 0.05, which is below the acceptable cutoff of 0.08, indicating that the model had a generally good fit (Table 2).

To verify the convergent validity of each factor, the standardized factor loading was required to be 0.5 or higher, the significance of the unstandardized regression coefficient (critical ratio) was required to exceed 1.97 (P<.05), the construct reliability had to be 0.7 or higher, and the average variance extracted (AVE) needed to be 0.5 or higher [39]. The results of the convergent validity test showed that the critical ratio values of the unstandardized regression coefficients ranged from 7.83 to 10.63, significantly exceeding the threshold [30], confirming that the measured items met the conditions necessary for validity testing. The standardized factor loadings for all factors met the criterion of 0.5 or higher. The AVE values for each factor ranged from 0.52 to 0.57, satisfying the minimum requirement of 0.50. The construct reliability values ranged from 0.85 to 0.90, all exceeding the minimum requirement of 0.70. Therefore, the DLS-CN demonstrated adequate convergent validity in terms of item construction. Through this, it was confirmed that 4 factors consistently measured the construct of "Digital Leadership Scale for Clinical Nurses," and that the 29 items consistently represented the corresponding factors. To assess discriminant validity, interconstruct correlation matrices were examined (Table 4). Discriminant validity was evaluated by checking whether the AVE of each factor was greater than the squared correlation coefficients between the factors [39]. In some cases, such as the ability to use digital technology–digital collaboration mindset, ability to use digital technology–organizational influence, digital safety management–digital collaboration mindset, and digital collaboration mindset–organizational influence, the squared correlations between factors exceeded the AVE of the individual constructs, indicating that discriminant validity was only partially satisfied.

Table . Correlations between variables and verification of construct validity (n=223).

Variables	Squared correlation (P value)				AVE[e]	CR[f]
	F1[a]	F2[b]	F3[c]	F4[d]		
F1	1	—[g]	—	—	0.57	0.91
F2	0.49 (<.001)	1	—	—	0.54	0.85
F3	0.66 (<.001)	0.53 (<.001)	1	—	0.52	0.90
F4	0.71 (<.001)	0.51 (<.001)	0.68 (<.001)	1	0.52	0.90

[a]F1: ability to use digital technology.

[b]F2: digital safety management.

[c]F3: digital collaboration mindset.

[d]F4: organizational influence.

[e]AVE: average variance extracted.

[f]CR: construct reliability.

[g]Not applicable.

Convergent Validity Testing

To verify convergent validity, the correlations with the Nursing Informatics Competency Measurement Tool [31] and the Self-Leadership Scale [32] were examined. The correlation with the Nursing Informatics Competency Scale was 0.82 (P<.001), and the correlation with the Self-Leadership Scale was 0.83 (P<.001). Since the correlation coefficients were all above 0.70, convergent validity was confirmed.

Reliability Testing

Reliability was tested using Cronbach α, with a criterion of 0.70 or higher [29,40]. Cronbach α for the 29 items measuring DLS-CN was 0.95.

By subfactor, Cronbach α was as follows: factor 1 (ability to use digital technology): 0.88; factor 2 (digital safety management): 0.78; factor 3 (digital collaboration mindset): 0.86; and factor 4 (organizational influence): 0.87.

Since Cronbach α for all subfactors and the overall tool exceeded 0.70, reliability was confirmed.

Final Items

In this study, the final version of the DLS-CN consisted of 29 items across 4 factors (Table 5).

By subfactor, the items were as follows: 8 items for the ability to use digital technology, 5 items for digital safety management,

XSL•FO
RenderX

8 items for a digital collaboration mindset, and 8 items for organizational influence.

XSL•FO
RenderX

Table . Final items of the convergence of Digital Leadership Scale for Clinical Nurses.

Factors and items	Item number
Ability to use digital technology	
I can provide patient-centered care by using digital technology.	1
I can explore ways to improve digital systems and processes for patient care and treatment.	46
I can explain the importance of actively using digital technology.	39
I can actively use new information obtained through digital technology to solve patients' health problems.	8
I can present new information obtained through digital technology to address issues in the nursing work environment.	9
I can systematically collect data on patients' health status and nursing needs using digital technology.	2
I can argue for the necessity of applicable resources (human, systems, equipment, etc) in the digitalizing clinical environment.	45
I understand the principles of digital technology necessary for nursing tasks and can apply them to patients.	4
Digital safety management	
I can learn methods for using and managing new digital technologies applicable to patient care.	22
I can maintain a secure environment for protecting patient privacy in the digitalizing clinical setting (eg, logging out of computers when changing locations).	14
I perform nursing tasks in accordance with organizational guidelines and procedures for managing personal information in the digital clinical environment.	16
I believe self-development is important for improving digital knowledge and technical skills related to nursing tasks.	26
I can request help from organizational members when encountering difficulties in the digital clinical environment.	30
Digital collaboration mindset	
I can assist organizational members who are struggling in the digital clinical environment.	31
I can logically express my opinions so that organizational members can understand them using digital collaboration systems (eg, EMR[a], PACS[b], in-hospital messenger, etc).	33
I feel comfortable exchanging real-time opinions through digital collaboration systems (eg, EMR, PACS, in-hospital messenger, etc).	37
I can understand organizational members' opinions easily when communicating via digital collaboration systems (eg, EMR, PACS, in-hospital messenger, etc).	34
I am not afraid of adapting to digital changes in the clinical setting.	20
If I do not know how to use a new digital technology, I can resolve the issue by using available resources.	21
I can proactively use digital technology to learn the latest nursing knowledge and skills.	25
I can accurately hand over information and nursing processes related to patients using digital collaboration systems (eg, EMR, PACS, in-hospital messenger, etc).	36
Organizational influence	
I am interested in new digital changes that benefit patient care.	19
I reflect on my values and norms that may affect patient privacy in the digital clinical environment.	12
I can encourage organizational members who attempt change by learning new digital technologies.	41
I understand the importance of using new digital technologies for advancing the nursing field.	18
I feel a sense of accomplishment when I effectively perform nursing tasks using newly acquired digital technology.	27
I can help improve organizational members' digital competencies for better work performance.	42
I strive to achieve organizational goals by introducing new digital technologies and resources.	38

XSL•FO
RenderX

Factors and items	Item number
I can present a vision (organizational future image) for performing nursing tasks using digital technology.	40

[a]EMR: electronic medical record.

[b]PCA: principal component analysis.

Discussion

Principal Findings

This study was conducted to develop a tool to measure digital leadership among clinical nurses, based on the instrument development steps proposed by DeVellis and Thorpe [24]. The research was carried out in 2 phases: the development of preliminary items and the evaluation of the instrument's validity and reliability.

During the item development phase, a conceptual framework was established by integrating the components of digital leadership identified in previous concept analyses and literature reviews on digital leadership in clinical nursing. Preliminary items were selected from the initial items through a content validity assessment by experts, followed by cognitive interviews and a pilot survey, which were used to derive the final set of preliminary items. Throughout this process, expert feedback and identified issues were systematically reviewed and refined.

In the instrument evaluation phase, data were collected from 446 clinical nurses through an online survey. A random sample of 223 responses was used for EFA, and another 223 responses were used for CFA. Construct validity was tested through these analyses, and convergent validity was assessed by examining the correlation with instruments measuring nursing informatics competency and self-leadership. Internal consistency reliability was confirmed with a Cronbach α of 0.95.

The final tool developed through this study consists of 29 items across 4 factors: 8 items for the ability to use digital technology, 5 items for digital safety management, 8 items for a digital collaboration mindset, and 8 items for organizational influence. The tool uses a 5-point Likert scale for self-report surveys, with a total score range from 29 to 145 points. Higher scores indicate a higher level of digital leadership among clinical nurses. Overall, the developed scale demonstrated strong reliability and validity for assessing digital leadership in clinical nursing practice. However, the tool did not fully satisfy the criteria for discriminant validity. Even when variables are used as conceptually independent constructs, correlations between related factors are common in practice [29]. This is consistent with findings from a study on the development of an integrated nursing leadership scale for Korean nurses [41], which also reported overlaps among constructs such as data management and communication technologies. In particular, relatively high correlations were observed between the ability to use digital technology and organizational influence, as well as between digital collaboration mindset and organizational influence. These results suggest that while the factors are conceptually distinct, they may be closely related in actual clinical settings where digital competencies are integrated with leadership behaviors. Although some aspects of discriminant validity were supported, these findings indicate that the boundaries between certain factors were not clearly distinguished. Despite these limitations, this tool has the advantage of including items that measure competencies needed to adapt to rapidly changing digital health care environments, such as work performance, learning, and adaptability [42]. Accordingly, future research is needed to more clearly distinguish the unique conceptual domains measured by each factor. Furthermore, the "digital safety management" factor included fewer items compared to other factors. Given the importance of patient safety and information security in digital health care environments, future research should consider expanding and refining items in this domain to enhance the comprehensiveness of the scale.

Given that leadership competencies required in digitally transforming health care environments are increasingly complex and diverse [13], this DLS-CN reflects the unique characteristics of Korean clinical settings and is the first scale of its kind to undergo statistical validation. Nevertheless, because data collection was conducted via an online survey using convenience sampling, caution is needed when interpreting the results. Additionally, the scale was developed for general clinical nurses and excluded nurses with less than 1 year of experience, nurse managers not involved in direct care, and those in administrative or educational roles, thereby limiting its generalizability. Moreover, the study did not include test-retest reliability, which represents another limitation.

Based on these findings, several implications can be drawn. In particular, the developed digital leadership scale should be applied to nurses from diverse countries and cultural backgrounds, as well as to nurses in various roles and settings, such as hospitals, community health centers, and educational institutions. Repeated studies are recommended to verify the reliability and validity of the instrument in diverse environments.

Future research should use this tool to assess digital leadership levels and develop tailored educational or intervention programs that enhance digital leadership according to workplace and contextual characteristics.

In addition, as this scale may be applicable not only to health care professionals but also to professionals in management, public administration, and related fields, further research should explore its broader applicability across disciplines.

Conclusions

The DLS-CN, developed through this study, is a systematic tool designed to measure and evaluate digital leadership among nurses in clinical practice settings. The DLS-CN consists of 4 core domains: ability to use digital technology, digital safety management, digital collaboration mindset, and organizational influence. It proposes a new standard for nursing leadership in health care environments driven by digital technologies.

XSL•FO
RenderX

In particular, the tool is useful for developing various educational programs aimed at identifying emerging nurse leaders and enhancing the digital leadership of clinical nurses in hospital environments where frequent changes in digital health care systems, technologies, and organizational structures occur. Furthermore, the development of this tool confirms that digital leadership extends beyond the mere acquisition of digital skills, encompassing the competence of clinical nurses to actively engage in and lead change. This is a meaningful contribution to the study.

Since this tool was developed based on data collected from clinical nurses in South Korea, and the items were created accordingly, further validation studies are needed in diverse health care settings and across cultural contexts. In the rapidly evolving digital health care environment, this tool can serve both as a foundational resource and a structured instrument for strengthening clinical nurses' digital leadership.

Acknowledgments

The text and tables in this paper were entirely generated by the authors without any contributions from generative artificial intelligence (AI). However, AI tools were used to assist in translating an author's original work into a secondary language for inclusion in the manuscript. This paper is based on the doctoral dissertation of JHL, titled *Development of a Digital Leadership Scale for Clinical Nurses*, submitted to Daegu Catholic University in 2024. Some parts of this manuscript have been revised and extracted from the original dissertation.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

References

1. KIM YM. Digital competencies in the era of digital transformation. Local Inform Mag 2020;125:50-55 [FREE Full text]
2. Martins H. Digital transformation and digital leadership. Healthc Inform Res 2019 Oct;25(4):350-351. [doi: 10.4258/hir.2019.25.4.350] [Medline: 31777680]
3. Fletcher M, Read C, D-Adderio L. Nurse leadership post COVID pandemic-a framework for digital healthcare innovation and transformation. SAGE Open Nurs 2023;9:23779608231160465. [doi: 10.1177/23779608231160465] [Medline: 36895712]
4. Oh EG. Perspectives on nursing profession for a post-COVID-19 new normal. Korean J Adult Nurs 2020;32(3):221. [doi: 10.7475/kjan.2020.32.3.221]
5. Reid L, Button D, Brommeyer M. Challenging the myth of the digital native: a narrative review. Nurs Rep 2023 Apr 4;13(2):573-600. [doi: 10.3390/nursrep13020052] [Medline: 37092480]
6. Kim KJ, Jang BY, Jung JY, Park OW. The coming of the 4th industrial revolution and the HRD issues for nurses—prospects and challenges [Article in Korean]. Korean J Hum Resour Dev 2018 Sep 30;21(3):137-159. [doi: 10.24991/KJHRD.2018.09.21.3.137]
7. Grailey K, Hussain R, Wylleman E, Ezzat A, Huf S, Franklin BD. Understanding the facilitators and barriers to barcode medication administration by nursing staff using behavioural science frameworks. A mixed methods study. BMC Nurs 2023 Oct 12;22(1):378. [doi: 10.1186/s12912-023-01382-x] [Medline: 37828520]
8. Brown J, Pope N, Bosco AM, Mason J, Morgan A. Issues affecting nurses' capability to use digital technology at work: an integrative review. J Clin Nurs 2020 Aug;29(15-16):2801-2819. [doi: 10.1111/jocn.15321] [Medline: 32416029]
9. Hwang W, Ha J, Park D. A survey on situation-related communication educational needs for novice intensive care unit nurses [Article in Korean]. J Korean Crit Care Nurs 2024 Feb;17(1):17-29. [doi: 10.34250/jkccn.2024.17.1.17]
10. Hensellek S. Digital leadership: a framework for successful leadership in the digital age. J Media Manag Entrep 2020;2(1):55-69. [doi: 10.4018/JMME.2020010104]
11. Cortellazzo L, Bruni E, Zampieri R. The role of leadership in a digitalized world: a review. Front Psychol 2019;10:1938. [doi: 10.3389/fpsyg.2019.01938] [Medline: 31507494]
12. Hughes R, Meadows MT, Begley R. AONL nurse leader competencies: core competencies for nurse leadership. Nurse Lead 2022 Oct;20(5):437-443. [doi: 10.1016/j.mnl.2022.08.005]
13. Laukka E, Pölkki T, Heponiemi T, Kaihlanen AM, Kanste O. Leadership in digital health services: protocol for a concept analysis. JMIR Res Protoc 2021 Feb 4;10(2):e25495. [doi: 10.2196/25495] [Medline: 33538702]
14. Kim MS, Shin SH, Seo E. The influences of self-leadership, role conflict, nursing work environment on nursing service quality in comprehensive nursing care service wards. J East-West Nurs Res 2021;27(1):22-31. [doi: 10.14370/jewnr.2021.27.1.22]

XSL•FO
RenderX

15. Lee JH, Jung SY. Concept analysis of digital leadership of clinical nurses. J Korea Acad-Ind Cooper Soc 2024 Mar 31;25(3):585-597. [doi: 10.5762/KAIS.2024.25.3.585]
16. Burgess JM, Honey M. Nurse leaders enabling nurses to adopt digital health: results of an integrative literature review. Nurs Prax Aotearoa N Z 2022;38(3):1-13. [doi: 10.36951/001c.40333]
17. De Leeuw JA, Woltjer H, Kool RB. Identification of factors influencing the adoption of health information technology by nurses who are digitally lagging: in-depth interview study. J Med Internet Res 2020 Aug 14;22(8):e15630. [doi: 10.2196/15630] [Medline: 32663142]
18. Laukka E, Hammarén M, Pölkki T, Kanste O. Hospital nurse leaders' experiences with digital technologies: a qualitative descriptive study. J Adv Nurs 2023 Jan;79(1):297-308. [doi: 10.1111/jan.15481] [Medline: 36300725]
19. Büyükbeşe T, Di˙kbaş T, Klei˙n M, Batuk Ünlü S. A study on digital leadership scale (DLS) development. Kahramanmaraş Sütçü İmam Üniversitesi Sosyal Bilimler Dergisi 2022;19(2):740-760. [doi: 10.33437/ksusbd.1135540]
20. Claassen K, Dos Anjos DR, Kettschau J, Broding HC. How to evaluate digital leadership: a cross-sectional study. J Occup Med Toxicol 2021 Oct 1;16(1):44. [doi: 10.1186/s12995-021-00335-x] [Medline: 34598724]
21. Heo ML, Im SB. Development of the patient caring communication Scale [Article in Korean]. J Korean Acad Nurs 2019 Feb;49(1):80-91. [doi: 10.4040/jkan.2019.49.1.80] [Medline: 30837445]
22. Ćwiklicki M, Klich J, Chen J. The adaptiveness of the healthcare system to the fourth industrial revolution: a preliminary analysis. Futures 2020 Sep;122:102602. [doi: 10.1016/j.futures.2020.102602]
23. Tiase VL, Henderson MD. The future of nursing in a digital age: planning for rapid change. In: Accessibility InformationBookNursing and Informatics for the 21st Century - Embracing a Digital World, 3rd Edition, Book 4, 1st edition: Productivity Press; 2022:171-186. [doi: 10.4324/9781003281047-10]
24. DeVellis RF, Thorpe CT. Scale Development: Theory and Applications, 5th edition: Sage Publications; 2021.
25. Schwartz-Barcott D, Kim HS. An expansion and elaboration of the hybrid model of concept development. In: Rodgers BL, Knafl KA, editors. Concept Development in Nursing: Foundations, Techniques, and Applications, 2nd edition: WB Saunders Company; 2000:129-159.
26. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health 2007 Aug;30(4):459-467. [doi: 10.1002/nur.20199] [Medline: 17654487]
27. Cho MS, Cho Y, Kim KH, Kwon IG, Kim MS, Lee JL. Development of clinical ladder system model for nurses: for tertiary care hospitals [Article in Korean]. J Korean Clin Nurs Res 2015;21(3):277-292. [doi: 10.22650/JKCNR.2015.21.3.277]
28. Cho MS, Kwon IG, Kim KH, Kim MS, Cho Y. Validity and applicability of clinical ladder system model for nurses [Article in Korean]. J Korean Clin Nurs Res 2017;23(3):281-292. [doi: 10.22650/JKCNR.2017.23.3.281]
29. Tak JK. Psychological Testing: An Understanding of Development and Evaluation Method: Hakjisa; 2007.
30. Woo JP. Concepts and Understandings of Structural Equation Models: Amos 4.0–20.0: Hannarae Publishing; 2012.
31. Jang SM, Kim J. Development of nursing informatics competence scale for Korean clinical nurses. Comput Inform Nurs 2022 Oct 1;40(10):725-733. [doi: 10.1097/CIN.0000000000000934] [Medline: 35524364]
32. Lee H. Development of a Self-Leadership Scale for Clinical Nurses (SLSN): Chonnam Univ; 2023.
33. Kline RB. Principles and Practice of Structural Equation Modeling, 4th edition: Guilford Press; 2015.
34. Lee EO, Im NY, Park HA, et al. Nursing Research and Statistical Analysis: Soomoonsa; 2009.
35. Kaiser HF. An index of factorial simplicity. Psychometrika 1974 Mar;39(1):31-36. [doi: 10.1007/BF02291575]
36. Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract Assess Res Eval 2005;10(7):1-9. [doi: 10.7275/jyj1-4868]
37. Beavers AS, Lounsbury JW, Richards JK, Huck SW, Skolits GJ, Esquivel SL. Practical considerations for using exploratory factor analysis in educational research. Pract Assess Res Eval 2013;18(6):1-13. [doi: 10.7275/qv2q-rk76]
38. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999;6(1):1-55. [doi: 10.1080/10705519909540118]
39. Noh KS. The Proper Methods of Statistical Analysis for Dissertation: SPSS & AMOS [Book in Korean]: Hanbit Academy; 2019.
40. Kim JW, Oh SW, Hong CH, et al. Clinical Guide to Psychological Assessment, 2nd edition: Hakjisa; 2014.
41. Cho SM, Choi J. Integral leadership in nursing: development and psychometric validation of a Korean version of the Integral Nursing Leadership Scale. Int J Nurs Stud 2024 Apr;152:104697. [doi: 10.1016/j.ijnurstu.2024.104697] [Medline: 38295669]
42. Rosser E, Westcott L, Ali PA, et al. The need for visible nursing leadership during COVID-19. J Nurs Scholarsh 2020 Sep;52(5):459-461. [doi: 10.1111/jnu.12587] [Medline: 32779857]

Abbreviations

AVE: average variance extracted
CFA: confirmatory factor analysis
DLS-CN: Digital Leadership Scale for Clinical Nurses
EFA: exploratory factor analysis
KMO: Kaiser-Meyer-Olkin

XSL•FO
RenderX

Edited by T Risling; submitted 09.Aug.2025; peer-reviewed by H Kim, H Jeon, N Akbar; revised version received 21.Apr.2026; accepted 26.May.2026; published 22.Jun.2026.

Please cite as:
Lee JH, Jung SY
Digital Leadership Scale for Clinical Nurses: Development and Validation of an Instrument
JMIR Nursing 2026;9:e82101
URL: https://nursing.jmir.org/2026/1/e82101
doi:10.2196/82101

© Ji-Hyeon Lee, Sun-Young Jung. Originally published in JMIR Nursing (https://nursing.jmir.org), 22.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing Digital Literacy Among Nursing Faculty Under the Artificial Intelligence–Enhanced Technological Pedagogical Content Knowledge Framework: Cross-Sectional Survey Analysis

Junyu Zhao[1*], RN, PhD; Xiaomin Hua[1*], BSc, RN, MSc; Rong Hu[1], RN, PhD; Chu Chen[2], RN, PhD; Wanxin Wu[1], BSCN, RN, MSc; Yaming Zhang[1], BSCN, RN, MSc; Yuhe Zhang[1], BSCN, RN, MSc; Aijing Wu[1], BSCN, RN, MSc; Yu Zhang[1], BSCN, RN, MSc

[1]School of Nursing, Fujian Medical University, No.1 North Xuefu Rd, Shangjie Town, Fuzhou, Fujian, China

[2]School of Health Management, Fujian Medical University, 1 Xuefu Road, Fuzhou, Fujian, China

[*]these authors contributed equally

Corresponding Author:
Rong Hu, RN, PhD
School of Nursing, Fujian Medical University, No.1 North Xuefu Rd, Shangjie Town, Fuzhou, Fujian, China

Abstract

Background: Digital literacy has become a core competence for nursing faculty to adapt to the digital transformation of education, improve teaching quality, and cultivate innovative nursing talents.

Objective: This study aimed to investigate the digital literacy status and influencing factors among nursing faculty in colleges and universities in Fujian province, China.

Methods: A cross-sectional survey analysis was conducted among 339 participants from July 2023 to June 2024 by selecting nursing faculty from 3 medical universities and 5 colleges in Fujian province. A general information questionnaire and a teacher digital literacy questionnaire (designed based on the artificial intelligence–enhanced technological pedagogical content knowledge theoretical model) were used to conduct the survey. The questionnaires included the following dimensions: an intelligent integration ethics layer (digital social responsibility), attention to the moral boundaries of applying technology; an awareness layer (digital awareness), the willingness to actively adapt to technological changes; a knowledge layer (digital technology knowledge and skills), mastery of the integration point of intelligent tools and subject teaching; an ability layer (digital application), the ability to design intelligent teaching plans; and a thinking layer (professional development; innovative thinking for the critical integration of technology). Statistical analysis included descriptive statistics, a 1-way ANOVA, Pearson correlation analysis, and multiple linear regression (using SPSS 28.0 software).

Results: The overall digital literacy of the nursing faculty was at a medium level (mean score 101.92, SD 16.47), with the highest score in digital awareness (mean 20.21, SD 9.43) and the lowest score in digital technology knowledge and skills (mean 9.68, SD 2.92). For the dimension of digital technology knowledge and skills, both age (β=−.142; *P*=.009) and years of teaching (β=−.147; *P*=.006) were significant negative predictors. Regarding digital application, age was found to be a significant negative predictor (β=−0.124; *P*=.02) but teaching experience was a positive predictor (β=0.123, *P*=.02). Similarly, age also had a significant negative impact on the professional development dimension (β=−.153; *P*=.005) and the overall digital literacy level (β=−.136; *P*=.01).

Conclusions: The digital literacy of nursing faculty is at a medium level, with especially low scores in technical knowledge. Age and teaching experience are key influencing factors. Recommendations to promote balanced development of awareness and skills include providing targeted training for senior faculty, integrating courses with clinical practice, optimizing the allocation of digital resources, promoting interdisciplinary cooperation, and strengthening ethical and safety training.

(JMIR Nursing 2026;9:e88016) doi:10.2196/88016

KEYWORDS

digital literacy; nursing faculty; higher education; influencing factors; professional development

XSL•FO
RenderX

Introduction

Background

With the rapid advancement of the digital age, the education sector is experiencing a profound digital transformation. In this context, the digital literacy of faculty members has emerged as a central topic in the digital transformation of education. Currently, the international community has established a relatively comprehensive development system for faculty digital literacy. Through policy frameworks, standard setting, resource support, and assessment mechanisms, this system has integrated faculty digital literacy into the faculty education system, thereby promoting the digital transformation of education [1-3]. The Chinese government has actively advanced educational modernization through initiatives such as the "China's Education Modernization 2035" blueprint [4], the World Digital Education Conference [5], and the national smart education public service platform. These endeavors indicate a global transition in education systems from information infrastructure to digital empowerment and intelligent innovation, representing a systematic breakthrough in the digital transformation of education [6].

Although digital technology has been preliminarily implemented in medical education, there are still significant gaps in its application to nursing education. In particular, research on how nursing faculty can effectively use digital technologies to enhance processes in nursing education and optimize practical training outcomes remains insufficiently studied. During the digital transformation of nursing education, the development of digital literacy among college nursing faculty has emerged as a crucial factor in ensuring the effectiveness of pedagogical digitization [7]. While most nursing faculty recognize the necessity of integrating digital technology into teaching practices, they also report facing significant challenges in achieving the effective integration of technological tools in teaching due to inadequate digital literacy levels [8] and imbalanced resource allocation [9,10]. Faculty members also demonstrate significant deficiencies in the design, implementation, and evaluation of digital teaching and learning processes [11,12]. Moreover, the existing digital education model has not yet achieved the expected results in stimulating students' learning motivation [13,14]. Previous studies have also confirmed [15,16] that digital technologies play a key role in clinical scenarios, simulation skills training, and distance learning. Their use enhances students' learning experiences and enables faculty to effectively explain complex operations and nursing concepts [17,18].

Technological pedagogical and content knowledge (TPACK) is a comprehensive framework that identifies well-defined types of knowledge required for effective teaching of subject knowledge enhanced by technology adaptation [19]. The AI-TPACK (artificial intelligence–technological pedagogical content knowledge) framework offers a vital analytical tool, with its 5-layer model (encompassing ethical, awareness, knowledge, competency, and cognitive dimensions) providing a systematic approach to assessing educators' digital literacy [20]. This model emphasizes the integration of technology, pedagogy, and ethical considerations.

However, existing research predominantly focuses on general disciplines, neglecting nursing-specific contexts, and often prioritizes technical operational skills over ethical and pedagogical integration. The results of a recently published cross-sectional study showed that the digital literacy of academic nurse educators was at a moderate level and revealed relevant influencing factors [21]. Although existing studies have initially revealed the value of digital literacy in nursing education, there are still research shortcomings in investigating the digital literacy level at nursing faculty institutions and analyzing the influencing factors.

Objective

This study exhibited a broader sample coverage during the selection of research participants. The research group encompassed both clinical teaching faculty and college nursing faculty, thereby representing an instructional design that fully aligned with the core characteristics of the dual faculty within the AI-TPACK theoretical framework of the nursing program. Through standardized data collection, the system of influencing factors and empirical findings that have been identified significantly augment the comprehensiveness and objectivity of the research results. Furthermore, in response to the pressing requirement for the integration of nursing professional development and digital teaching, this study aimed to analyze the digital literacy status of nursing faculty in colleges and universities and examine their influencing factors through a cross-sectional survey, thereby providing a robust foundation for the formulation of precise and effective competency development strategies and the enhancement of digital teaching quality.

Methods

Ethical Considerations

This study was approved by the Bioethics Review Committee of Fujian Medical University (FJMU174) and conducted in strict compliance with the Declaration of Helsinki. All participants provided written informed consent after receiving detailed explanations regarding the research objectives, procedures, and potential risks. Participants received the appropriate compensation for their time and effort upon completion of the study. To ensure data confidentiality, personal information was anonymized and stored securely using encrypted storage methods.

Instrument Design

A cross-sectional survey was conducted from July 2023 to June 2024. This study was designed to comprehensively capture the current state of digital literacy among nursing faculty at 5 clinical colleges and 3 medical universities in Fujian province, China, thereby facilitating an effective assessment of the current situation and its associated influencing factors without the need for a long-term follow-up study. The inclusion criteria for this study were as follows: (1) nursing faculty in undergraduate colleges and universities or those involved in clinical nursing teaching; (2) ≥1 year of nursing teaching experience; and (3)

XSL•FO
RenderX

provision of informed consent to participate in this study. The exclusion criteria included faculty who were not on duty due to sick leave, personal leave, maternity leave, or further studies. Our sample size was computed using the Soper online sample size calculator, with an expected effect size of 0.15, a statistical power of 0.85, a probability level of 0.05, and a minimum sample size of 140. In total, 339 participants were included in this study, thereby meeting the requirements of statistical efficacy.

Data Measurements

The questionnaire for this study contained 2 parts and took approximately 15 to 20 minutes to complete. As a core research instrument, it effectively enables standardized data collection from nursing faculty in universities in Fujian province. In addition, the tool assessed multiple dimensions of digital literacy and digital technology participation in a structured and easily analyzed format. The tool is aligned with China's digital literacy agenda and thus demonstrates appropriateness and reliability at both the cultural and semantic levels.

Section A: General Information

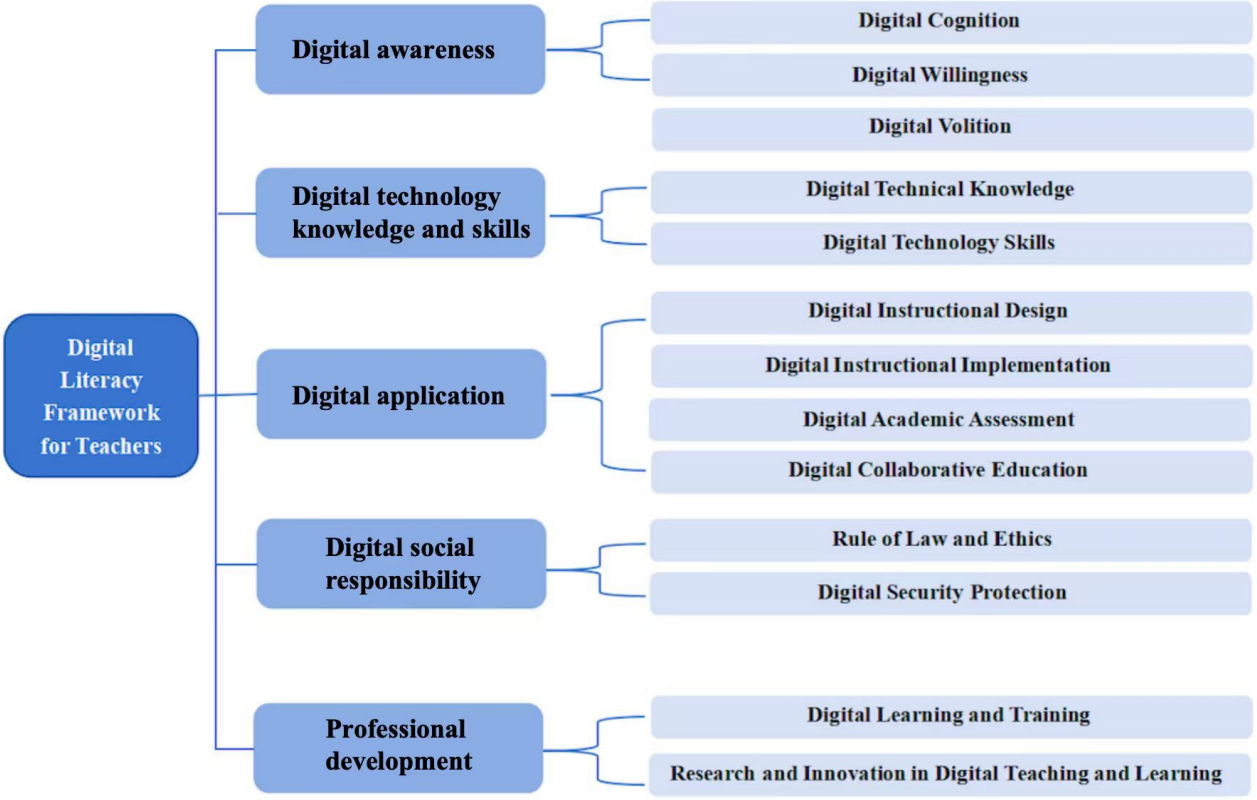
The general information questionnaire investigated the basic demographic characteristics of nursing faculty in colleges and universities. It included gender, age, education level, type of teaching, years of teaching experience, and professional title.

Section B: Nursing Faculty Digital Literacy Questionnaire

We adopted the Teacher Digital Literacy Industry Standard (2022) [6] and adjusted the questionnaire based on the AI-TPACK theoretical framework to align with our research objectives. It included the intelligent integration ethics layer (digital social responsibility), attention to the moral boundaries of technology application; the awareness layer (digital awareness), the willingness to actively adapt to technological changes; the knowledge layer (digital technology knowledge and skills), mastery of the integration point of intelligent tools and subject teaching; the ability layer (digital application), the ability to design intelligent teaching plans; and the thinking layer (professional development), the innovative thinking for the critical integration of technology. The questionnaire included 5 first-level dimensions, 13 second-level dimensions, and 28 third-level dimensions, mainly investigating the level of digital literacy among nursing faculty in colleges and universities. A 5-point Likert scale was used, with scores ranging from 1 for "strongly disagree" to 5 for "strongly agree," and total scores ranging from 28 to 140, with higher scores indicating higher levels of digital literacy among university nursing faculty.

Following iterative discussions among the research team that culminated in the development of the preliminary questionnaire, a rigorous expert validation process was conducted. The selection criteria mandated the inclusion of senior professionals holding associate professor or higher academic ranks with a minimum of 10 years of professional experience in relevant disciplines. The expert panel was strategically composed to encompass cohorts of both nursing education specialists and educational technology experts, ensuring comprehensive domain coverage. Considering the discipline-specific characteristics of nursing and the practice-oriented nature of instructional scenarios, this study systematically refined 28 core dimensions from the educational industry standards through a 2-round Delphi expert consultation process (Figure 1).

Figure 1. Digital literacy framework for nursing faculty.



XSL•FO
RenderX

The reliability of the questionnaire was assessed by pilot testing among 80 college nursing faculty. The Cronbach α of the faculty digital literacy questionnaire was tested to be 0.931, and the Kaiser-Meyer-Olkin value was 0.837, indicating a good level of reliability and validity and fully supporting the questionnaire's scientific rigor in terms of cultural adaptation and semantic validity.

Data collection for this study was officially initiated after ethics approval by the relevant institutional review board, and the questionnaires were collected from July 2023 to June 2024. The research team used the Questionnaire Star platform to administer the electronic questionnaire, which was sent to the nursing faculty in the form of a QR code after obtaining administrative approval from the participating organizations (including academic institutions and hospital nursing departments). Standardized procedures were strictly followed during survey implementation: a uniformly prepared informed consent form and standardized instructions were used, and anonymized data collection was performed after obtaining written informed consent from the study participants. Two master's degree nursing students independently verified the data through the platform's backend, eliminating questionnaires with logical errors and regular responses to ensure the validity of the datasets. Finally, the raw data were stored in an encrypted laptop, and only the researchers in the project team had access to the data.

Data Analysis

Statistical analysis was conducted using SPSS (version 28.0; IBM Corp). Normally distributed continuous variables, including dimension scores, were expressed as mean (SD). Categorical variables such as gender and professional titles were described using frequencies and percentages. Bivariate correlations between variables were examined using Pearson correlation analysis, and multiple linear regression analysis was conducted to identify the predictors of digital literacy dimensions. Before interpreting the regression coefficients, collinearity diagnostics were performed to ensure the stability of the model. An α level of .05 was considered statistically significant.

Results

Sample Characteristics

The study population comprised 339 nursing faculty members from higher education institutions. In total, 343 questionnaires were distributed, yielding a response rate of 98.83% (n=339). Among the respondents, 302 (89.1%) identified as female and 37 (10.9%) as male. Age distribution showed a predominant concentration in the ≤45-years age group (n=326, 96.2%). The vast majority (n=336, 99.2%) held bachelor's degrees or higher educational qualifications. Teaching experience was primarily within the 0- to 10-year range for 218 participants (64.3%). Additionally, 328 (96.6%) participants held intermediate-level academic titles or higher. Further demographic characteristics are detailed in Table 1.

XSL•FO
RenderX

Table . Comparison of digital literacy scores of nursing faculty in higher education institutions with different characteristics (N=339).

Variables	Digital awareness	Digital technology knowledge and skills	Digital application	Digital social responsibility	Professional development
Gender, mean (SD)					
Male (n=37)	4.04 (0.88)	3.25 (0.72)	3.34 (0.94)	3.71 (0.71)	3.95 (0.91)
Female (n=302)	4.04 (0.88)	3.22 (0.79)	3.32 (0.87)	3.72 (0.87)	3.97 (0.84)
t test (*df*)	−0.025 (337)	0.214 (337)	0.133 (337)	−0.076 (337)	−0.173 (337)
P value	.98	.83	.90	.94	.86
Age (y), mean (SD)					
≤35 (n=217)	4.04 (0.87)	3.27 (0.78)	3.40 (0.86)	3.74 (0.85)	4.03 (0.83)
36-45 (n=109)	4.11 (0.86)	3.22 (0.74)	3.21 (0.90)	3.69 (0.84)	3.96 (0.85)
≥46 (n=13)	3.57 (0.90)	2.51 (0.87)	2.91 (0.84)	3.67 (0.92)	3.06 (0.54)
F test (*df*)	2.250 (2, 336)	5.958 (2, 336)	3.163 (2, 336)	0.149 (2, 336)	8.323 (2, 336)
P value	.11	.003[b]	.04[a]	.86	<.001
Education level (y), mean (SD)					
≤12 (n=3)	4.13 (0.12)	4.11 (0.38)	3.78 (0.59)	4.44 (0.42)	3.80 (0.92)
13-15 (n=234)	4.01 (0.90)	3.21 (0.77)	3.30 (0.89)	3.69 (0.83)	3.94 (0.87)
15-16 (n=63)	4.11 (0.82)	3.24 (0.80)	3.27 (0.83)	3.84 (0.89)	4.06 (0.83)
≥17 (n=39)	4.13 (0.86)	3.21 (0.82)	3.48 (0.86)	3.67 (0.94)	4.01 (0.76)
F test (*df*)	0.413 (3,335)	1.307 (3,335)	0.824 (3,335)	1.257 (3,335)	0.397 (3,335)
P value	.74	.27	.48	.29	.76
Teaching experience (y), mean (SD)					
≤5 (n=218)	4.07 (0.83)	3.29 (0.73)	3.22 (0.90)	3.75 (0.80)	4.00 (0.84)
6-10 (n=89)	4.01 (0.97)	3.19 (0.79)	3.49 (0.86)	3.63 (0.96)	3.93 (0.88)
≥11 (n=32)	3.98 (0.94)	2.89 (1.03)	3.47 (0.69)	3.78 (0.87)	3.89 (0.79)
F test (*df*)	0.241 (2,336)	3.862 (2,336)	3.573 (2,336)	0.675 (2,336)	0.363 (2,336)
P value	.79	.02[a]	.03[a]	.51	.70
Professional title, mean (SD)					
Primary (n=11)	3.93 (1.04)	3.27 (0.61)	3.25 (0.73)	3.62 (0.65)	4.13 (0.72)
Intermediate (n=289)	4.04 (0.86)	3.24 (0.79)	3.29 (0.89)	3.73 (0.86)	3.96 (0.86)
Senior (n=39)	4.05 (0.96)	3.09 (0.82)	3.53 (0.81)	3.72 (0.82)	4.02 (0.82)
F test (*df*)	0.097 (2,336)	0.633 (2,336)	1.332 (2,336)	0.079 (2,336)	0.282 (2,336)
P value	.91	.53	.26	.92	.75

[a]*P*<.05.

[b]*P*<.01.

Digital Literacy Score Among Nursing Faculty

The study revealed that nursing faculty exhibited a moderate level of overall digital literacy (mean 3.64, SD 0.59). Among the 3 dimensions assessed, digital awareness demonstrated the highest proficiency (mean 4.04, SD 0.88), whereas digital technology knowledge and skills had the lowest score (mean 3.23, SD 0.78). Digital application (mean 3.32, SD 0.88), digital social responsibility (mean 3.72, SD 0.85), and professional development (mean 3.97, SD 0.85) were also demonstrated at a moderate level. Further details are shown in Table 2.

XSL•FO
RenderX

Table . Faculty scores on digital literacy level 1 dimensions (N=339).

Dimension	Items, n (%)	Score, mean (SD)	Score per item, mean (SD)
Digital awareness	5	20.21 (9.43)	4.04 (0.88)
Digital technology knowledge and skills	3	9.68 (2.92)	3.23 (0.78)
Digital application	9	29.86 (10.30)	3.32 (0.88)
Digital social responsibility	6	22.33 (6.44)	3.72 (0.85)
Professional development	5	19.84 (5.13)	3.97 (0.85)

Comparison of Digital Literacy Scores of Nursing Faculty With Different Characteristics

The results showed statistically significant differences among faculty members across different age groups in the dimensions of digital technology knowledge and skills (P=.003), digital application (P=.04), and professional development (P<.001). Significant differences were also observed among faculty members with different years of teaching experience in the dimensions of digital technology knowledge and skills (P=.02) and digital application (P=.03).

Factors Influencing the Dimensions of Faculty Digital Literacy

Variables with statistically significant differences in each dimension of digital literacy of nursing faculty in colleges and universities with different characteristics (age and years of teaching experience) were selected as independent variables, and each dimension was analyzed using multiple linear regression models as the dependent variable. The variance inflation factor for all independent variables was close to 1 (ranging from 1.003 to 1.007), which is well below the threshold of 5, indicating no multicollinearity issues in the model.

The results of the regression analysis are presented in Table 3. The result showed that age and years of teaching experience were significant predictors of specific dimensions of digital literacy, while sex, education level, and professional titles did not show statistically significant effects. Specifically, for the dimension of digital technology knowledge and skills, both age (β=−.142; P=.009) and years of teaching (β=−.147; P=.006) were significant negative predictors. Regarding digital application, age was found to be a significant negative predictor (β=−.124; P=.02), but teaching experiences show a positive predictor (β=.123; P=.02), suggesting that teaching experience may play a positive role in promoting teachers' digital application capabilities, possibly due to accumulated practical experience in integrating digital tools into teaching over time, whereas age might be associated with differences in adaptability to new digital application scenarios. Similarly, age also had a significant negative impact on the professional development dimension (β=−.153; P=.005) and the total digital literacy level (β=−.136; P=.01).

Table . Standardized coefficients (β) for factors influencing digital literacy among nursing faculty (N=339)[a].

Variables	Digital consciousness	Digital technology knowledge and skills	Digital application	Digital social responsibility	Professional development
Gender	0.007	−0.01	−0.002	0.006	0.015
Age (y)	−0.03	−0.142[b]	−0.124[c]	−0.03	-0.153[b]
Education level	0.058	0.0003	0.046	0.017	0.046
Teaching experiences	−0.04	−0.147[b]	0.123[c]	−0.023	−0.054
Professional title	0.013	−0.067	0.075	0.009	−0.008

[a]All variance inflation factor values were close to 1 and below the threshold of 5, indicating that there was no issue of multicollinearity among the variables.

[b]P<.01.

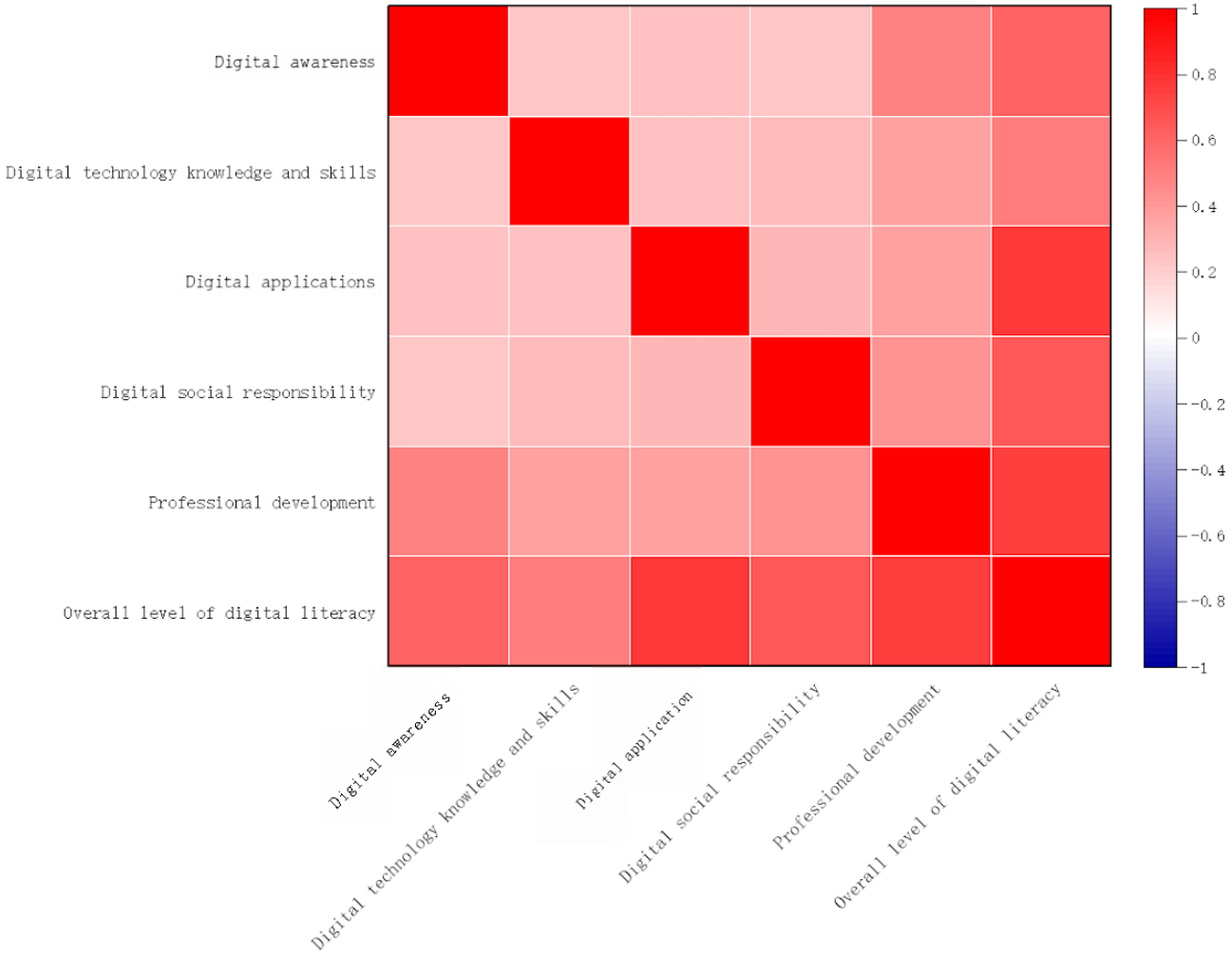
[c]P<.05.

Correlation Analysis of the Dimensions of Digital Literacy Competence of Nursing Faculty

Correlation analysis showed that the overall digital literacy level of nursing faculty in colleges and universities was significantly and positively correlated with all dimensions (P<.001). In addition, the dimensions were positively interrelated, indicating mutual reinforcement among them. Details are presented in Figure 2.

XSL•FO
RenderX

Figure 2. Correlation coefficients between digital literacy and dimensions of nursing faculty.



Discussion

Principal Findings

Our study's findings on nursing faculty's digital literacy scores reveal a nuanced landscape of technological competence, with significant implications for nursing education and practice. The moderate overall digital literacy score (mean 3.64, SD 0.59) suggests that while nursing faculty possess a foundational understanding of digital tools, there remains considerable room for enhancement, particularly in the realm of digital technology knowledge and skills (mean 3.23, SD 0.78), which emerged as the weakest area. This finding is consistent with another study on intergenerational differences in digital competence among faculty [22].

This deficit may hinder faculty's ability to fully leverage advanced technologies in clinical training and patient care, potentially limiting the integration of innovative educational methods such as virtual simulations or AI-driven learning platforms.

The higher proficiency in digital awareness (mean 4.04, SD 0.88) indicates that nursing faculty are generally cognizant of the importance and ethical considerations surrounding digital technology use. This awareness is crucial for maintaining patient confidentiality and navigating the digital health care environment responsibly [23,24]. However, the average scores in digital applications (mean 3.32, SD 0.88), digital social responsibility (mean 3.72, SD 0.85), and professional development (mean 3.97, SD 0.85) suggest a gap between theoretical understanding and practical application. Although faculty may recognize the value of digital tools in education, their limited skills in digital technology knowledge and skills may prevent them from effectively implementing these tools in real-world scenarios [25,26].

The statistically significant differences in digital literacy scores across age groups and years of teaching experience highlight the dynamic nature of technological competence in nursing education. Older faculty members and those with longer teaching experience exhibited lower scores in digital technology knowledge and skills and digital applications, likely due to generational differences in technology exposure and the rapid evolution of digital tools in health care [27,28]. To address this gap, medical schools should develop tailored digital training curricula specifically designed for senior faculty. These programs should be structured to accommodate their unique learning needs and pace, ensuring gradual proficiency in basic digital technology knowledge and skills [29,30].

Furthermore, the findings reveal that faculty members' years of teaching experience also exert a certain influence on their digital literacy levels, which is reflected in a negative correlation

XSL•FO
RenderX

with the dimension of digital technology knowledge and skills, and a positive correlation with the dimension of digital application. Pairwise comparative analysis focusing on the digital technology knowledge and skills dimensions indicated that faculty with more than 11 years of teaching experience exhibited lower willingness to learn digital technology knowledge and skills than those with <5 years of experience. Consequently, this reduced willingness translated to lower levels of mastery of relevant technological knowledge and skills, consistent with the results of previous studies [31]. This phenomenon may be attributed to the dual professional obligations of collegiate nursing faculty, who must concurrently fulfill clinical responsibilities and pedagogical duties [32]. Empirical investigations have demonstrated a statistically significant positive correlation between teaching tenure and burnout prevalence, particularly among faculty with more than a decade of pedagogical experience. Additionally, although experienced faculty possess rich pedagogical expertise accumulated through prolonged teaching practice, they often exhibit diminished motivation for ongoing professional development in emerging digital technologies. Such practitioners' technological knowledge is frequently confined to foundational competencies acquired during initial training phases, resulting in a growing disconnect with rapidly evolving digital innovations [33,34]. This stagnation impedes the systematic advancement of faculty digital technological knowledge and pedagogical skills, thereby constraining their ability to align with contemporary educational transformation requirements and ultimately hindering their capacity to cultivate digital literacy commensurate with the demands of modern education digital paradigm shift.

Additionally, the study showed that faculty members' years of teaching experience were positively correlated with the digital application dimension. This seems to contradict the finding that age was negatively correlated with the digital application dimension; however, we should analyze the effect of years of teaching experience on the level of faculty' digital literacy, all other influences being equal. This phenomenon may be attributed to some interrelated factors. Nursing faculty who have been teaching for a long time usually show significant advantages in the application of digital technology by virtue of their rich accumulation of teaching experience [35]. Research has shown that they are able to systematically screen high-quality digital educational resources (eg, innovative teaching tools such as virtual simulation experiment platforms) that meet the teaching objectives based on the laws of students' cognitive development, rather than unilaterally pursuing technological novelty to gain short-term attention [36]. Empirical research has shown that experienced faculty members are more adept at realizing the organic integration of traditional teaching methods with digital resources, such as visual case analysis method and multimodal narrative teaching method, among other innovative practices [37]. This model of digital application based on teaching a needs-driven approach has led to a significant professional advantage in the dimension of technology application.

The regression analysis provided nuanced insights into the factors shaping nursing faculty's digital literacy, with age and years of teaching emerging as the most robust predictors, while demographic variables such as gender, education level, and professional title showed no statistically significant effects. First, the consistent negative associations between age and multiple dimensions of digital literacy—including digital technology knowledge and skills, digital application, professional development, and overall digital literacy—align with prior research on generational differences in digital competency [38]. This pattern likely stems from a combination of cognitive, attitudinal, and structural factors: older faculty may face age-related declines in technological learning efficiency, experience anxiety or avoidance toward emerging digital tools due to past unsuccessful attempts, and prioritize traditional pedagogical approaches honed through decades of practice over innovative digital methods. Compounding these challenges, institutional resource allocation often prioritizes younger faculty for digital training, leaving senior educators with limited opportunities to update their skills amid heavy administrative and research responsibilities [39,40]. The significant negative relationship between years of teaching and digital technology knowledge and skills further underscores the impact of professional tenure on digital adaptation. While experienced faculty may develop sophisticated digital application strategies rooted in pedagogical expertise, their motivation to acquire new technical knowledge tends to wane over time. This stagnation can be attributed to factors such as burnout from long-term dual clinical and teaching responsibilities [41], as well as a reliance on foundational technological skills acquired during initial training, which may become increasingly outdated in the face of rapid digital evolution [42].

Notably, the absence of significant effects for gender, education level, and professional title suggests that digital literacy disparities in this context are primarily driven by generational and career-stage factors rather than demographic or hierarchical differences. This finding challenges assumptions that higher academic credentials or professional status correlate with greater digital proficiency, highlighting the need for targeted interventions that transcend traditional institutional hierarchies. The low variance inflation factor values (1.003 - 1.007) confirm the absence of multicollinearity, ensuring the reliability of the regression model. This strengthens confidence in the observed relationships between age, teaching tenure, and digital literacy outcomes.

Limitations

Despite the insightful findings, this study is not without limitations that warrant consideration. First, the data were exclusively collected from nursing faculty in universities and clinical colleges located in Fujian province. While the sample provides valuable regional insights, its geographic focus restricts the generalizability of the results, particularly beyond southern China. Future research would benefit from expanding the sampling frame to include diverse geographic regions, institutional types, and health care settings to enhance the external validity of conclusions. Second, the cross-sectional research design used in this study allows for the examination of correlations between variables at a single time point but cannot establish definitive causal relationships. A longitudinal

XSL•FO
RenderX

research design, which tracks changes in digital literacy over time alongside potential influencing factors, would be necessary to disentangle these causal pathways. Finally, the reliance on self-reported questionnaires may introduce response bias, as participants might overestimate their digital skills due to social desirability or lack objective awareness of their competency gaps. Subjective perceptions of digital literacy could thus compromise the accuracy and authenticity of the data. To mitigate this limitation, future studies should integrate objective assessment tools.

Conclusions

This study highlights the critical importance of cultivating digital literacy among nursing educators, whose overall proficiency remains at a moderate level. Gaps persist in the dimensions of digital technology knowledge and skills, necessitating efforts to bridge theory and practice while enhancing applied operational capabilities. Age significantly influences nursing faculty's digital literacy, with younger instructors demonstrating greater advantages in digital teaching, underscoring the necessity for continuous professional development. The impact of teaching experience proves more complex, which further emphasizes the urgent need for medical schools to develop differentiated strategies to enhance digital literacy. Against this backdrop, strengthening digital technology training, implementing systematic instruction on digital ethics and information security, and cultivating educators' digital social responsibility are key to enhancing nursing education quality. For senior faculty, tailored training programs should prioritize gradual, context-specific skill-building that aligns with their learning pace and pedagogical needs, while addressing attitudinal barriers through peer mentoring and success case sharing. For midcareer faculty, interventions should focus on bridging the gap between existing pedagogical expertise and emerging digital tools, fostering innovative integration rather than technical replacement. By addressing age-related and tenure-related disparities, institutions can cultivate a more inclusive digitally competent nursing faculty workforce, better positioned to meet the demands of modern health care education. Future research should examine regional and institutional variations in nursing educators' digital literacy to explore personalized training strategies for targeted improvement.

Acknowledgments

The authors thank all the participants who contributed to this study. Junyu Zhao and Xiaomin Hua are co-first authors.

Funding

This study was supported by the Education and Teaching Project of the Fujian Provincial Department of Education (grant MJJXMS2025001) and the Fujian Provincial Education Planning Project–General Project (grant FJJKBK24-007).

Data Availability

The datasets used and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: Junyu Zhao, Rong Hu, Xiaomin Hua, Chu Chen
Data curation: Junyu Zhao, Rong Hu
Formal analysis: Xiaomin Hua, Chu Chen
Investigation: Wanxin Wu, Yaming Zhang, Yuhe Zhang, Aijing Wu, Yu Zhang
Methodology: Junyu Zhao, Rong Hu, Chu Chen
Project administration: Junyu Zhao, Rong Hu, Chu Chen
Supervision: Junyu Zhao, Rong Hu, Chu Chen
Validation: Junyu Zhao, Xiaomin Hua, Rong Hu, Chu Chen
Visualization: Junyu Zhao, Xiaomin Hua, Rong Hu, Chu Chen
Writing – original draft: Junyu Zhao, Xiaomin Hua
Rong Hu, Chu Chen
Writing – review and editing:
Rong Hu, Chu Chen

Co-Corresponding Author: Chu Chen
School of Health Management, Fujian Medical University, Fuzhou, Fujian, China,
86 5912286526

Conflicts of Interest

None declared.

References

1. The global education community has adopted and released the Education 2030 action framework. UNESCO. 2023. URL: https://www.unesco.org/zh/articles/quanqiujiaoyujiecainabingfabujiaoyu2030xingdongkuangjia [accessed 2025-03-17]

XSL•FO
RenderX

2. ISTE standards: for educators. International Society for Technology in Education. URL: https://iste.org/standards/educators [accessed 2025-03-17]
3. Digital Education Action Plan 2021-2027. European Commission. 2021. URL: https://education.ec.europa.eu/focus-topics/digital-education/actions [accessed 2025-03-17]
4. The CPC Central Committee and the State Council issued the "China Education Modernization 2035" plan. The Central People's Government of the People's Republic of China. 2019. URL: https://www.gov.cn/zhengce/2019-02/23/content_5367987.htm [accessed 2026-03-21]
5. World Digital Education Conference: digital transformation and future of education. Ministry of Education of the People's Republic of China. URL: http://en.moe.gov.cn/features/2023WorldDigitalEducationConference [accessed 2025-03-17]
6. Wang H, Baek J. A systematic review on teacher digital literacy in higher education. World J Soc Sci Res 2023;10(1):38. [doi: 10.22158/wjssr.v10n1p38]
7. Notice from the Ministry of Education on issuing the educational industry standard for "Teachers' Digital Literacy". Ministry of Education of the People's Republic of China. 2022. URL: http://www.moe.gov.cn/srcsite/A16/s3342/202302/t20230214_1044634.html [accessed 2026-03-21]
8. Alanoglu M, Aslan S, Karabatak S. Do teachers' educational philosophies affect their digital literacy? The mediating effect of resistance to change. Educ Inf Technol (Dordr) 2022;27(3):3447-3466. [doi: 10.1007/s10639-021-10753-3] [Medline: 34602847]
9. Lokmic-Tomkins Z, Choo D, Foley P, Dix S, Wong P, Brand G. Pre-registration nursing students' perceptions of their baseline digital literacy and what it means for education: a prospective COHORT survey study. Nurse Educ Today 2022 Apr;111:105308. [doi: 10.1016/j.nedt.2022.105308] [Medline: 35240398]
10. Alshutwi S, Alsharif F, Shibily F, Wedad M A, Almotairy MM, Algabbashi M. Maintaining clinical training continuity during COVID-19 pandemic: nursing students' perceptions about simulation-based learning. Int J Environ Res Public Health 2022 Feb 15;19(4):2180. [doi: 10.3390/ijerph19042180] [Medline: 35206368]
11. Alqahtani FM. Nurse educators and faculty members challenges towards online teaching during COVID-19 pandemic crisis: a qualitative descriptive study. Nurs Open 2024 Nov;11(11):e70062. [doi: 10.1002/nop2.70062] [Medline: 39497483]
12. Kleib M, Arnaert A, Nagle LM, et al. Digital health education and training for undergraduate and graduate nursing students: scoping review. JMIR Nurs 2024 Jul 17;7:e58170. [doi: 10.2196/58170] [Medline: 39018092]
13. Haanes GG, Nilsen E, Mofossbakke R, Wighus M, Ravik M. Digital learning in nursing education: lessons from the COVID-19 lockdown. BMC Nurs 2024 Sep 11;23(1):646. [doi: 10.1186/s12912-024-02312-1] [Medline: 39261882]
14. Bao W. COVID-19 and online teaching in higher education: a case study of Peking University. Hum Behav Emerg Technol 2020 Apr;2(2):113-115. [doi: 10.1002/hbe2.191] [Medline: 32510042]
15. Lin R, Yang J, Jiang F, Li J. Does teacher's data literacy and digital teaching competence influence empowering students in the classroom? Evidence from China. Educ Inf Technol (Dordr) 2023;28(3):2845-2867. [doi: 10.1007/s10639-022-11274-3] [Medline: 36068819]
16. Erdat Y, Sezer Ceren RE, Ozdemir L, Uslu-Sahan F, Bilgin A. Influence of technical, cognitive and socio-emotional factors on digital literacy in nursing students assessed using structural equation modeling. Nurse Educ Today 2023 Nov;130:105937. [doi: 10.1016/j.nedt.2023.105937] [Medline: 37639879]
17. Aydınlar A, Mavi A, Kütükçü E, et al. Awareness and level of digital literacy among students receiving health-based education. BMC Med Educ 2024 Jan 8;24(1):38. [doi: 10.1186/s12909-024-05025-w] [Medline: 38191385]
18. Nylén-Eriksen M, Stojiljkovic M, Lillekroken D, et al. Game-thinking; utilizing serious games and gamification in nursing education - a systematic review and meta-analysis. BMC Med Educ 2025 Jan 29;25(1):140. [doi: 10.1186/s12909-024-06531-7] [Medline: 39881301]
19. Mishra P, Koehler MJ. Technological pedagogical content knowledge: a framework for teacher knowledge. Teach Coll Rec 2006;108(6):1017-1054. [doi: 10.1111/j.1467-9620.2006.00684.x]
20. Yang P, Liu Y, Li Y. The mediating role of self-regulation in fostering Intelligent-TPACK and ethics in physical education teacher education students. Sci Rep 2025 Oct 1;15:34126. [doi: 10.1038/s41598-025-21795-9]
21. Li P, Tan R, Yang T, Meng L. Current status and associated factors of digital literacy among academic nurse educators: a cross-sectional study. BMC Med Educ 2025 Jan 3;25(1):16. [doi: 10.1186/s12909-024-06624-3] [Medline: 39754150]
22. Wu Q, Wang Y, Lu L, Chen Y, Long H, Wang J. Virtual simulation in undergraduate medical education: a scoping review of recent practice. Front Med (Lausanne) 2022;9:855403. [doi: 10.3389/fmed.2022.855403] [Medline: 35433717]
23. Nes AA, Steindal SA, Larsen MH, Heer HC, Lærum-Onsager E, Gjevjon ER. Technological literacy in nursing education: a scoping review. J Prof Nurs 2021;37(2):320-334. [doi: 10.1016/j.profnurs.2021.01.008] [Medline: 33867086]
24. Sánchez-Cruzado C, Santiago Campión R, Sánchez-Compaña MT. Teacher digital literacy: the indisputable challenge after COVID-19. Sustainability 2021;13(4):1858. [doi: 10.3390/su13041858]
25. Gutiérrez-Ángel N, Sánchez-García JN, Mercader-Rubio I, García-Martín J, Brito-Costa S. Digital literacy in the university setting: a literature review of empirical studies between 2010 and 2021. Front Psychol 2022 Sep 6;13:896800. [doi: 10.3389/fpsyg.2022.896800] [Medline: 36176798]

XSL•FO
RenderX

26. Ignatowicz A, Slowther AM, Elder P, et al. Ethical implications of digital communication for the patient-clinician relationship: analysis of interviews with clinicians and young adults with long term conditions (the LYNC study). BMC Med Ethics 2018 Feb 23;19(1):11. [doi: 10.1186/s12910-018-0250-0] [Medline: 29475437]
27. Stoltman JJ, Terplan M. Privacy and digital health: causes for concern and a way forward. J Addict Med 2025;19(4):353-355. [doi: 10.1097/ADM.0000000000001496] [Medline: 40249080]
28. Schlicht L, Wendsche J, Melzer M, Tschetsche L, Rösler U. Digital technologies in nursing: an umbrella review. Int J Nurs Stud 2025 Jan;161:104950. [doi: 10.1016/j.ijnurstu.2024.104950] [Medline: 39603090]
29. Hassanein S, El Arab RA, Abdrbo A, et al. Artificial intelligence in nursing: an integrative review of clinical and operational impacts. Front Digit Health 2025 Mar 7;7:1552372. [doi: 10.3389/fdgth.2025.1552372] [Medline: 40124108]
30. Jobst S, Lindwedel U, Marx H, et al. Competencies and needs of nurse educators and clinical mentors for teaching in the digital age - a multi-institutional, cross-sectional study. BMC Nurs 2022 Aug 28;21(1):240. [doi: 10.1186/s12912-022-01018-6] [Medline: 36031618]
31. Gomez FCJ, Trespalacios J, Hsu YC, Yang D. Exploring teachers' technology integration self-efficacy through the 2017 ISTE Standards. TechTrends 2022;66(2):159-171. [doi: 10.1007/s11528-021-00639-z] [Medline: 34378007]
32. Ljungbeck B, Sjögren Forss K, Finnbogadóttir H, Carlson E. Content in nurse practitioner education - a scoping review. Nurse Educ Today 2021 Mar;98:104650. [doi: 10.1016/j.nedt.2020.104650] [Medline: 33203544]
33. Yao J, Abdullah MN. Five decades of teacher burnout research (1970-2024): a comprehensive bibliometric analysis. Acta Psychol (Amst) 2025 Sep;259:105318. [doi: 10.1016/j.actpsy.2025.105318] [Medline: 40749650]
34. Taylor L, Tello KT. Exploring the relationship between wellness behaviors and burnout amongst university faculty and staff. J Healthy Eat Act Living 2024 Dec 1;4(3):162-173. [Medline: 40385263]
35. Chan MM, Wan AW, Cheung DS, et al. Integration of artificial intelligence in nursing simulation education: a scoping review. Nurse Educ 2025;50(4):195-200. [doi: 10.1097/NNE.0000000000001851] [Medline: 40168182]
36. Mijakoski D, Cheptea D, Marca SC, et al. Determinants of burnout among teachers: a systematic review of longitudinal studies. Int J Environ Res Public Health 2022 May 9;19(9):5776. [doi: 10.3390/ijerph19095776] [Medline: 35565168]
37. Belay AA, Gasheya KA, Engdaw GT, Kabito GG, Tesfaye AH. Work-related burnout among public secondary school teachers is significantly influenced by the psychosocial work factors: a cross-sectional study from Ethiopia. Front Psychol 2023;14:1215421. [doi: 10.3389/fpsyg.2023.1215421] [Medline: 37465483]
38. Loureiro F, Sousa L, Antunes V. Use of digital educational technologies among nursing students and teachers: an exploratory study. J Pers Med 2021 Oct 8;11(10):1010. [doi: 10.3390/jpm11101010] [Medline: 34683151]
39. Hakvoort L, Dikken J, Cramer-Kruit J, Nieuwenhuyzen KM, van der Schaaf M, Schuurmans M. Factors that influence continuing professional development over a nursing career: a scoping review. Nurse Educ Pract 2022 Nov;65:103481. [doi: 10.1016/j.nepr.2022.103481] [Medline: 36327594]
40. Manninen K, Björling G, Kuznecova J, Lakanmaa RL. Ethical Coffee Room: an international collaboration in learning ethics digitally. Nurs Ethics 2020 Jul 15:969733020934145. [doi: 10.1177/0969733020934145] [Medline: 32666868]
41. Christianson J, Bekhet A, Guttormson J, Singh M, Johnson NL. Why we quit: relationships between work-related burnout, work rewards, and intention to leave in nursing. NEC 2024;42(1):7. [doi: 10.62116/NEC.2024.42.1.7]
42. Meum TT, Koch TB, Briseid HS, Vabo GL, Rabben J. Perceptions of digital technology in nursing education: a qualitative study. Nurse Educ Pract 2021 Jul;54:103136. [doi: 10.1016/j.nepr.2021.103136] [Medline: 34246884]

Abbreviations

AI: artificial intelligence
AI-TPACK: artificial intelligence–technological pedagogical content knowledge

Edited by L Donelle; submitted 18.Nov.2025; peer-reviewed by B Dovdon, K Tuamsuk; revised version received 08.Feb.2026; accepted 09.Feb.2026; published 11.Jun.2026.

Please cite as:
Zhao J, Hua X, Hu R, Chen C, Wu W, Zhang Y, Zhang Y, Wu A, Zhang Y
Assessing Digital Literacy Among Nursing Faculty Under the Artificial Intelligence–Enhanced Technological Pedagogical Content Knowledge Framework: Cross-Sectional Survey Analysis
JMIR Nursing 2026;9:e88016
URL: https://nursing.jmir.org/2026/1/e88016
doi:10.2196/88016

© Junyu Zhao, Xiaomin Hua, Rong Hu, Chu Chen, Wanxin Wu, Yaming Zhang, Yuhe Zhang, Aijing Wu, Yu Zhang. Originally published in JMIR Nursing (https://nursing.jmir.org), 11.Jun.2026. This is an open-access article distributed under the terms of

XSL•FO
RenderX

the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Informatics Competency and Technology Self-Efficacy Levels Among Undergraduate Nursing Students: Cross-Sectional Study

Nader Alnomasy[1*], PhD; Habib Alrashedi[1*], PhD; Sharifah Alsayed[2*], PhD; Petelyne Pangket[3*], PhD; Razan Alsayed[4*], MSN

[1]Medical Surgical Department, College of Nursing, University of Ha'il, PO Box 2240, Hail, Saudi Arabia

[2]Medical Surgical Department, College of Nursing, King Saud bin Abdulaziz University for Health Sciences, Jeddah, Saudi Arabia

[3]Medical Surgical Department, College of Nursing, Taif University, Taif, Mecca Region, Saudi Arabia

[4]Nursing Education Department, King Fahd Armed Forces Hospital, Jeddah, Mecca Region, Saudi Arabia

[*]all authors contributed equally

Corresponding Author:
Nader Alnomasy, PhD
Medical Surgical Department, College of Nursing, University of Ha'il, PO Box 2240, Hail, Saudi Arabia

Abstract

Background: The Saudi Arabian health care sector is transforming under Vision 2030 with the goal of digitizing its services. This necessitates a digitally prepared nursing workforce; however, there is evidence suggesting that nursing students have limited informatics competency and that these skills are minimally covered in their training programs.

Objective: This study measured the baseline informatics competency and technology self-efficacy of undergraduate nursing students at the University of Hail, Saudi Arabia.

Methods: Using a descriptive cross-sectional design, data were collected from 243 undergraduate nursing students at the University of Hail through an online survey. The survey measured demographics, informatics competency (Canadian Nurse Informatics Competency Assessment Scale), and digital technology self-efficacy. Data analysis used descriptive statistics, 2-tailed *t* tests, ANOVA, and hierarchical regression analysis.

Results: Students reported a moderate level of informatics competency, with a mean Canadian Nurse Informatics Competency Assessment Scale item score of 2.57 (SD 0.84) on a scale from 1 to 4. They also showed moderate to high self-efficacy for digital technology, with a mean item score of 2.7 (SD 0.56) out of 4. Informatics competency scores were substantially higher in students with prior informatics training and frequent electronic health record exposure. In the underpowered exploratory hierarchical regression model, self-efficacy for digital technology showed a small positive association with informatics competency that approached statistical significance; however, the overall model was not statistically significant and explained only a limited proportion of the variance.

Conclusions: This single-site study of undergraduate nursing students at the University of Hail found moderate self-reported informatics competency despite higher digital technology self-efficacy, indicating a need for strengthened informatics education to support the Saudi Vision 2030. While underpowered, exploratory analyses and self-report data mean that these institution-specific findings should be interpreted cautiously and used primarily to guide future multisite and mixed methods research.

(JMIR Nursing 2026;9:e85283) doi:10.2196/85283

KEYWORDS

informatics competency; digital health; nursing education; self-efficacy; electronic health records

Introduction

The digital health revolution continues to transform health care delivery and organizational structures. Advances in artificial intelligence (AI), telemedicine, and electronic health records (EHRs) now support safer, more efficient, and data-driven care [1,2]. Consequently, preparing a globally competent health informatics workforce is essential for implementing these innovations and enhancing clinical outcomes [3,4].

Nurses comprise the largest portion of the health care workforce and are in a pivotal position to integrate technology into direct patient care. Informatics competency, the integration of computers to manage and communicate data [3], is a key skill that nurses must develop. Equally important is technology self-efficacy, or nurses' confidence in using digital tools in clinical and educational settings [5]. Despite the wide recognition of their importance, the inclusion of nursing informatics in undergraduate curricula varies considerably worldwide. Studies from North America, Europe, Africa, and

XSL•FO
RenderX

the Asia-Pacific region have shown significant differences in course content, learning opportunities, and assessment methods used to measure informatics competency among nursing students [6-8].

International initiatives such as Technology Informatics Guiding Education Reform and the Canadian Nurse Informatics Competency Assessment Scale (C-NICAS) promote the formal integration of informatics into nursing education. However, their adoption varies widely across different programs and countries. This study used the C-NICAS as an assessment tool and conceptual framework to identify key informatics competency domains among undergraduate nursing students. Despite the global use of such frameworks, empirical studies have shown that students and new graduates consistently self-report only basic to intermediate proficiency, often lacking hands-on experience with clinical information systems. This highlights the persistent gap between curriculum design and real-world practice [4].

As part of Vision 2030, Saudi Arabia has launched a nationwide digital health initiative to modernize health care using digital tools, telemedicine, and AI [9-11]. Initiatives such as the Healthcare Sector Transformation Program aim to transform service delivery, improve accessibility, and enhance the quality and efficiency of care, reflecting a national focus on developing a data-driven, patient-centered workforce. However, studies have shown that Saudi nursing students and practicing registered nurses have lower informatics competency than their international counterparts, particularly those in countries with more advanced digital technologies and infrastructure [12-15]. Despite significant reform efforts, the integration of informatics into Saudi nursing education has not yet reached the level observed in many other high-income countries, indicating the need for targeted policy and curriculum intervention.

Although many global models advocate for the integration of informatics into nursing education, very little research has addressed this topic in Saudi Arabia. Few studies have investigated how prepared nursing students are for digital health roles or how their informatics competency relates to their confidence with technology, both of which are important factors for successfully joining today's digital health care workforce. These research gaps need to be filled so that evidence can guide faculty training and improve nursing curricula, especially given the Vision 2030 focus on digital health. This study measured the baseline informatics competency and digital readiness of nursing students at the University of Hail, Saudi Arabia. The goal was to provide local data that address a known gap in preparing Saudi Arabia's future nurses to work in digitally transformed health care settings.

Methods

Study Design

This study adopted a descriptive cross-sectional design to assess perceived informatics competency and digital technology self-efficacy among undergraduate nursing students using an online self-assessment survey.

Study Population and Sampling

Voluntary participation in an electronic survey allowed for data collection from undergraduate nursing students at a university in Hail, covering students from the first to fourth years (levels 1 - 4) of the Bachelor of Nursing program. The target population was all undergraduate nursing students enrolled in the Bachelor of Nursing program at the University of Hail during the 2024 - 2025 academic year (levels 1 - 4 and interns). The source population comprised all students with an active institutional email address on the official program mailing list at the time of data collection. Therefore, the accessible sampling frame consisted of this official mailing list of currently enrolled undergraduate nursing students, which was used to distribute a single open invitation to participate in the study.

Students were eligible if they (1) were currently enrolled as undergraduate nursing students in levels 1 to 4 at the University of Hail, (2) were able to read and understand English (the language of instruction), and (3) had an active institutional email account. Students on academic suspension, on long-term leave, or who had withdrawn from the program at the time of data collection were excluded from the study.

A convenience sampling approach was used in which all eligible students from the source population were invited to participate. This strategy was practical, cost-effective, and enabled the coverage of the broadest possible segment of the accessible student population. In this context, convenience sampling reflects all students who chose to respond to the open email invitation rather than a preselected subgroup. All invitations were distributed via the official undergraduate nursing program's email list. Because the list is updated dynamically, the exact number of invitations sent could not be confirmed, but it approximately matched the total number of eligible students. The survey was conducted using Google Forms. A total of 260 students started the survey, of whom 93.5% (243/260) completed it and were included in the analysis. In total, 6.5% (n=17) of the responses were excluded due to substantial missing data (more than 20% of the items were unanswered). Because the total number of active student email accounts at the time of the survey could not be verified, the exact response rate could not be calculated. However, program records indicated that the mailing list approximately matched the total number of enrolled and eligible students.

For inferential analyses involving year level, only students in levels 1 to 4 were included, excluding interns because their internship year is structurally distinct from the numbered academic levels. To reduce potential selection bias, all eligible undergraduate nursing students on the institutional mailing list were invited to participate using the recruitment protocol described in the Data Collection section.

Study Size

A priori power analysis indicated that at least 315 respondents were required to achieve 80% power at a 95% confidence level for medium effects in ANOVA, 2-tailed *t* tests, and regression analyses. Therefore, the final analytic sample of 243 respondents was underpowered for confirmatory inferential analyses but

XSL•FO
RenderX

was considered adequate for descriptive and exploratory correlational analyses.

Instrument

The questionnaire was divided into 3 sections to capture comprehensive participant information. The first section collected general demographic data, including age, sex, academic year, prior experience with electronic medical records, and self-assessed computer proficiency. These variables provided a contextual background for interpreting the findings in the subsequent sections.

The second section assessed participants' informatics competency using the 21-item C-NICAS version 2 with permission from Kleib and Nagle [3]. This student-focused instrument has been adapted and validated for undergraduate nursing students across multiple international contexts, supporting its suitability for this population and enabling cross-study comparison. Originally developed and psychometrically validated among registered nurses in Alberta, Canada, the scale has also undergone cross-cultural validation, supporting its use in international benchmarking.

The C-NICAS includes 4 subscales: foundational information and communications technology (ICT) skills, information and knowledge management, professional responsibility and accountability, and use of ICTs in the delivery of care. These domains align with the core competencies relevant to the digital health ecosystem in Saudi Arabia. The instrument was administered in English, consistent with the language of instruction at the study site. Participants self-reported their competency using a 4-point Likert scale ranging from 1 ("not competent") to 4 ("very competent"). The total scores ranged from 21 to 84 and were categorized into 4 levels: not competent (21-41), somewhat competent (42-62), competent (63-83), and very competent (84). The scale demonstrated strong internal consistency (Cronbach α=0.93 overall and 0.74 - 0.89 for the subscales).

The final section measured digital technology self-efficacy using a revised 17-item version of the questionnaire by Hughes [16]. This scale complements the C-NICAS by assessing students' confidence in their ability to use digital technology. Responses were recorded on a 4-point Likert scale ranging from 1 ("strongly disagree") to 4 ("strongly agree"). In total, 12 items were reverse coded, and 5 items (1, 7, 8, 11, and 16) were positively coded. The mean scores ranged from 1 to 4, with higher scores indicating greater self-efficacy. The scale demonstrated strong reliability, with the Cronbach α exceeding 0.90 in both the original study and the sample in this study.

Both informatics competency and digital technology self-efficacy were assessed using self-report measures. Although both instruments have strong psychometric properties, self-reported data may be subject to bias and inaccurate self-assessment, potentially leading to over- or underestimation of actual competency and confidence levels.

Data Collection

An ethical recruitment plan was developed. Nursing students were invited through email, which contained an invitation letter, a survey link, an information sheet, and a consent statement. The invitation highlighted the purpose of the research, the importance of the study, and the possible implications of the study for nursing informatics research.

Reminder emails using a predetermined protocol were sent to encourage response; the first reminder was sent 24 to 48 hours after the initial email, followed by another one 3 to 7 days later, with a total of 3 to 5 reminders sent throughout the data collection period. Google Forms was used for the survey because of its accessibility and reasonable data storage capacity.

Statistical Analysis

Data analysis was performed using SPSS (version 29.0; IBM Corp) and the Python software (Python Software Foundation). Item-level missing data were minimal (<5% for all variables). Continuous items with less than 5% of missing data were imputed using item-specific means and categorical variables using the mode. Given the low proportion of missing data, this approach was unlikely to bias the descriptive estimates, although it may have slightly attenuated the variance in the inferential analyses.

Statistical assumptions were assessed prior to the inferential analyses. Descriptive statistics, including measures of central tendency (mean, median, and mode) and variability (SD), were calculated for key variables, such as C-NICAS and digital technology self-efficacy scores. Inferential analyses included Pearson correlation to examine relationships among variables, independent-sample *t* tests, and one-way ANOVA to assess group differences.

Hierarchical multiple regression was conducted using 4 models to identify the predictors of informatics competency. Potential confounders (age, sex, academic level, prior informatics training, and previous EHR exposure) were included; however, given the underpowered sample and limited variables, residual and unmeasured confounding (eg, informal IT training or personal interest in technology) cannot be excluded.

The significance level was set at an α value of .05. Effect sizes were interpreted using the Cohen standards for correlations (small=0.10; medium=0.30; large=0.50), and *F* change statistics were used to assess changes in R^2 values across regression models. As the final sample was below the a priori target (N=243 vs 315), all inferential analyses should be considered exploratory and hypothesis generating rather than confirmatory.

Ethical Considerations

Institutional review board approval from the University of Hail (H-2024-437) was obtained prior to data collection. Participation was entirely voluntary and no financial compensation was provided to the participants. All procedures were conducted in accordance with the ethical standards of the Research Ethics Committee and the 1964 Declaration of Helsinki and its subsequent amendments.

Informed consent was obtained from all students prior to accessing the online survey and questionnaire. Participation was entirely voluntary, and students had the right to withdraw at any time without adverse academic consequences. To preserve participants' anonymity, no personal identifiers (such as names

XSL•FO
RenderX

or email addresses) were collected through Google Forms, and the researchers did not link any responses to individual participants. The dataset was stored on a secure, password-protected server, and access was limited to the research team members.

This study posed an extremely low risk as it involved noninvasive procedures and presented no more than minimal risk to the participants. The participants' normal coursework was not affected, and it was clearly stated that their decision to participate or not would have no impact on their academic status or their relationship with faculty members.

Results

Table 1 presents the study's 243 undergraduate nursing student participants from Saudi Arabia, with most being young adults aged 21 to 25 years (n=116, 47.7%) and only 7% (n=17) aged 26 years and above. The sample was predominantly female (n=159, 65.4%). Students were distributed across academic levels, with interns accounting for the largest group (n=63, 25.9%) and year level 8 students representing the smallest group (n=13, 5.3%). The participants were nearly equally divided between those with informatics training (n=124, 51%) and those without (n=119, 49%). Comfort with digital technology varied, with 41.2% (n=100) expressing discomfort and 22.2% (n=54) remaining neutral on the subject. EHR experience ranged from 40.3% (n=98) having occasional exposure to 29.6% (n=72) having no prior experience with EHRs. Most participants demonstrated good technology accessibility, with 82.3% (n=200) using technology daily for educational purposes and 65% (n=158) reporting excellent home internet access. Regarding career intentions, most (n=130, 53.5%) planned to pursue clinical care, whereas 13.6% (n=33) remained undecided about their future nursing career paths. A comparison between respondents and nonrespondents was not possible because identifiable sampling frame data were unavailable; therefore, some nonresponse bias cannot be excluded.

XSL•FO
RenderX

Table . Demographic characteristics (N=243).

Category	Values, n (%)
Age (years)	
≤20	110 (45.3)
21-25	116 (47.7)
≥26	17 (7)
Sex	
Male	84 (34.6)
Female	159 (65.4)
Year level	
First	69 (28.4)
Second	72 (29.6)
Third	39 (16)
Fourth	63 (25.9)
Have you previously completed any informatics-related training or coursework?	
No	119 (49)
Yes	124 (51)
Rate your comfort level with digital technology (eg, computers, smartphones, and EHRs[a])	
Very uncomfortable	47 (19.3)
Uncomfortable	53 (21.8)
Neutral	54 (22.2)
Comfortable	89 (36.6)
Previous exposure to EHRs	
No experience	72 (29.6)
Yes, occasionally (monthly or less)	98 (40.3)
Yes, frequently (weekly or daily)	73 (30)
Do you own or regularly have access to a personal computer or laptop?	
No	49 (20.2)
Yes	194 (79.8)
How often do you use technology (smartphones, tablets, or computers) for educational purposes?	
Monthly	11 (4.5)
Weekly	32 (13.2)
Daily	200 (82.3)
Internet accessibility at home	
Fair	19 (7.8)
Good	66 (27.2)
Excellent	158 (65)
Intended future nursing career area	
Undecided	33 (13.6)
Administration or leadership role	14 (5.8)
Academic or teaching role	35 (14.4)
Nursing informatics	14 (5.8)
Public health or community nursing	17 (7)
Clinical care (hospital or clinical setting)	130 (53.5)

XSL•FO
RenderX

[a]EHR: electronic health record.

Table 2 presents the self-reported C-NICAS item scores from the 243 undergraduate nursing students (mean 2.57, SD 0.84 on a scale from 1-4); the median was 2.50 (2.00-3.00). Regarding the total number of C-NICAS scores, the largest percentage of students (n=116, 47.7%) achieved a score of "somewhat competent"; 21.4% (n=52), or nearly 1 in 5, achieved a score of "not competent"; 23.9% (n=58) achieved a score of "competent"; and only 7% (n=17) achieved a score of "very competent." When reviewing the subscales, the "foundational ICT skills" subscale had the highest average C-NICAS item score (2.62, SD 0.93), followed by the "use of ICT in the delivery of patient care" (2.60, SD 0.85) and "professional and regulatory accountability" (2.56, SD 0.88) subscales, whereas the "information and knowledge management" subscale had the lowest average C-NICAS item score (2.51, SD 0.86). Most students reported that they rated themselves as somewhat competent on 3 out of the 4 subscales (n=75, 30.9% - n=106, 43.6%), and most students who evaluated themselves as not competent did so on the "foundational ICT skills" subscale (n=58, 23.9%) and least frequently in the "use of ICT in the delivery of patient care" subscale (n=43, 17.7%). Overall, students self-reported moderate levels of informatics competency; however, many rated themselves as only somewhat competent in practically all areas of informatics competency.

Table . Nursing students' perceptions of informatics competency (Canadian Nurse Informatics Competency Assessment Scale [C-NICAS] mean item scores; N=243).

Scale and subscale	Item score (1-4), mean (SD)	Item score (1-4), median (IQR)	Range	Not competent, n (%)	Somewhat competent, n (%)	Competent, n (%)	Very competent, n (%)
NI[a] competency (overall C-NICAS)	2.57 (0.84)	2.50 (2.01–3.10)	1-4	52 (21.4)	116 (47.7)	58 (23.9)	17 (7)
Foundational ICT[b] skills	2.62 (0.93)	2.66 (2.00–3.33)	1-4	58 (23.9)	75 (30.9)	78 (32.1)	32 (13.2)
Information and technology management	2.51 (0.86)	2.50 (2.00–3.00)	1-4	57 (23.5)	103 (42.4)	57 (23.5)	26(10.7)
Professional and regulatory accountability	2.56 (0.88)	2.50 (2.00–3.17)	1-4	52 (21.4)	102 (42)	58 (23.9)	31 (12.8)
Use of ICT in the delivery of patient care	2.60 (0.85)	2.66 (2.00–3.17)	1-4	43 (17.7)	106 (43.6)	65 (26.7)	29 (11.9)

[a]NI: nursing informatics

[b]ICT: information and communications technology.

The mean self-efficacy score for the use of digital technologies was 2.7 (SD 0.56) on a scale from 1 to 4. This indicates that, as a group, the nursing students reported a moderate to high level of confidence in using digital technologies. Because the mean score of 2.7 is above the scale midpoint, the students' digital technology self-efficacy can be considered strong.

Table 3 presents the data on informatics competency and digital technology self-efficacy among Saudi nursing students. There were no statistically significant differences in the C-NICAS total scores by year level for students in levels 3 to 8 (P=.48) or in the digital technology self-efficacy mean item scores (P=.87). The mean C-NICAS total scores ranged from 48.3 (SD 15.6) in level 3 to 55.4 (SD 17.9) in level 7, whereas the digital technology self-efficacy mean item scores ranged from 2.5 (SD 0.5) to 2.7 (SD 0.2) on a scale from 1 to 4. Exploratory t test analysis showed that students with prior informatics training had higher mean C-NICAS total scores (55.0, SD 15.8) than those without such training (50.6, SD 15.2; P=.005), whereas their mean digital technology self-efficacy item scores were similar (2.6, SD 0.4 vs 2.5, SD 0.5; P=.53). Exploratory ANOVA further indicated that students with no EHR experience had the lowest mean C-NICAS total scores (48.4, SD 16.1) compared with higher scores among those with occasional (53.4, SD 17.2) or frequent (54.9, SD 13.4) EHR use (P=.002); in contrast, digital technology self-efficacy mean item scores remained relatively stable across groups with EHR use (2.5, SD 0.4 to 2.6, SD 0.4; P=.40).

XSL•FO
RenderX

Table . Overall competency and self-efficacy levels by year level, prior informatics training, and routine electronic health record (EHR) use.

Variable	C-NICAS[a] total score (21-84), mean (SD)	*P* value	DT-SE[b] item score (1-4), mean (SD)	*P* value
Year level		.48		.87
3	48.3 (15.6)		2.6 (0.4)	
4	53.1 (16.2)		2.6 (0.4)	
5	53.5 (16.7)		2.5 (0.5)	
6	53.2 (14.3)		2.6 (0.5)	
7	55.4 (17.9)		2.7 (0.2)	
8	53.7 (15.7)		2.6 (0.4)	
Informatics training		.005		.53
Yes	55.0 (15.8)		2.6 (0.4)	
No	50.6 (15.2)		2.5 (0.5)	
EHR use		.002		.40
No experience	48.4 (16.1)		2.6 (0.4)	
Occasional	53.4 (17.2)		2.5 (0.4)	
Frequent	54.9 (13.4)		2.6 (0.4)	

[a]C-NICAS: Canadian Nurse Informatics Competency Assessment Scale.

[b]DT-SE: digital technology self-efficacy.

Table 4 illustrates the preliminary correlational findings showing a positive association between digital technology self-efficacy and 3 of the informatics competency subscales. The strongest correlation was with foundational ICT skills (r=0.181; P=.002), followed by professional responsibility and accountability (r=0.157; P=.002) and use of ICT in care delivery (r=0.113; P=.02). However, the strongest correlations were between the competency subscales themselves, indicating a good level of internal connection. Specifically, the strongest correlation was between professional accountability and use of ICT in care (r=0.794; P=.001), followed closely by the correlation between foundational ICT skills and professional accountability (r=0.769; P=.001) and between foundational ICT skills and use of ICT in care (r=0.765; P=.001).

Table . Correlation between competency and self-efficacy.

	DT-SE[a]	Foundational ICT[b] skills	Professional and regulatory accountability	Use of ICT in the delivery of patient care
DT-SE				
r	1	0.181	0.157	0.113
P value	—[c]	.002	.002	.02
Foundational ICT skills				
r	0.181	1	0.769	0.765
P value	.002	—	.001	.001
Professional and regulatory accountability				
r	0.157	0.769	1	0.794
P value	.002	.001	—	.001
Use of ICT in the delivery of patient care				
r	0.113	0.765	0.794	1
P value	.02	.001	.001	—

[a]DT-SE: digital technology self-efficacy.

[b]ICT: information and communications technology.

[c]Not applicable.

Table 5 presents the outcomes of an exploratory hierarchical multiple regression analysis examining potential predictors of informatics competency measured using the C-NICAS score. In model 1, age and sex were entered, yielding an R value of

XSL•FO
RenderX

0.013 and an adjusted R^2 value of 0.005 ($F_{2,243}$=1.611; P=.20), with no significant predictors. In model 2, year level (fourth-eighth) and internship status were added, resulting in an R value of 0.031 and an adjusted R^2 value of −0.002 ($F_{8,237}$=0.950; P=.48), and again, no significant predictors emerged. In model 3, prior EHR experience and informatics training were added, yielding an R value of 0.054 and an adjusted R^2 value of 0.013 ($F_{10,235}$=1.330; P=.22), with no variables reaching statistical significance. In the final model (model 4), the mean digital technology self-efficacy score was included, resulting in an R value of 0.069 and an adjusted R^2 value of 0.022 ($F_{11,234}$=1.577; P=.11). Digital technology self-efficacy showed at most a weak, borderline association with informatics competency (B=4.34; P=.05), accounting for only a negligible proportion of the variance; thus, it cannot be considered a robust predictor in this sample. Given the modest effect size, nonsignificant overall model, and sample size (N=243) falling below the a priori target (N=315), these regression results should be interpreted with caution and viewed as exploratory and hypothesis generating rather than confirmatory. Overall, the hierarchical models explained only a small proportion of the variance in informatics competency scores.

Table . Hierarchical multiple regression models predicting informatics competency[a].

Model	Variables entered	R^2	Adjusted R^2	F test (df)	P value	Significant predictors (final model)
1	Age and sex	0.013	0.005	1.611 (2, 243)	.20	None
2	+ year level (fourth-eighth) and internship	0.031	−0.002	0.950 (8, 237)	.48	None
3	+ EHR[b] exposure and informatics training	0.054	0.013	1.330 (10, 235)	.22	None
4	+ mean digital technology self-efficacy score	0.069	0.022	1.577 (11, 234)	.11	Digital technology self-efficacy (B=4.34; P=.05)

[a]None of the hierarchical regression models reached conventional statistical significance for the overall F test, and the explained variance was small (maximum adjusted R^2=0.022); therefore, these findings, including the trend-level association for digital technology self-efficacy, should be interpreted as exploratory and preliminary.

[b]EHR: electronic health record.

Discussion

Principal Findings

Overview

This study assessed the baseline informatics competency of undergraduate nursing students at a single institution in Hail, Saudi Arabia, based on self-reported informatics competency scores and revealed a moderate level of competency. Most students rated themselves as somewhat competent. These results suggest that current competency levels may be insufficient to fully support the digital health priorities articulated in the Saudi Vision 2030, indicating the need to further strengthen informatics education and experiential exposure to digital health systems. In addition, this study found that students with previous informatics education had notably higher C-NICAS scores than those without.

Informatics Competency Levels

A moderate level of informatics competency, particularly lower scores in the “information and knowledge management” and “professional responsibility and accountability” subscales, suggests that students may have basic ICT and documentation abilities but lack proficiency in the critical appraisal and interpretation of digital health data, application of ethical and regulatory standards, and more advanced use of information to support clinical decision-making. These deficits are likely related to limited systematically integrated informatics content within undergraduate curricula and restricted, structured exposure to digital health systems, particularly EHRs.

Several studies in Saudi Arabia have reported that nursing students possess basic informatics knowledge but do not consistently achieve high competency [17,18]. The reported deficits are greatest in clinical informatics and applied computer skills, highlighting the gap between knowledge and practice [17,18]. Students proficient in everyday digital technologies often struggle to transfer these abilities to clinical settings that require specific informatics skills [18].

Other studies have reinforced the distinction between digital literacy and informatics competency [19,20]. Nursing students may report moderate or high digital health literacy; however, literacy alone does not ensure competency in critical appraisal, data interpretation, or safe clinical application [19]. Although there have been initiatives to integrate digital skills into the curriculum, the lack of rich experiential learning limits students’ progression to competent or advanced informatics levels [20]. Considering these findings and the C-NICAS results in our study, Saudi nursing programs should substantially expand informatics content across all years of study. The lowest mean scores in the “information and knowledge management” subscale underscore the need for modules that teach students to manage, retrieve, and critically appraise health information from electronic databases. These competencies can be addressed

XSL•FO
RenderX

through dedicated courses and by embedding informatics objectives into existing theory and clinical units, consistent with recent international recommendations [21].

Experiential learning should include structured, practice-focused engagement with digital health systems, particularly EHRs, as students with previous EHR exposure in this study demonstrated higher competency. This aligns with evidence that greater EHR integration is associated with improved informatics knowledge and skills [22,23]. Targeted faculty development is also needed so that educators can confidently model and assess informatics competency in both classroom and clinical settings.

Academic-practice partnerships that provide students with supervised access to real-time informatics environments can help bridge the gap between classroom learning and clinical practice. Aligning these curricular changes with the Saudi Vision 2030 health strategy is essential for building a digitally competent nursing workforce capable of supporting high-quality, safe care across the kingdom [9,23,24].

Digital Technology Self-Efficacy

The moderately high digital technology self-efficacy found in this study likely stems from students' heavy use of digital tools, widespread internet access, and regular use of online resources. Although digital technology self-efficacy showed a small positive correlation with informatics competency, the hierarchical regression model was not statistically significant, explaining only 2% of the variance. Consequently, digital technology self-efficacy was not a strong predictor of informatics competency in this study. These results should be viewed as hypothesis generating, suggesting a weak association that requires confirmation in larger samples. Instead of making causal claims, the focus remained on robust descriptive and bivariate findings such as differences related to prior informatics training and EHR exposure.

A trend toward increased willingness to use digital technologies for educational and clinical purposes is evident among frequent users, consistent with the findings of another study [25]. Recent studies indicate that nursing students generally show acceptable preparedness for AI-based health care and a positive disposition toward digital health tools [14,25,26]. In Saudi Arabia, nursing students typically view e-learning favorably and are proficient in technology-rich environments, particularly in blended learning contexts [14]. Prior experience and digital literacy remain key indicators of self-efficacy and preparedness in undergraduate nursing education [26].

Nevertheless, digital technology self-efficacy varies considerably [14,27]. Many students remain uncomfortable with digital technologies, especially when transitioning to clinical settings [27]. Furthermore, some are dissatisfied with e-learning due to a lack of in-person interaction and support [14]. Technical difficulties, inconsistent design, and inadequate training can also decrease confidence and satisfaction [14].

These findings suggest that access and exposure are insufficient. Students face obstacles in applying personal technological knowledge to professional contexts and require more support [14,26]. Digital technology self-efficacy can be strengthened by providing consistent hands-on digital literacy education, individualized assistance, and cooperative blended learning environments. Nursing programs should also build strong technological foundations and systematically collect student feedback to improve these environments [14,26]. This approach will help students develop the confidence and skills necessary to use digital tools effectively in both study and practice.

Role of EHR Experience

Early experience with EHRs significantly influences nursing students' informatics competency. There was a positive relationship between EHR experience and informatics competency: students who reported frequent or occasional hands-on use had higher competency scores than those with little or no experience. Practicing informatics in real time enhances the translation of abstract knowledge into practical applications, and providing students with authentic EHR experiences contributes to their clinical preparation [28]. Exposure to educational EHR platforms requires technical proficiency and has broader implications for student learning.

Institutions emphasizing EHR use can expect improved student performance in informatics assessments and application skills, resulting in graduates who are better prepared for practice [22]. Structured exposure to EHR systems may also enhance students' documentation skills, clinical reasoning, and compliance with professional and regulatory requirements. As the demand for advanced informatics skills increases, all students will need structured and equitable access to educational EHR systems [28].

Digital literacy and informatics competency gaps are most apparent when students transition from classrooms to clinical settings. Many students who do not receive instruction on EHR use feel unprepared to document clinical data or perform other informatics tasks [28]. Furthermore, obstacles during clinical rotations are multifaceted, including limited EHR access, unclear learning expectations, and variable guidance from preceptors [28].

Similar concerns exist in the Saudi context, where inequitable student access to EHRs at clinical sites has been identified [15]. Nursing programs should provide comprehensive longitudinal training programs that address EHR use. Furthermore, programs should incorporate simulated or laboratory-based experiences that replicate authentic informatics tasks such as data retrieval, clinical decision-making, and regulatory documentation [21].

Finally, partnering with health care organizations can ensure that students have supervised EHR access and clear learning objectives during clinical rotations [28]. These strategies help bridge the gap between basic digital literacy and professional informatics competency, enhancing readiness for real-world practice, increasing patient safety, and creating long-term career opportunities in the digital health care field.

Implications

These results have significant implications for nursing education and policy in Saudi Arabia. Considering the moderate level of informatics competency, curricular reform is needed to align nursing students with the digital elements of Vision 2030. This can be achieved not only by viewing faculty as presenters of

XSL•FO
RenderX

subject matter but also by emphasizing authentic learning experiences. This could be further aided by incorporating a specific informatics module or student engagement activities that rely on EHRs and digital health technologies, which have demonstrated successful clinical experiences in American and Canadian programs. Finally, it is of utmost importance to formalize partnerships between universities and practice settings to ensure that students have authentic experience with real-time informatics in practice settings. These initiatives could yield a digitally competent workforce that can use technology in patient care and contribute to health initiatives in the future.

Conclusions

This single-site survey found that undergraduate nursing students at the University of Hail showed noticeable gaps in informatics competency even though many reported moderate to high confidence in using digital technologies. This highlights the need to revisit the current curriculum considering Saudi Vision 2030 and its emphasis on digital health. The findings also suggest that being comfortable with everyday technology does not necessarily mean that students are prepared with the informatics skills required in clinical settings. Students with prior exposure to informatics content or experience with EHRs generally demonstrated better competency levels. However, the study sample was smaller than originally planned, and the statistical analyses were not sufficiently powered. Therefore, any observed relationship, including the weak link between digital technology self-efficacy and informatics competency, should be interpreted with caution and considered exploratory.

It is also important to note that this study relied on a single-site convenience sample and self-reported data, both of which limit the generalizability of the findings. These factors may also lead to an overestimation of the actual competency levels. Therefore, the results are best understood as specific to this context rather than reflective of all nursing students in Saudi Arabia. Nevertheless, this study offers useful baseline information that can support local curriculum improvements and inform future research, particularly larger multisite or mixed methods studies aimed at strengthening nursing informatics education in the country.

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Hübner U, Shaw T, Thye J, et al. Technology Informatics Guiding Education Reform - TIGER. Methods Inf Med 2018 Jun;57(S 01):e30-e42. [doi: 10.3414/ME17-01-0155] [Medline: 29956297]
2. O'Connor S, Hubner U, Shaw T, Blake R, Ball M. Time for TIGER to ROAR! Technology Informatics Guiding Education Reform. Nurse Educ Today 2017 Nov;58:78-81. [doi: 10.1016/j.nedt.2017.07.014] [Medline: 28918322]
3. Kleib M, Nagle L. Development of the Canadian Nurse Informatics Competency Assessment Scale and evaluation of Alberta's registered nurses' self-perceived informatics competencies. Comput Inform Nurs 2018 Jul;36(7):350-358. [doi: 10.1097/CIN.0000000000000435] [Medline: 29668498]
4. Harerimana A, Wicking K, Biedermann N, Yates K. Click, care, connect: incorporating nursing informatics into undergraduate nursing curricula. Inform Health Soc Care 2026;51(1):2-28. [doi: 10.1080/17538157.2025.2582852] [Medline: 41289030]
5. Hariyati RT, Handiyani H, Wildani AA, Afriani T, Nuraini T, Amiruddin MH. Disparate digital literacy levels of nursing manager and staff, specifically in nursing informatics competencies and their causes: a cross-sectional study. J Healthc Leadersh 2024;16:415-425. [doi: 10.2147/JHL.S470456] [Medline: 39473455]
6. Nagle LM, Kleib M, Furlong K. Digital health in Canadian schools of nursing part A: nurse educators' perspectives. Qual Adv Nurs Educ 2020;6(1). [doi: 10.17483/2368-6669.1229]
7. Kaihlanen AM, Gluschkoff K, Kinnunen UM, Saranto K, Ahonen O, Heponiemi T. Nursing informatics competences of Finnish registered nurses after national educational initiatives: a cross-sectional study. Nurse Educ Today 2021 Nov;106:105060. [doi: 10.1016/j.nedt.2021.105060] [Medline: 34315050]
8. Kleib M, Nagle LM, Furlong KE, Paul P, Duarte Wisnesky U, Ali S. Are future nurses ready for digital health?: Informatics competency baseline assessment. Nurse Educ 2022;47(5):E98-E104. [doi: 10.1097/NNE.0000000000001199] [Medline: 35324499]
9. Mani ZA, Goniewicz K. Transforming healthcare in Saudi Arabia: a comprehensive evaluation of Vision 2030's impact. Sustainability 2024;16(8):3277. [doi: 10.3390/su16083277]
10. Aboalshamat K, Alhuzali R, Alalyani A, et al. Medical and dental professionals readiness for artificial intelligence for Saudi Arabia Vision 2030. Int J Pharm Res Allied Sci 2022;11(4):52-59. [doi: 10.51847/NU8y6Y6q1M]

XSL•FO
RenderX

11. Alsufyani AM, Alforihidi MA, Almalki KE, Aljuaid SM, Alamri AA, Alghamdi MS. Linking the Saudi Arabian 2030 vision with nursing transformation in Saudi Arabia: roadmap for nursing policies and strategies. Int J Afr Nurs Sci 2020;13:100256. [doi: 10.1016/j.ijans.2020.100256] [Medline: 33072514]
12. Elmorshedy H, AlAmrani A, Hassan MH, Fayed A, Albrecht SA. Contemporary public image of the nursing profession in Saudi Arabia. BMC Nurs 2020;19:47. [doi: 10.1186/s12912-020-00442-w] [Medline: 32528229]
13. Brinjee D, Thobaity AA, Ahmari WA, Almalki M. Current issues facing disaster nursing education in Saudi Arabia. Research Square. Preprint posted online on Mar 9, 2021. [doi: 10.21203/rs.3.rs-290535/v1]
14. Alshammari A, Fayez Alanazi M. Use of technology in enhancing learning among nurses in Saudi Arabia; a systematic review. J Multidiscip Healthc 2023;16:1587-1599. [doi: 10.2147/JMDH.S413281] [Medline: 37313273]
15. Alghamdi R, Albloushi M, Alzahrani E, Aldawsari A, Alyousef S. Nursing education challenges from Saudi nurse educators' and leaders' perspectives: a qualitative descriptive study. Int J Nurs Educ Scholarsh 2019 Feb 23;16(1). [doi: 10.1515/ijnes-2018-0027] [Medline: 30798324]
16. Hughes JE. Descriptive indicators of future teachers' technology integration in the PK-12 classroom: trends from a laptop-infused teacher education program. J Educ Comput Res 2013;48(4):491-516. [doi: 10.2190/EC.48.4.e]
17. Almarwani AM, Yacoub MI. Examining the effect of an educational program on nursing students' informatics competencies. Nurs Educ Perspect 2023;44(6):E59-E61. [doi: 10.1097/01.NEP.0000000000001106] [Medline: 36792354]
18. Abou Hashish EA, Alnajjar H. Digital proficiency: assessing knowledge, attitudes, and skills in digital transformation, health literacy, and artificial intelligence among university nursing students. BMC Med Educ 2024 May 7;24(1):508. [doi: 10.1186/s12909-024-05482-3] [Medline: 38715005]
19. Hassan Mekawy S, Ali Mohamed Ismail S, Zayed Mohamed M. Digital health literacy (DHL) levels among nursing baccalaureate students and their perception and attitudes toward the application of artificial intelligence (AI) in nursing. Egypt J Health Care 2020 Mar;11(1):1266-1277. [doi: 10.21608/ejhc.2020.274757]
20. Alqahtani N. English language usage and academic achievement among nursing students: a cross-sectional study. SAGE Open Nurs 2022;8:23779608221109364. [doi: 10.1177/23779608221109364] [Medline: 35837245]
21. Raghunathan K, McKenna L, Peddle M. Baseline evaluation of nursing students' informatics competency for digital health practice: a descriptive exploratory study. Digit Health 2023;9:20552076231179051. [doi: 10.1177/20552076231179051] [Medline: 37274371]
22. Choi J, Bove LA, Tarte V, Choi WJ. Impact of simulated electronic health records on informatics competency of students in informatics course. Healthc Inform Res 2021 Jan;27(1):67-72. [doi: 10.4258/hir.2021.27.1.67] [Medline: 33611878]
23. Young Y, Alharthy A, Hosler AS. Transformation of Saudi Arabia's health system and its impact on population health: what can the USA learn? Saudi J Health Syst Res 2021 Oct 5;1(3):93-102. [doi: 10.1159/000517488]
24. Suleiman AK, Ming LC. Transforming healthcare: Saudi Arabia's Vision 2030 healthcare model. J Pharm Policy Pract 2025;18(1):2449051. [doi: 10.1080/20523211.2024.2449051] [Medline: 39845746]
25. Al-Olaimat K, Salameh B, Abdulhalim Alqadi R, et al. Readiness and acceptance of nursing students regarding AI-based health care technology on the training of nursing skills in Saudi Arabia: cross-sectional study. JMIR Nurs 2025 Jul 30;8:e71653. [doi: 10.2196/71653] [Medline: 40737490]
26. Alkhuzaimi F, Wilson CB, Wong WY. Key factors influencing undergraduate nursing students' perceptions of the use of learning management systems: a systematic literature review. BMC Nurs 2025 Mar 26;24(1):323. [doi: 10.1186/s12912-025-02962-9] [Medline: 40140804]
27. Ghaythan Alessa M, Fawaz AlQahtani J, Mohammed Kharma S, Salman Alamri E, Mohideen Nagoor Thangam M. Bridging the practice gap with the virtual learning -awareness and readiness among the new nursing graduates. J Pioneer Med Sci 2025 May 5;14(4):81-88. [doi: 10.47310/jpms2025140411]
28. Whitt KJ, Allen CL, Hogg CW, et al. The use of electronic health records in advanced practice nursing education: a scoping review. J Prof Nurs 2024;50:83-94. [doi: 10.1016/j.profnurs.2023.11.007] [Medline: 38369377]

Abbreviations

AI: artificial intelligence
C-NICAS: Canadian Nurse Informatics Competency Assessment Scale
EHR: electronic health record
ICT: information and communications technology

XSL•FO
RenderX

Edited by E Borycki, J Sarvestan; submitted 04.Oct.2025; peer-reviewed by UM Kinnunen, Y Liu; revised version received 10.May.2026; accepted 11.May.2026; published 08.Jun.2026.

Please cite as:
Alnomasy N, Alrashedi H, Alsayed S, Pangket P, Alsayed R
Informatics Competency and Technology Self-Efficacy Levels Among Undergraduate Nursing Students: Cross-Sectional Study
JMIR Nursing 2026;9:e85283
URL: https://nursing.jmir.org/2026/1/e85283
doi:10.2196/85283

© Nader Alnomasy, Habib Alrashedi, Sharifah Alsayed, Petelyne Pangket, Razan Alsayed. Originally published in JMIR Nursing (https://nursing.jmir.org), 8.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Potential for Smartphone-Based Reflection to Support Socioemotional Competence Development in Student Nurses' Clinical Placement: Mixed Methods Field Trial

Ivo Bril-Broers[1,2], MSc; Nick Degens[3], PhD; Joke Fleer[1], Prof Dr

[1]Department of Health Sciences, Section Health Psychology, University of Groningen, A. Deusinglaan 1, Groningen, The Netherlands

[2]Research Group Human-Centered Technologies, Cluster Arts, Design & Built Environment, Hanze University of Applied Sciences, Groningen, The Netherlands

[3]Research Group (Dis)Connected Technology & Creativity, University of the Arts Utrecht, Utrecht, The Netherlands

Corresponding Author:
Ivo Bril-Broers, MSc
Department of Health Sciences, Section Health Psychology, University of Groningen, A. Deusinglaan 1, Groningen, The Netherlands

Abstract

Background: Socioemotional competencies (SEC) are essential for nurses to provide quality care and to maintain their own health. Historically underrepresented, nursing programs are changing their curricula to incorporate SEC development more strongly. Although it is recommended to practice SEC during clinical placement, studies on how to support nursing students in developing these competencies are lacking.

Objective: Our study explores second-year nursing students' perceived and actual use of a reflection tool during their first 20-week clinical placement. Designed to support SEC development, the study focuses on how students incorporate the application in their learning process for SEC development.

Methods: An embedded mixed method study was conducted with 15 students from a nursing program in the Netherlands. Qualitative surveys and semistructured interviews were used to capture students' experiences with the application both during and at the end of the clinical placement. Log data from the application provided insight into how the app was used. The mixed methods data were analyzed using warranted assertion analysis.

Results: The app helped students relate learning goals to clinical experiences, identify learning opportunities, and experience a sense of growth. Features that removed practical barriers to reflecting were experienced as most beneficial. Although SEC-focused use was limited (n=7), involved students used the app over a longer period of time and reported greater benefits. Nevertheless, 13 students stopped using the app halfway through their clinical placement, as their attitude shifted from learning to working.

Conclusions: This study gives a rich description of how nursing students used a reflection app to structure their learning process during clinical placement. It reveals that smartphone-based reflection supports students in clinical placement learning. Specifically, the app lowers barriers to engaging in structural reflection and decreases the mental burden of self-regulated learning. However, more guidance and a multimodal intervention approach seem necessary for SEC development. Further research into how to account for students' shifting attitudes toward learning is needed.

(JMIR Nursing 2026;9:e91083) doi:10.2196/91083

KEYWORDS

mobile app; smartphone; education; mixed methods; field trial; technology enhanced learning; nursing students; clinical placement; nurse; design research; social-emotional competence; soft skills; clinical practice; nursing education; reflection

Introduction

Background

Socioemotional competencies (SEC) play a large role in the day-to-day work of a nurse, both to provide quality care as well as to maintain their own well-being [1-5]. As our understanding of the importance of these nursing competencies grows, so too does the interest in how to effectively teach [3,6-8] and assess [1,9] SEC in nursing programs.

A recent systematic review underlines the importance of developing SEC during clinical practice [10], showing nursing students are unprepared for the social dynamics of the workplace otherwise. Traditionally, however, the focus of clinical placement has been on improving technical skills [11-14]. Concurrently, nontechnical skills such as SEC are expected to

XSL•FO
RenderX

be learned through a social and more implicit learning process [15-17]. The review's findings suggest supporting SEC development within the clinical learning environment requires specific didactical attention. Yet, there is little research on what clinical placement needs or how to implement it there.

Research findings indicate that SEC require nursing programs to adopt active and innovative learning methodologies in their curricula [8,18,19]. Studies show that when reflection is used as an active learning methodology, SEC development is supported [20-23]. Reflection helps students to process their clinical experiences and learn from them [23]. Specific to SEC, reflection can support the development of interprofessional communication and collaboration skills [20,21], as well as a better understanding of oneself and the workplace experience [22-24]. Thus, using reflection specifically to make the implicit learning process of SEC more explicit warrants further exploration.

Over the past decades, studies on reflection within clinical placement have explored the potential of smartphone technology for its ease of use and ability to support students' development when workplace coaching is limited [11,14,25,26]. When it comes to clinical skills, research shows this to be an effective approach to improve learning during clinical placement [26-28]. Moreover, many studies show nursing students, educators, and preceptors favor digital apps over paper-based methods of reflection [13,27-31]. Reasons for this preference include greater flexibility, a more convenient method of managing one's reflections, and better visualization of one's progress [29]. Although smartphone-based reflection can support clinical skill development, there is little research on how such apps can support SEC development.

This study aims to address this gap and explore the potential of smartphone-based reflection for SEC development during clinical placement. The study is part of a larger project in which we developed a reflection app called Nurse-IT in co-creation with nursing students and educators. The app aims to help students structurally reflect on their clinical experiences for SEC development purposes.

In the development process of the app, design guidelines from both theory and practice have been translated to concrete design features (eg, daily and weekly reflections). Given the lack of applied research on this topic, our study's main focus is to explore the in situ impact of a SEC-focused reflection app on students' learning processes. By focusing on students' experiences with the app from a learning process perspective, we better understand not just the app's impact, but also how it interacts with the learning environment in which it is introduced. Doing so provides a more nuanced and practical understanding of the developed app and the underpinning design guidelines. This leads us to the following design research question: *How do the app's features impact nursing students' socioemotional competency development during clinical placement?*

To answer this question, we conduct an embedded mixed methods field trial of the Nurse-IT app. By gathering a diverse set of data during the field trial, we aim to get a detailed view on how nursing students incorporate the app's structural reflection in their learning process. We also measure their perception and usage of the app throughout the clinical placement to better understand how usage develops over time.

The App

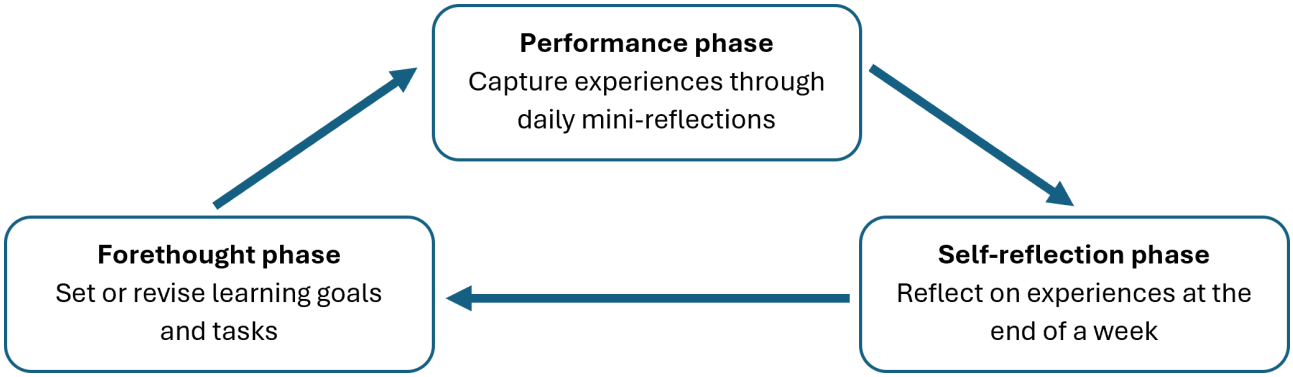
Design Guidelines

Designing educational interventions requires incorporating and balancing both theory and practice. We began with an overarching learning process based on theory and clinical placement as a learning context. This was developed further into a prototype through co-creation. We briefly summarize the most salient design guidelines and the app itself here.

Theory

The clinical learning experience focuses on learning-by-doing and real-life practice, with students encountering many social situations during a shift. Part of the app's goal is to support students in capturing those moments to be able to reflect and learn from them, facilitating further experiential learning [32]. Moreover, the aim is to empower them to structurally and intentionally work on their SEC, resulting in a self-regulated learning cycle [33] (Figure 1). Finally, as an educational app is dependent on students using it consistently, its design needs to make the primary task (ie, reflection) easy and enjoyable. For this, we used the principles of persuasive systems design [34].

Figure 1. Self-regulated learning cycle which starts at the forethought phase, to then monitor oneself in the performance phase, and to evaluate oneself in the self-reflection phase after which the cycle begins anew.



XSL•FO
RenderX

Practice

Nursing students and educators of a nursing program in the Netherlands were involved to make sure that the app's design and user experience align with their needs [11,14,30,35]. Although the overarching learning process was determined beforehand, the co-creative process was necessary to translate it to a design that corresponds with the reality of doing a clinical placement. One outcome of the entirety of the co-creative process was the formulation of five design guidelines:

1. Provide structure, but allow for flexibility: The app needs to scaffold goal-setting and reflection by providing preset questions [36]. Those questions should allow students to express themselves in a way that suits their needs.
2. Low time-investment when capturing experiences: The performance phase requires students to structurally capture their experiences. Given the high workload of clinical placement, capturing those experiences should be quick and easy to do. Even if that means having a low level of reflection.
3. Looking back, but also planning ahead: The app should facilitate students in planning follow-up actions based on what they learned from their reflections.
4. Help me see, monitor, and appreciate my growth: The app needs to summarize and visualize the student's growth toward their learning goals. In doing so, the students should feel in control of their learning process [35].
5. No one has access but me, but I can share things if I want: To further underline student autonomy and learner agency, the app's contents should only be visible to the student [37,38]. Nothing should be assessed unless the student wants to share something themself.

Features of the App

Forethought Phase: Learning Goals and Tasks

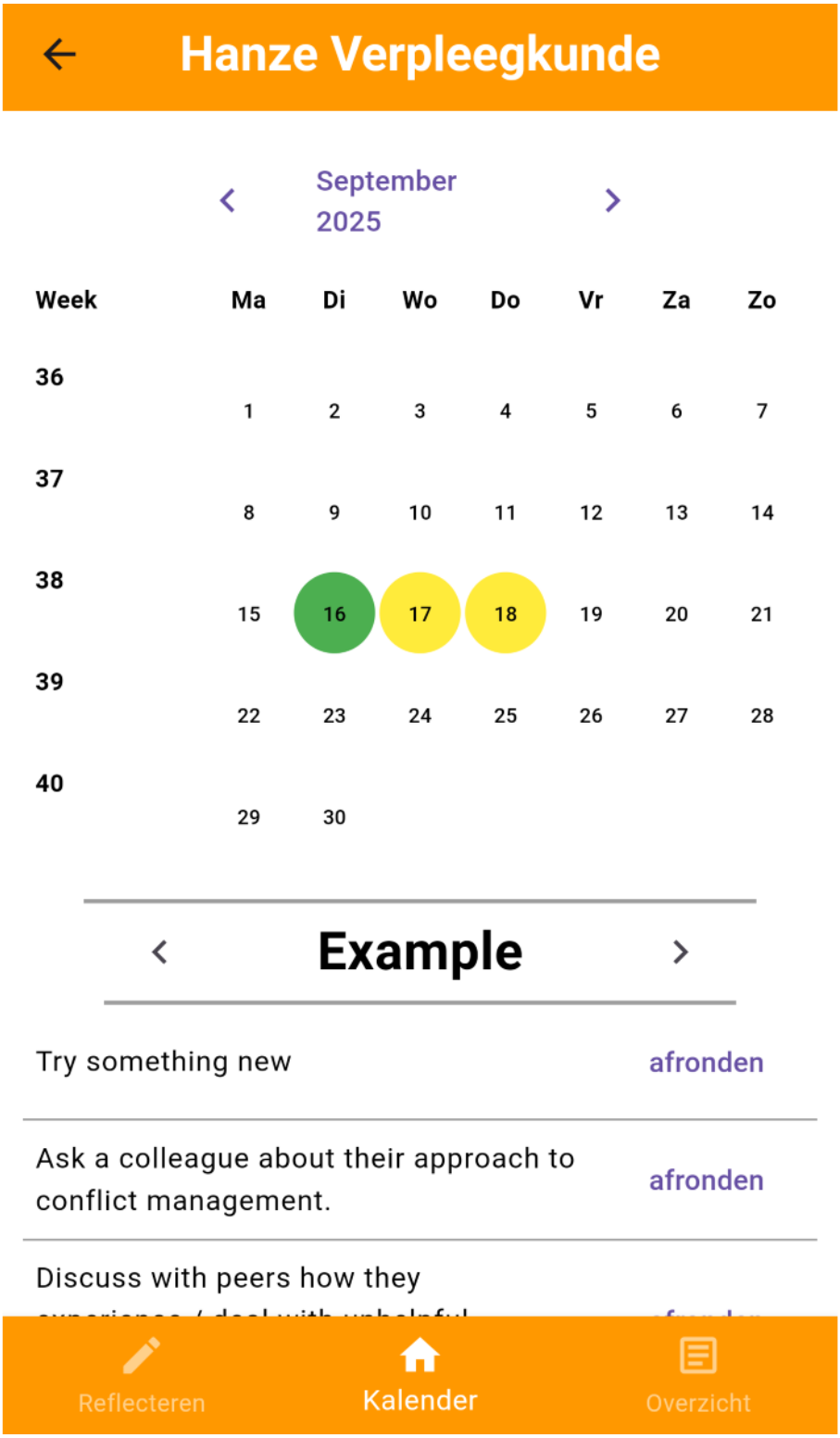
Clinical placement requires students to reach specific learning outcomes. Goal-setting helps structure this learning process and represents the forethought phase of self-regulated learning. Moreover, setting goals allows the students to monitor their progress and stay aware of the goals they are working on [39]. Consequently, learning goals are a core feature of the app: students formulate their own or use the curriculum's competencies. They then add a target grade they want to work toward and determine the period in which they want to work on the learning goal (guideline 3).

Another aspect of the forethought phase is planning. Setting a goal is one thing, but breaking it down into smaller, more achievable parts (ie, proximal goals) helps focus students' efforts toward achieving the goal and raises self-efficacy [39]. This also ties into experiential learning's phase of active experimentation. The app facilitates this by allowing students to add tasks to their learning goals (guideline 1).

To further help students be aware of their learning goals and respective tasks, they are placed on the home screen of the app (Figure 2). Here they can navigate between learning goals and check off tasks (guideline 4).

XSL•FO
RenderX

Figure 2. Home screen of the app. The calendar shows a month with 3 completed daily reflections, each represented by a colored dot. The word 'Example' represents a learning goal, with the lines of text underneath each representing an individual learning task.



Performance Phase: Daily Reflections

Daily reflections represent the performance phase of self-regulated learning. The main goal of daily reflections is to have the student self-monitor and evaluate after each shift (guideline 4). To ensure students are willing to do so, we favor a low time investment over achieving deeper reflection (guideline 2) [34]. In doing so, we aim to achieve two things: (1) students record more situations and build up a reservoir of experiences which they can reflect on, and (2) students engage in a frequent and structured way of reflection-on-action and thus have better awareness of their progress (ie, self-monitoring) [33,39].

The daily reflection's design reflects these goals by using features that provide immediate low-level insights and help summarize an experience from both a descriptive (eg, title and tags) and an emotional level (eg, emojis). In terms of evaluation,

XSL•FO
RenderX

the student (1) grades themselves on the learning goal they are writing the reflection for, (2) can write a *tip* for themselves, something they could do differently next time, and (3) can write a *top* (ie, a positive comment) about what the user felt went well. Finally, the student can use a text field to write about the shift in whichever way they want (called "free writing"). It is to be noted that we only require the student to give a grade to their shift; all other features are optional. This way we follow the design guideline of providing structure but allowing for flexibility (guideline 1).

Self-Reflection Phase: Weekly Reflections

With the daily reflections covering the performance phase, the weekly reflections represent the self-reflection phase and the revisiting of the forethought phase of self-regulated learning. Consequently, the main goal of the weekly reflections is twofold: (1) to achieve deep reflection and use different experiences gained through daily reflections, and (2) to review one's progress and make plans for the week to come (guideline 3).

This is translated into the app through a three-step process, aligning with experiential learning principles (and guideline 1):

1. A student looks back at the week and describes what they are proud of and what they found difficult.
2. They choose one of the 2 and reflect on why they feel that way, what situation caused those feelings and how they dealt with the situation, what they learned from the experience, and how they would approach similar situations next time.
3. They review their learning goals and select those that relate to the week's reflection. Learning goals and tasks can be marked as complete, edited, or newly created.

Ideally, a weekly reflection helps students identify when they are stuck in their learning process and need to reach out to others for help. As the reflection app keeps the students' learning goals and reflections private to themselves, only the student can act on app-related insights. This, combined with findings of our co-creative process and earlier work that show students can have trouble identifying such situations [17], made us aim to create triggers for help-seeking behavior.

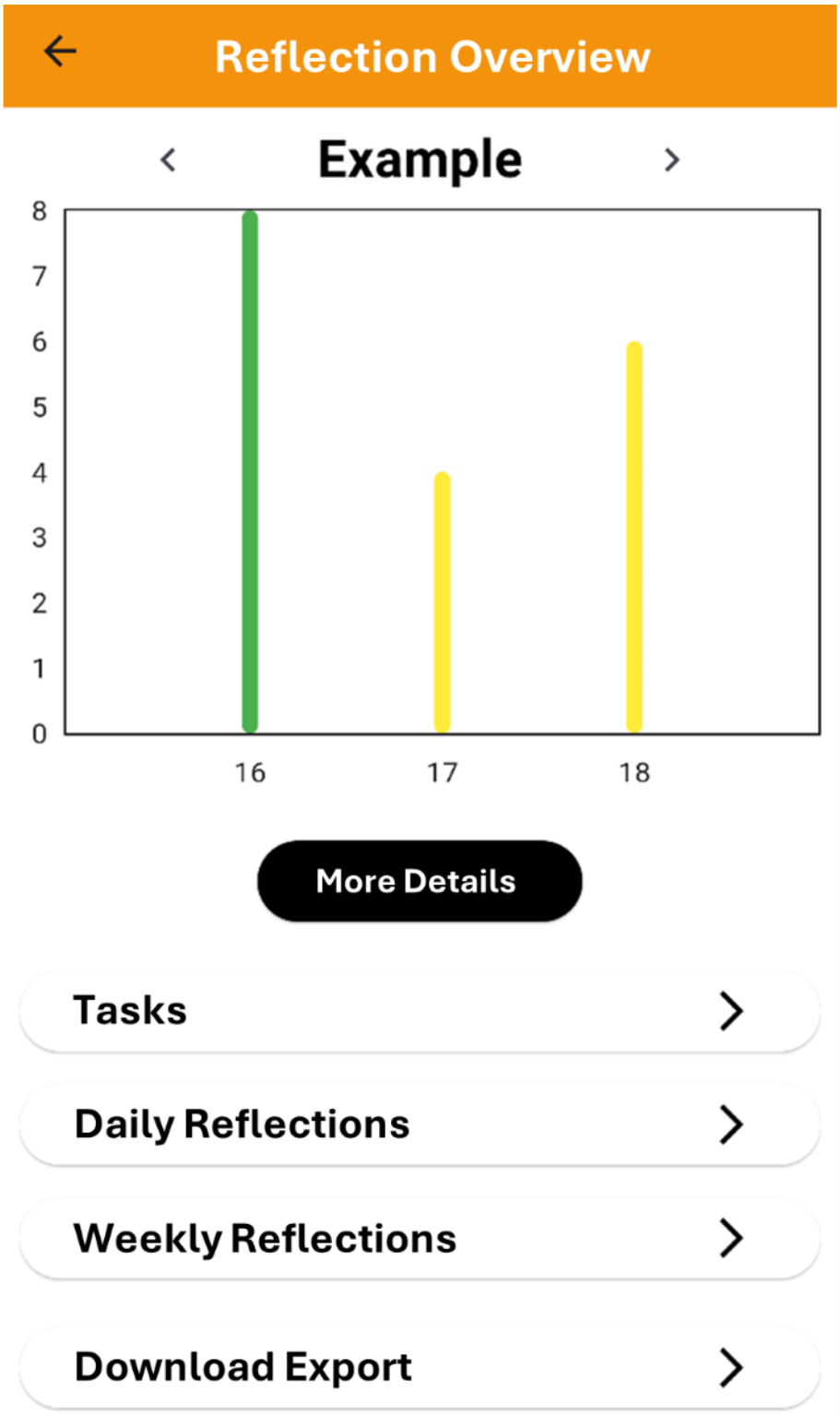
Learning Goal Overview

Students can review their active and completed learning goals through the overall learning goal overview screen. Here, they can also review their progress on a specific learning goal through the learning goal overview (guideline 4), which harbors all the related daily and weekly reflections, tasks, and visualizes the daily reflection grades over time (Figure 3). As per the design guideline concerning sharing data, this page also allows students to export their learning goals for documentation or coaching purposes (guideline 5).

Having discussed the features of the app, Figure 4 summarizes their connection to the design guidelines.

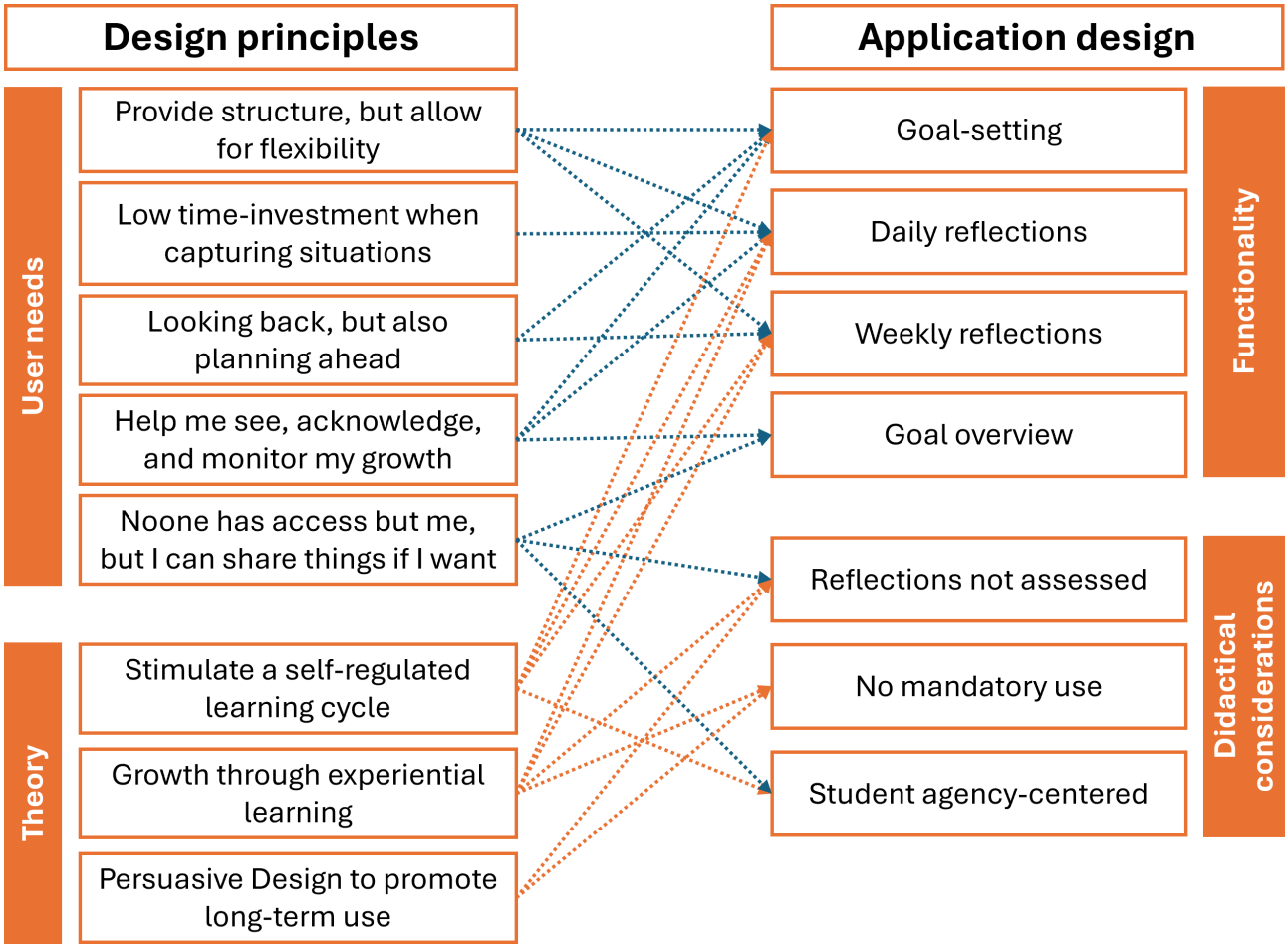
XSL•FO
RenderX

Figure 3. Overview of an example learning goal. The graph shows the grades of three daily reflections, the color indicating the grade they gave themselves (1 - 3 red, 4 - 6 yellow, 7 - 10 green). The 'tasks,' 'daily reflections,' and "weekly reflections" buttons lead to their respective overview pages. Download export provides a spreadsheet version of the learning goal's activities.



XSL•FO
RenderX

Figure 4. Overview of how the design guidelines are instantiated in the design of the app.



Methods

Study Design

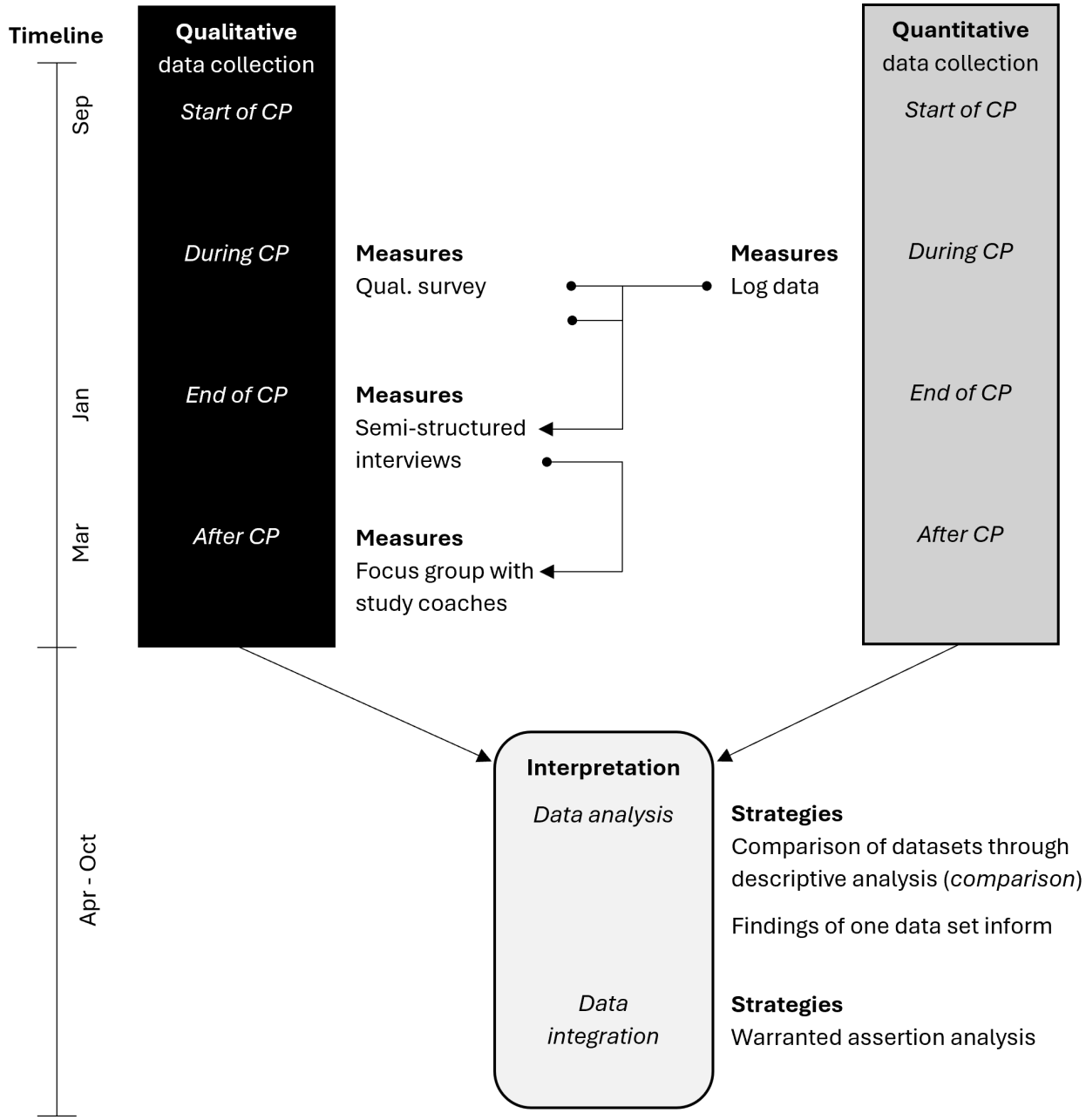
The study used an embedded mixed methods research approach within a pragmatist paradigm to capture the potential richness of a naturalistic field trial. The goal of mixed methods research is to integrate knowledge of several methods, both quantitative and qualitative, allowing meta-inferences to be drawn [40]. These meta-inferences aim to explain phenomena in a way that exceeds the sum of a study's parts (ie, each individual method). Thus, describing the selection of methods as well as the purpose and timing of integrating results in mixed methods studies is important.

For this study, reflection contents, open-ended surveys, and interviews give a qualitative perspective on how students used the tool, how that usage developed over time, and a rich description of their context of use. Embedded within this qualitative study are log data. These give a quantitative perspective on how participants engaged with the tool. By combining perspectives, our understanding is informed both from the actual usage of the tool, as well as the participants' context-rich experiences with the tool. Together, both strands help illuminate the app's role within the student's learning process: which parts of the app are used, what needs do they meet, and why?

It is recommended to visualize the mixed methods design in a diagram to condense its complexity in an explainable manner [41]. In Figure 5, we provide a chronological, vertical diagram of the above-mentioned methods. Arrows between methods highlight how insights from one method informed the data collection of another (ie, *building* [40]). Data interpretation (ie, analysis and integration) is simplified in this diagram and will be discussed in the Data Interpretation section.

XSL•FO
RenderX

Figure 5. Visualization of the embedded mixed methods study design. A predominantly qualitative study embedded with a quantitative method to answer the research question from both a perceived and actual use of the app. CP: clinical placement.



Measures

Actual Use of the App

Log Data

In addition to participants' own assessment of their engagement, log data were recorded of in-app activity to capture actual use of the tool. Participants' sessions, the creation of a learning goal or daily reflection, and the usage of tags were logged. Additionally, participants' learning goals, tasks, daily reflections, and weekly reflections were also stored in a database for both quantitative and qualitative analysis.

For the purposes of determining whether students were still active, in-app activity was monitored on a weekly basis. If students had been inactive for over 2 weeks, they were considered to have stopped using the app.

Reflection and Learning Goal Contents

Contents of the reflections and learning goals were included in the study to better understand how the features were used. This helps to further interpret the log data and possibly distinguish different patterns of use (eg, whether SEC were set as learning goals).

Experience Using the App

Qualitative Survey

To understand participants' experiences with the app, we developed a small survey with seven open-ended questions (refer to Multimedia Appendix 1 for the survey questions). These focused on the daily and weekly reflection features, as well as how students experienced the balance between time and the insights they gained from using the app. This survey was set to be distributed three times: the third, fifth, and seventh

XSL•FO
RenderX

week of the clinical placement. The qualitative survey helped capture the early development of the student's engagement with the tool. Due to a severe technical issue further described in the limitations section, only the first 2 surveys were administered. The first on September 19, 2024, the second on October 10, 2024.

Posttrial Semistructured Interview

Semistructured interviews were conducted online between December 18, 2024, and January 22, 2025. Our goal was to understand the role of the app in participants' learning process during the clinical placement (refer to Multimedia Appendix 2 for the interview schema). Participants were reminded to give their open and honest opinion about the app. As other measures already provided a clearer picture of the actual use and, to an extent, what it was used for, the interviews focused on:

1. Participants' context of learning
   - The clinical placement as learning environment
   - Their perspective on the clinical placement (eg, expectations and motivation)
2. Participants' learning process
   - What skills did they focus on, and how did they approach learning itself?
   - How structured and self-regulated was their process?
3. How did those related to the app
   - How did the context of learning influence app use?
   - What was the role of the app in the participants' learning process?

To be able to compare learning processes of active users and students who did not, or hardly, use the app, students who dropped out during the trial were also invited for this interview. For these cases, the last topic focused on why the app was not, or hardly, used.

Posttrial Focus Group With Study Coaches

After the interviews, a coach-perspective was required to verify whether certain observations were commonly seen in the broader student population, or just our sample. Of our invitation to the entire coaching team of this specific clinical placement, 4 coaches agreed to participate. All gave verbal consent (refer to Multimedia Appendix 3 for the topic list). The focus group was conducted on March 18, 2025, took approximately 45 minutes, and focused on (1) students' changing attitude toward learning during the clinical placement, (2) differences between identified groups of students, and (3) level of challenge of the second-year clinical placement.

Data Interpretation

The goal of mixed methods research is to find meta-inferences through the integration of qualitative and quantitative findings. One strategy to do so is through *warranted assertion analysis* [42], which consists of four steps:

1. Organizing the raw data of each individual data set
2. These data sets are then read as a whole, repeatedly, to come to warranted assertions (ie, meta-inferences)
3. Evidence is collected for each warranted assertion
4. Warranted assertions are refined iteratively by looking for disconfirming evidence

This inductive approach toward mixed methods analysis was chosen as each of our measures shed light on a different aspect of the research question. Although some revealed more than others, they needed to be treated equally valuable for meta-inferences to be drawn. Put in Smith's [42] words:

> *That is, whether data happen to be in the form of words or in the form of numbers should not materially affect the process of constructing meaning from them. Both forms of data are symbolic and neither is inherently "harder" or "softer" than the other.*

We will first report our initial analysis on a measure-to-measure basis. We then describe how we integrated qualitative and quantitative data to reach warranted assertions. The findings section will describe the findings of each measure, after which we discuss our warranted assertions in the Discussion section.

Data Analysis Per Measure

Log data were analyzed quantitatively through descriptive measures (medians, means, and SDs when applicable) to see students' actual usage of the app. Students' activity was quantified through (1) the number of days between first and last use; (2) the number of sessions within that time period; (3) the number of learning goals, daily reflections, and weekly reflections created; and (4) the number of tags and emoticons used (both total and average per session).

Learning goal contents were analyzed to see what topics students used and whether they were SEC or clinical skill-related. This was done by (1) creating an overview of all learning goals, categorized per participant; (2) comparing learning goal phrasing; and (3) cross-referencing it with the corresponding daily reflections to see how the learning goal was used in practice. An additional category (introductory period) was later added to refine the categorization.

During this round, reflection contents were only analyzed to get a first impression of how students used both the daily and weekly reflections. This was done by (1) creating an overview of an individual participant's daily and weekly reflections, (2) reading the reflections, (3) summarizing the participant's approach to reflecting (what fields were used and overall focus of the reflection).

The qualitative survey data were analyzed for perceived ease of use and utility of the daily and weekly reflections. Per question, responses were analyzed using content analysis. These results were then compared between the 2 questionnaires to analyze potential differences.

The interview schema was used as a lens to structure the semistructured interview analysis. Each transcript was analyzed separately to describe each student's individual learning process and how they incorporated the app.

Data Integration

The data integration process consisted of repeatedly engaging with the results of the data analyses and the raw data. Performing the data analysis and comparing the results between measures

XSL•FO
RenderX

was the first step in identifying potential warranted assertions. Then, using the semistructured interview findings as a starting point, these potential warranted assertions were refined, and new ones were found by *importing* [40] the other methods' findings one by one. For example, importing the reflection and learning goal contents helped us understand patterns of use in relation to students' approaches to learning and their learning environment.

Throughout this process, the integration shifted between a student-specific focus and an overarching one. Doing so allowed us to scrutinize the warranted assertions for disconfirming evidence (ie, if a student's learning process did not fit an assertion), resulting in a more nuanced analysis.

Recruitment

Procedure

The sample consisted of Dutch nursing students of the Hanze University of Applied Sciences in the Netherlands. Inclusion criteria were students attending their second year and who were going on their first 20-week-long clinical placement.

Students were recruited using convenience sampling through the following means:

First, a call (and several reminders) was posted on the study program's educational intranet, specifically the clinical placement page where students of our sample were enrolled. It featured a brief description of what the field trial entailed and its goals, a link to sign up, and a reference to the information letter.

Second, an information letter was shared on the same educational intranet page as the call. This letter included an explanation of the goals of the study, what participation entailed, how data would be managed, the potential benefits and risks of participation, and how students could drop out during the study.

Third, an information session was held by the first author (IBB) following a general information session about the clinical placement. In this session, the basic features of the app, the setup of the field trial (eg, what data would be collected and when), and how to sign up were explained.

Finally, the start-up session of the field trial was deliberately planned to follow the main start-up of the clinical placement organized by the study program. This allowed students who had missed the call before the summer break to still sign up. Additionally, as this was the start-up to the curriculum's first clinical placement, this also guaranteed our participants met our inclusion criteria.

The start-up session was given by the first author (IBB), someone the students had no previous relation with and who does not fulfill a role within the curriculum. The session included a reiteration of the study's goals and data collection, an explanation of the features of the app, how students could use it to showcase course-related competencies, and a hands-on support session to help students get the app running on their phones.

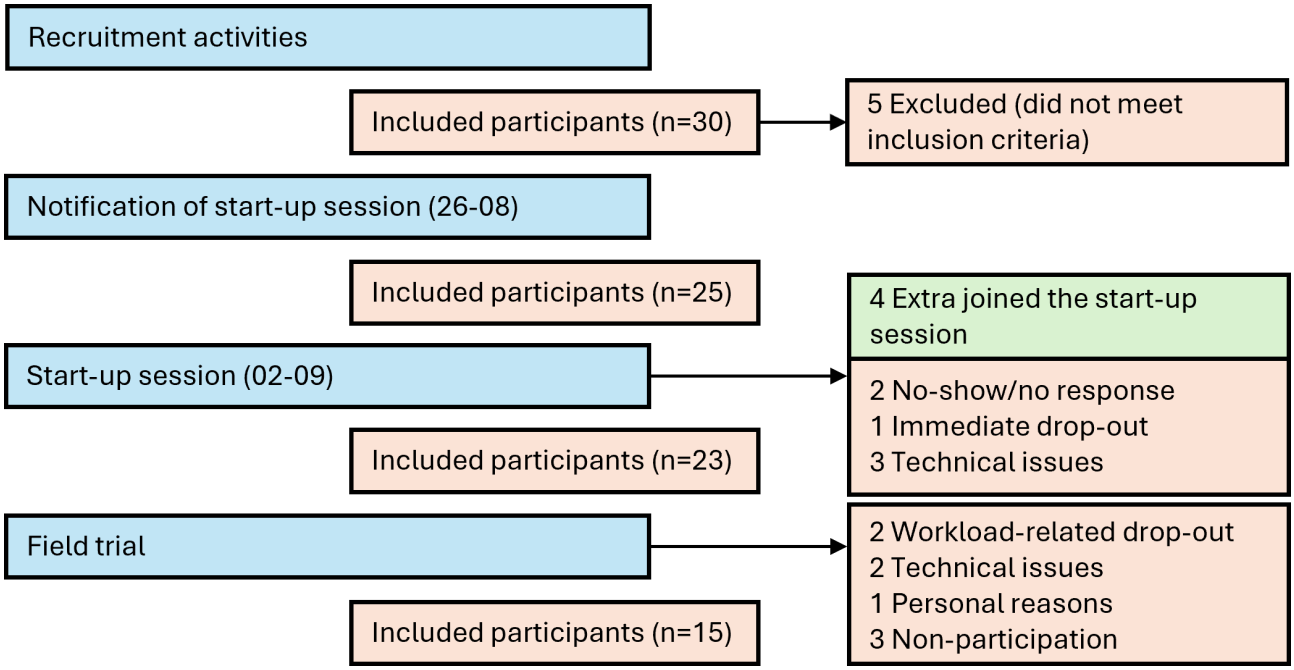
Participants

In total, 34 participants signed up for the study, but only 23 participants started. Moreover, 5 were excluded during the screening process as they were 3rd- or 4th-year students, respectively, doing their second and third clinical placement. In total, 6 participants dropped out at the start-up session.

In the first 3 weeks of the trial, we had 5 more dropouts: 2 due to the workload of the clinical placement, 2 due to unsolvable technical difficulties, and 1 due to personal reasons. Moreover, it became clear near the end of our trial that 3 participants had hardly been active in the app, had not filled in any surveys, and did not respond to our interview request. This brings the number of active participants down to 15 (refer to Multimedia Appendix 4 for the demographic data). Figure 6 gives an overview of the process.

XSL•FO
RenderX

Figure 6. An overview of the recruitment process for the field trial.



Ethical Considerations

A risk to study participation was the added workload of maintaining learning goals and reflections in the app, with no guarantee that this work would aid in passing the clinical placement itself. As the clinical placement is partially graded based on personal growth with regard to several learning outcomes (including 2 personal learning goals), we arranged with the clinical placement coordinator that the activity within the app could be used as evidence for such growth. In doing so, we aimed to lower the risk of participation and ensured that students gained something by participating. We prepared alternative material in case technical errors would occur (same format and questions but in separate Word files).

Ethical approval was obtained from the Hanze Ethical Advisory Committee before the commencement of this study (heac.T2024.032).

Informed consent was obtained at the end of the start-up session. Data were stored on the cloud storage of the researchers' university. As multiple data gathering tools were used, participant data were pseudonymized to allow for data comparison and importation. The translation key was stored in a separate cloud storage which only the lead researcher had access to. The data will be deleted 5 years after the project's completion. Due to low response rates to the first questionnaire, a small financial compensation was added to the study. The reasoning behind this choice is discussed further in the Limitations section.

Results

Actual Use of the App

Log Data

All students used learning goals, with the second most used feature being the daily reflections (14/15, 93.3%). Weekly reflections were barely used (8/15, 53.3%), of which only 1 used it multiple times. Emojis were used by a smaller group (6/15, 40%), with half of those users using at least one emoji per daily reflection. Tags were more broadly used (12/15, 80%; mean 1.76, SD 1.63), but intensity of use varied. Moreover, 3 students used tags and emojis frequently, with on average more than 3 tags and 1 emoji per daily reflection.

App usage varied strongly (mean 46.7, SD 20 d), with 3 outliers who used the app for more than 2 months (not the same 3 as the tag- and emoji-users). Most students used the app for 6 weeks in the first half of their clinical placement (9/15, 60%). No one completed goals, with overall goal use showing 3 patterns: creating all learning goals at the start (4/15, 26.7%), creating learning goals throughout clinical placement (6/15, 40%), or only creating a few at the start (5/15, 33.3%). Due to logging issues, we were unable to accurately determine session lengths.

Reflection and Learning Goal Contents

Qualitatively, learning goal contents show three categories: focusing on clinical skills, the introductory period, and SEC goals (refer to Table 1 for an overview). Only half of the students set SEC goals (n=7), whereas nearly all of them set clinical skill (n=13) and introductory period (n=12) goals.

XSL•FO
RenderX

Table . An overview of the 3 identified goal categories, how many students used them, and illustrative examples.

Goal categories	n	Illustrative goals
Clinical skills	13	• "improving practical skills" [Participant 264] • "independently doing the handover to the next shift" [Participant 451]
Introductory period	12	• "getting to know the patients" [Participant 269] • "getting up to speed with protocols" [Participant 380]
Socioemotional competences	7	• "setting boundaries" [Participant 960] • "clearly and empathically react to and give feedback" [Participant 528]

Reflection contents for clinical skill-focused goals included summaries of executed procedures and self-assessments of what went well or needs to be improved. Similarly, introductory period-focused reflections mainly consisted of self-assessments. SEC-focused reflections were often more elaborate, with a stronger focus on emotions and actions. Examples are illustrated in Table 2.

Table . An overview of the three goal categories, their predominant characteristics, and example quotes.

Goal categories and reflection content	Illustrative quotes
Clinical skills	
Summaries	"I washed people only when asked to do so. I emptied catheters. I helped people go to the toilet. I gave medication to a few people, but it was busy so my colleague did most of it herself. I restocked cupboards with towels myself." [Participant 664]
Self-assessment	"Next time, I should ask if I can do something, like wash their back or face." [Participant 269]
Introductory period	
Self-assessment	"I notice that I am starting to feel comfortable with the various clients and that I can talk with colleagues." [Participant 309]
Socioemotional competences	
Focus on emotions and actions.	"I had to do the caretaking tasks for a lady, after which she gave me one task after another while I was still working on the first one. She also said this in a very commanding tone. After this, I stopped and told her that I can only do one thing at a time and that I am doing my best. It frustrated me and made me feel uncomfortable. After that, she apologized and was much calmer. The colleague who was in the resident's living room also said that I had understood the situation well and responded appropriately. This made me feel very good and confident." [Participant 202]

The tip- and top-fields were used often and similarly across goals. The tip-field often being a small, future-action-oriented reflection (eg, "project more confidence" [Participant 402]) related to the described experience, and the top being a reflection on what went well (eg, "indicated more clearly when I didn't understand something" [Participant 246]).

Experience Using the App

Qualitative Surveys

First Survey

The 9 participants who filled out the first version also filled in the second. Moreover, 6 other participants only filled in the second version. Although this limits our potential to compare changes within individual participants, there is enough data to look at more overarching trends.

In line with the log data, results of the first questionnaire show participants appreciate the daily reflection feature (questions 1, 1a, and 3). Specifically, students mention its ease of use:

> *It brings me a lot. And the questions make it easy, so it takes less time.* [Participant 907]

Additionally, they like how the formatted questions provide focus (eg, tip or top):

> *That you can give tips and tops so that you can clearly see what you are doing right and wrong.* [Participant 246]

Finally, students appreciate that these reflections are brief but still provide small, valuable insights:

XSL•FO
RenderX

> *It helps me process things by writing down what has happened, while typing, I have sometimes come to realize things.* [Participant 202]

Some participants remark that it is hard to remind themselves to use the app after a shift, asking for a reminder feature (question 3):

> *It might be nice to have the option to receive notifications. I notice that I quickly forget about it myself.* [Participant 309]

The few students who have used the weekly reflections were positive (question 2 and 2a):

> *Quite a lot, actually! It's nice to have a slightly clearer structure and to answer the questions.* [Participant 926]

However, the majority of the students mention they still have to try out the weekly reflection. Generally, the first questionnaire shows that about a third of the participants engaged little with the app.

Second Survey

The second survey also shows participants are positive about the daily-reflection feature (questions 1 and 1a). They describe the added value more clearly:

> *I reflect on things I learned that day and things that are quite special, which I would not have thought about without the questions.* [Participant 202]

The week reflections remain underused, even though participants say they are interested. Those that have used it appreciate its reflective format (questions 2 and 2a).

> *Especially step 2 in the weekly reflection (the analysis). By asking follow-up questions about the questions from step 1, you really start to think deeply about why you feel a certain way. The questions also make me feel thankful!* [Participant 451]

About half of the participants state they get value from the overview page of their learning goals. The way the other half uses the app is not supported by the app's design (question 4).

> *Personally, I don't pay much attention to the charts, tips, and tops. This is because I mainly use free writing.* [Participant 163]

> *I hadn't really reflected on learning objectives yet; I had mainly written reflections about the days themselves.* [Participant 363]

The second survey shows participants appreciate the incremental, iterative growth of the daily reflection- and learning goal-features. This supports them to be more mindful of their learning process and goals (question 3):

> *This way, I remember much better what I've done, and I often find that I've done more than I think.* [Participant 664]

Posttrial Semistructured Interviews and Coach Focus Group

In total, 13 participants responded to our call for a posttrial interview, 2 of whom had not used the tool but had enrolled for the study at the start. We specifically included nonusers to better understand the perspective of students who had the opportunity to use the app but chose not to engage with it.

Fading App Usage

Nearly all app-using participants remarked that usage faded over time. A commonality among the reasons given was the transition students experienced from an initially challenging environment to one in which they experienced more confidence in their capabilities and were less in need of ways to structure their experiences. This decreasing urgency is mirrored by the decrease of in-app activity as seen in the log data.

Combining students' accounts of this transition with the reflection contents and learning goal categories, we see one distinct usage pattern which affects this (refer to Table 3 for a joint display). The pattern consists of students who used the app predominantly for clinical skill learning goals by logging performed procedures. They describe the transition as moving from a "learner" to a "worker" attitude or "just doing the work."

XSL•FO
RenderX

Table . Joint display showing students' pattern of use compared with their learning goals per category and an illustrative quote of their reason for quitting the app (this table only includes students who were interviewed).

Pattern of use and participant	Learning goals	Illustrative quote for quitting the app
Clinical skill		
246	• CS[a]=2 • IP[b]=2 • SEC[c]=0	"But after a month, I had a pretty good handle on it, and now it's just work, work, work."
269	• CS=6 • IP=1 • SEC=0	"Yes, I've become so ingrained in that department that it no longer feels like I can really learn much more, so to speak."
451	• CS=1 • IP=1 • SEC=0	"Because I was getting more into the flow of work, I was, like, less focused on the school's learning objectives. It became more like work instead of [learning]."
664	• CS=1 • IP=3 • SEC=0	"It's all pretty much the same. (.) I didn't really do much with it, so I noticed, like, that as the internship went on, I'd start thinking, well, should I write something down again this afternoon, or not? Or should I? Or what am I going to do? And then decided I didn't feel like it."
960	• CS=5 • IP=0 • SEC=5	"Well, it's because I'm really focusing on the new part of my internship. That I'm putting a lot of interest and energy into learning what they do there. So then that extra part about how I should work on myself, I kind of disregarded it."
Socioemotional competences or journaling		
163	• CS=0 • IP=1 • SEC=0	"And at some point you think, if it's been two or three days, 'Oh well, that's not fresh in my memory anymore, is it?' I'll just make sure to do it another time. But then, that 'another time' sometimes never came."
363	• CS=0 • IP=2 • SEC=1	Severe technical issue
402	• CS=1 • IP=1 • SEC=1	Severe technical issue
528	• CS=4 • IP=5 • SEC=5	Severe technical issue
Other		
309	• CS=1 • IP=3 • SEC=0	"I quickly had this intuitive sense of how things were going, so I didn't feel like I needed it."
907	• CS=0 • IP=1 • SEC=0	"I don't know, I just didn't manage to find my way around in the app all the time"

[a]CS: clinical skill.

[b]IP: introductory period.

[c]SEC: socioemotional competences.

> *Yes, I don't think about it as much. That's a shame, of course, because you want to keep pushing forward and keep learning. But over the past few weeks, I've noticed that I'm just working along with the team. I'm thinking less about like, what went well and what could go better.* [Participant 451]

XSL•FO
RenderX

For some of these students, this shift in attitude coincided with a sense of having learned all there is to learn for this clinical placement.

> *Well, I'm trying to challenge myself to work in other departments to see something different. But basically, yes, there are so few nursing procedures that when there were a lot, I could set a learning goal for each one. Yes, I've become so ingrained in that department that it no longer feels like I can really learn much more, so to speak.* [Participant 269]

Most students who used the app for SEC learning goals or as a more personal journal of their learning experience quit due to the severe technical issue. For them, the app provided a consistent place for experiencing a sense of growth and peace of mind, and kept them from forgetting their experiences.

> *Yeah, maybe it's a little weird, but like it gives me some peace of mind or something. Because sometimes you're in the middle of an internship day and think to yourself: why am I doing all of this anyway? And when you're busy in the app, for me it brought a sense of, oh, right, I see now, I have been working on this after all, and I have grown a bit. Or at least stayed consistent with it, so it did give me some peace of mind, like, oh, all of this does have a purpose.* [Participant 528]

> *But in the end, it did provide an overview of everything you've learned. Because I think if you didn't use the app or didn't think about it every day, you'd end up trying to remember everything later on. Like, oh yeah, I think I learned this too. But if you really know what you did each day… Because at the end it all just becomes one big blur, so to speak, now you really know a bit about when you learned what.* [Participant 402]

Consequences of Fading App Usage

For all students, no longer using the app meant their conscious learning shifted to unconscious and, for some, more passive learning. The coaches also remarked on this:

> *They learn a lot in the second half, but they are less aware of it because they are also more focused on working.*

Coaches acknowledged the transition being a common occurrence during clinical placement. Coaches mentioned that a low level of detail in a student's plan of approach, written at the start of the clinical placement, often leads to this passive attitude. Other reasons came down to:

1. Whether the working environment fostered students' learning as opposed to treating them as workload-relief
2. The desire of the individual student to continuously develop their SEC, as the first clinical placement's bar is low in that regard
3. Whether students proactively sought opportunities and activities to learn from, as opposed to a more reactive or passive stance.

In line with the coaches' observation, most participants remarked that no longer using the app also resulted in less awareness of their personal and professional growth.

> *Yes, well, that was indeed the case. In the beginning, I used it a lot, and then the halfway evaluation came, and then I noticed that, yes, at least. In the beginning, I had a clear picture, but those last 2, 3, 4 weeks. It was just like I was working. As if it was just normal for me to do certain things. And only then I thought, yes, some things are actually not normal at all for us to already do.* [Participant 163]

Other Avenues for Reflection

The 2 students who had not used the app seemed to already have a structured learning approach: meaningful situations were noted in a physical journal and reflected upon. Additionally, both could identify and act upon relevant "learning moments" during their shift. This indicates that students who can self-regulate their learning might have less need for this app.

Other students shifted their reflective needs to colleagues or peers, depending on the clinical placement's own learning environment. Like the app, colleagues in a learning-inducive environment could provide students with insight into their learning and progress. However, those discussions rarely focused on SEC.

Due to the severe technical issue, some students chose to use our nonapp alternative, albeit only briefly. The main reasons for students to prefer the app came down to (1) a clearer connection between learning goals and daily experiences, (2) a stronger sense of achievement upon seeing their progress in the app, and (3) the ease of use of a predefined format for reflection.

Discussion

Principal Findings

This study explored how and why nursing students used the Nurse-IT reflection app's features to structure their learning process for developing SEC during clinical placement. Using an embedded mixed methods approach and a naturalistic setting, we were able to capture a detailed and realistic picture of how the app was used and how this usage changed during the clinical placement. We will first discuss how students used the app to structure their learning. We then highlight how that pattern of use differed from our intended use and what we can learn from it. Finally, we discuss students' apparent changing needs during clinical placement and what that means for the design of such reflective applications.

Role of the App in Student Learning

The overarching design goal of the app was to stimulate learner agency through scaffolding a self-regulated learning process (Figure 1). We modeled this learning process by defining *learning goals and tasks* (forethought phase), *daily reflections* (monitoring phase), and *weekly reflections* (self-reflection phase and forethought phase). This section discusses how and why parts of this process supported students in structuring their learning process.

XSL•FO
RenderX

Having the app facilitate a self-regulated learning process alleviated students of the burden of setting up such a process themselves. This is seen in how students valued the app for lowering their mental workload. For example, writing experiences down meant no longer having to remember them. Finally, students remarked how the preformatted questions made it easy for them to engage in goal-setting and reflection. These findings underline the importance of applying persuasive systems design principles when designing such tools [34]. Thus, taking away practical barriers makes engaging in self-regulated learning manageable for students.

Structural engagement with the app made students more aware of their learning process (eg, remembering learning goals during a shift). In turn, this awareness helped them identify learning opportunities and connect those to their learning goals. Other studies focusing on supporting reflection in nursing students show similar effects [12,28]. Survey and interview findings show students valued these effects the most. Overall, this highlights that structurally making one's learning explicit enables students to act on it.

Moreover, and in line with other studies, students who structurally used the app also remarked how its various overviews helped them see and appreciate their growth [12,35]. This boosted confidence and motivated app usage. This effect was best seen in students who used the app as a reflective journal or to capture and evaluate experiences in relation to SEC-related learning goals. In contrast, students who only recorded clinical procedures appreciated their development, but noted it became repetitive and questioned its long-term value. Thus, making one's learning explicit also motivates further action, especially when it is focused on SEC-related learning.

Arguably, this could be achieved through more affordable, physical notebooks. However, students who switched to our nonapp alternative said they missed how the app structured their learning process. For example, the link between learning goals, tasks, and daily reflections, and how the app summarized their growth. As a result, they stopped using the alternative, further underlining the benefits of smartphone-based support systems [29-31]. Thus, empowering students to have an overview and awareness of their learning process scaffolds a self-regulated learning process.

Discrepancy Between Intended Use and Actual Use of the App

Overview

Our intended learning process included all features of the app. However, most students only used the learning goals (forethought phase) and daily reflections (monitoring phase). The weekly reflections were only used sporadically. Given the beneficial effects discussed in the previous section, this raises questions about the necessity of our intended learning process and the potential consequences of how students used the app. We will first discuss the assumptions behind our weekly reflection's design, then relate those to the students' favored pattern of use, the consequences of that use, and its implications for designing reflection apps for clinical placement in nursing programs.

Weekly Reflection Design Assumptions

We designed the daily reflection format to get to surface-level insights, with weekly reflections providing deeper reflection. Thus, although the daily reflections have some self-judgment (a part of the self-reflection phase), we believed the weekly reflections would be necessary for self-reflection.

Aside from depth of reflection, the weekly reflections also dictate a pace for the self-regulated learning cycle. Even though this was not technically enforced within the app, the term "weekly" indicates a pattern.

Finally, the weekly reflections consist of 2 steps—a guided reflection on an experience of the past week (self-reflection phase) and a check-up on one's goals to prepare for the following week (forethought phase). With the latter, the weekly reflections represent the start of a new self-regulated learning cycle. Thus, we assumed students needed scaffolding for reviewing their goals, completing them, and setting new ones.

Students' Pattern of Use and Its Consequences

A key aspect of how students used the app differently lies in how they approached their learning goals. Whereas we intended students to complete them and set new ones, the majority of students only set them at the start. Moreover, no one completed a learning goal. Instead, students set large learning goals, which they worked on throughout clinical placement. This negatively impacted the weekly reflections' relevance.

Consequently, students' self-regulated learning cycle took on a different form: students used the app to monitor their progress on these large goals. Their self-reflection and forethought phases aligned closely with the daily reflections' features: how they are performing (eg, grade and top-field) and whether they need to adjust their approach (eg, tip-field and free writing). Additionally, the interview findings show self-regulated learning also took place *outside* the app (eg, reflecting with colleagues or peers). Overall, this indicates the weekly reflection was not needed for self-regulated learning.

This pattern of use had 2 consequences. First, students with few goals who did not set new goals were prone to quit the app and transition to a working attitude. For these students, only using daily reflections hindered their self-regulated learning. Second, without structural moments of deeper reflection, students potentially stay stuck in superficial learning, ultimately not reaching their learning goals [37].

Lessons for Designing Reflection Apps for Use in Clinical Placement Learning

The key finding in this unintended pattern of use is that supporting students through the self-regulated learning phases does not require in-app, phase-specific activities. In our case, supporting students in remembering and monitoring their learning goals seemed to already enable self-regulated learning. This highlights that students can go through these phases *outside* of the app, and that self-regulated learning is not a strictly linear process [43,44]. It is unsurprising that, given the nature of clinical placement learning, learning cannot be completely scripted. Thus, a reflection app for clinical placement learning

XSL•FO
RenderX

should focus on enabling self-regulated learning and avoid a rigid process of in-app activities.

When it comes to enabling self-regulated learning, our findings show the importance of removing practical barriers. One barrier was forgetfulness. Students wanted reminders for their daily reflections, with some quitting the app after a period of inactivity. Similarly, students could be reminded to review their goals to maintain an intentional learning attitude. Thus, the intensive nature of clinical placement requires reflection apps to reduce the cognitive load of managing one's learning process.

Another barrier ties to our weekly reflection—ease of access. Students who did write weekly reflections found them to be useful but lacked triggers to reflect more often. Those who did not use them were interested yet lacked a trigger to try them. One could address this by facilitating students to write deeper reflections when they realize they want to during the daily reflection. One unimplemented feature was to have the free writing field be interchangeable with reflection frameworks (eg, ALACT by Korthagen et al [45]). Future research into such features is needed to understand how a reflection app can facilitate students' shifting reflection needs.

Students' Changing Needs Throughout Clinical Placement

A final key finding was the decline in app use, with nearly all students no longer using the app after the halfway point (10 wk). Interestingly, other studies with similar app interventions do not report this decline in use while having similar, mostly positive experiences.

One reason could be that those studies featured shorter clinical placement periods, such as 3 weeks [27,28], 6 weeks [20,46,47], or 8 weeks [12]. Yet, some studies that spanned a similar period of time do not report a decline in use [13,31]. From what we can tell, the difference there lies in whether using the tool was a required part of clinical placement learning. Shajani et al [13] digitized an existing, required reflection exercise, and Slade and Downer [31] measured students' perceptions of electronic portfolios already being used in a curriculum. Thus, one reason for our dropout rate could be that it was optional to use.

Looking at the interview and focus group findings, we believe the decline is partially due to students experiencing a transitional phase during clinical placement. Students valued the structure provided by the app in the *initial, tumultuous phase*. Students' reflections and goals reflect this as they focused mostly on learning protocols, how the team operates, and building rapport with patients. Similarly, this is when they learned clinical skills through hands-on experience. Thus, the urgency of the learning goals was high—"I need to function within the team!"—and one's day-to-day growth was more palpable—"I am doing things I could not do last week!"

As this phase ended, *how* students used the app influenced further use. Students whose learning goals focused on clinical skills lacked urgency to further monitor their progress and quit the app. These students often did not manage their learning goals, whereas students with a large set of goals or who used a journaling approach stayed active longer. This indicates certain types of use are more conducive to self-regulated learning and should be advised.

Study coaches remarked that it is harder to get students to critically reflect on their learning process in the second half of the clinical placement. This is seen in how students lost awareness of their learning process and stopped reflecting after app use faded. Moreover, some students missed the oversight and, in hindsight, felt like they would have benefited from further use. This indicates that the work-oriented attitude and its unintentional learning process can hinder self-regulated learning.

The interviewees described clinical placement as being more conducive to learning the starting-phase skills. Most colleagues were willing and capable of helping students reach these goals, for example, through modeling procedures, giving clinical feedback, or explaining protocols. Contrastingly, and in line with other research [15,17], SEC—such as giving feedback or conflict management—were harder to discuss and thus found less uptake by students. Additionally, such skills bear less urgency and are more abstract than, for example, learning a new procedure. Thus, clinical placement as a learning environment and the nature of learning SEC within such an environment also influenced whether students were able to use it for that purpose.

Given the above finding, we believe other intervention components are needed to prepare students for practicing SEC during clinical placement. The app's goal-setting and reflective structure then enables further learner agency and self-regulated learning of SEC. For example, only the few students in our field trial with clear goals regarding SEC managed to work on them. The study coaches acknowledged this, saying students who start with clearer goals are more likely to maintain a learning attitude. Simulation learning could help prepare students to practice with SEC before clinical placement [3,19,48,49]. Subsequently, SEC becomes more recognizable and their urgency clearer, enabling students to set concrete goals and tasks.

Overall, the app's structure fulfills students' early clinical placement learning needs. However, the app's current design and how it was introduced to students do not support the more intentional approach needed for prolonged use. Looking at a recent similar study [12], we believe a stronger integration with the coaching process could help stimulate a more intentional approach (eg, revising or setting new learning goals). Thus, the design and implementation of a reflection app like ours needs to consider these changing needs to reach its intervention goals.

Limitations

Two data collection issues occurred during the field trial. First, we observed low response rates to our first survey. Students struggled to keep up with our intended pace for surveys, as they were overwhelmed by the first weeks of clinical placement. Moreover, the surveys were the only part of the field trial that did not benefit their own learning process. We took two measures to address the low response rates:

1. Financial compensation: Students would receive a gift coupon of €10 (US $11.40) if they filled in at least 2 of the 3 surveys. Although this could introduce response bias, we aimed to mitigate this by framing the incentive as fair

XSL•FO
RenderX

compensation for their time spent on the surveys. Additionally, research shows incentives to be an effective tool to increase response rates with no clear indication of lower response quality [50].

2. More time between surveys: We slowed down the pace of surveys to allow students more time to fill them in.

The second issue was technical and required us to stop the trial in the 13th week. As a result, we could not send out the third qualitative survey nor gather further log data. At that time, there were only 2 active users left, meaning comparatively little log data were lost. Similarly, the third survey would have had a low response rate. Nevertheless, the technical issues did bar us from measuring how students used the reflections to prepare for their clinical placement assessment, a question which could be addressed in future research.

The open setup of the study made it harder for us to evaluate our intended reflective process. We did explain the idea behind the types of reflection and the self-regulated learning cycle as a potential way to use the app. However, we did not require this pattern of use due to the exploratory nature of the study; we were interested in how they themselves would use it. Thus, although our approach led to insights in the variety of ways students used the app in their learning process, our study lacked insight into the potential effect of our intended reflective process.

Similarly, the small sample size, lack of a control group, and design of our study limited our capability to measure the effect of the app on student learning and SEC development. Nevertheless, this means our findings should be interpreted with caution and for their descriptive value of students' experiences with the app and its role in their learning process. Future experimental research should look into the effect of reflection apps on SEC development during clinical placement. Finally, our small sample size and focus on one specific clinical placement in one specific nursing program limit the generalizability of our findings. We addressed this inherent limitation by supporting transferability of our findings to different contexts. For example, transparency in the design of the app and its underlying guidelines helps other researchers understand the intent behind design choices and gives context to our findings. Similarly, the mixed methods approach provides a rich description of the context, which further supports other researchers' understanding of the relevance of our findings to their context. In doing so, we aim to support future research on this topic.

Conclusions and Implications for Practice

The study's exploratory and naturalistic nature allowed us to follow students' developments throughout the clinical placement and the role of the app within it. The findings show that supporting students in structuring and monitoring their clinical placement learning helps them during the initial phase of the clinical placement. Although limited, our study indicates that capturing experiences as opposed to logging clinical procedures is key in using reflection for SEC development purposes. Concurrently, we see that the social norms of the workplace impact students' SEC development, indicating interventions should address this in their design.

Future research should investigate how educational interventions can adapt to students' changing attitudes toward learning as the clinical placement progresses. Additionally, for SEC development specifically, research into multicomponent intervention designs is needed to help students practice these competencies during clinical placement. Our findings can inform designers and nursing educators of the complexities in designing for nursing students' SEC development during clinical placement.

Acknowledgments

We are grateful to Ms Mirjam Palsma for her assistance with the recruitment phase and for allowing the field trial to be conducted during her course. No generative AI was used during the creation of this manuscript.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: IBB (lead), ND (equal), JF (equal)
Data curation: IBB
Formal analysis: IBB
Investigation: IBB
Methodology: IBB (lead), ND (supporting), JF (supporting)
Project administration: IBB
Supervision: ND (lead), JF (equal)
Writing – original draft: IBB (lead), ND (supporting), JF (supporting)
Writing – review & editing: IBB (lead), ND (supporting), JF (supporting)

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Qualitative survey.
[DOCX File, 16 KB - nursing_v9i1e91083_app1.docx]

Multimedia Appendix 2
Interview schema.
[DOCX File, 21 KB - nursing_v9i1e91083_app2.docx]

Multimedia Appendix 3
Coach focus group verbal consent and focus group topic list.
[DOCX File, 18 KB - nursing_v9i1e91083_app3.docx]

Multimedia Appendix 4
Demographic data.
[DOCX File, 17 KB - nursing_v9i1e91083_app4.docx]

References

1. Gandía-Carbonell N, Losilla JM, Viguer P. Strategies to assess and promote the socio-emotional competencies of university students in the socio-educational and health fields: a scoping review protocol. Int J Educ Res 2022;112:101922. [doi: 10.1016/j.ijer.2021.101922]
2. Carvalho VS, Guerrero E, Chambel MJ. Emotional intelligence and health students' well-being: a two-wave study with students of medicine, physiotherapy and nursing. Nurse Educ Today 2018 Apr;63:35-42. [doi: 10.1016/j.nedt.2018.01.010] [Medline: 29407258]
3. Hurley J, Hutchinson M, Kozlowski D, Gadd M, van Vorst S. Emotional intelligence as a mechanism to build resilience and non-technical skills in undergraduate nurses undertaking clinical placement. Int J Ment Health Nurs 2020 Feb;29(1):47-55. [doi: 10.1111/inm.12607] [Medline: 31127972]
4. Afsar B, Cheema S, Masood M. The role of emotional dissonance and emotional intelligence on job-stress, burnout and well-being among nurses. IJISCM 2017;9(2):87. [doi: 10.1504/IJISCM.2017.087952]
5. Soriano-Vázquez I, Cajachagua Castro M, Morales-García WC. Emotional intelligence as a predictor of job satisfaction: the mediating role of conflict management in nurses. Front Public Health 2023;11:1249020. [doi: 10.3389/fpubh.2023.1249020] [Medline: 38026285]
6. Napolitano F, Calzolari M, Di Pietro S, et al. Pedagogical strategies to improve emotional competencies in nursing students: a systematic review. Nurse Educ Today 2024 Nov;142:106337. [doi: 10.1016/j.nedt.2024.106337] [Medline: 39154590]
7. Widad A, Abdellah G. Reflecting on soft skills development sessions: utilizing reflective journaling to enhance nursing students' soft skills. Teaching and Learning in Nursing 2024 Apr;19(2):e344-e349. [doi: 10.1016/j.teln.2023.12.012]
8. Leal LA, Cassiano C, Cruchinho P, et al. Strategical pedagogy for the development of socio-emotional competences in nursing students. Nurs Rep 2024 Oct 8;14(4):2837-2849. [doi: 10.3390/nursrep14040208] [Medline: 39449445]
9. Pires S, Monteiro S, Pereira A, Chaló D, Melo E, Rodrigues A. Non-technical skills assessment for prelicensure nursing students: an integrative review. Nurse Educ Today 2017 Nov;58:19-24. [doi: 10.1016/j.nedt.2017.07.015] [Medline: 28818708]
10. Jamaludin TSS, Nurumal MS, Ahmad N, Muhammad SN, Chan CM. Soft skill elements in structured clinical nursing assessment for undergraduate nursing students: a systematic review. Enfermería Clínica 2021 Apr;31:S58-S62. [doi: 10.1016/j.enfcli.2020.10.019]
11. Armour T, Coffey E, Manias E, Redley B, Nicholson P. Development of mobile educational applications designed for nurses: a narrative review. Nurse Educ Today 2025 Apr;147:106576. [doi: 10.1016/j.nedt.2025.106576] [Medline: 39862477]
12. Johnsen HM, Nes AAG, Haddeland K. Experiences of using a digital guidance and assessment tool (the technology-optimized practice process in nursing application) during clinical practice in a nursing home: focus group study among nursing students. JMIR Nurs 2024 Sep 10;7:e48810. [doi: 10.2196/48810] [Medline: 39255477]
13. Shajani Z, Laing CM, O'Rae A, Burkett J, Peters K. Utilization and perception of a digital clinical tracking tool in undergraduate nursing education. Teaching and Learning in Nursing 2024 Jul;19(3):e494-e499. [doi: 10.1016/j.teln.2024.02.009]
14. Tickle N, Creedy DK, Carter AG, Gamble J. The use of eportfolios in pre-registration health professional clinical education: an integrative review. Nurse Educ Today 2022 Oct;117:105476. [doi: 10.1016/j.nedt.2022.105476] [Medline: 35850091]

XSL•FO
RenderX

15. Thomas J, Jinks A, Jack B. Finessing incivility: the professional socialisation experiences of student nurses' first clinical placement, a grounded theory. Nurse Educ Today 2015 Dec;35(12):e4-e9. [doi: 10.1016/j.nedt.2015.08.022] [Medline: 26358630]
16. Felstead I. Role modelling and students' professional development. Br J Nurs 2013;22(4):223-227. [doi: 10.12968/bjon.2013.22.4.223] [Medline: 23448985]
17. Bril I, Boer HJ, Degens N, Fleer J. Nursing students' experiences with clinical placement as a learning environment for assertiveness: a qualitative interview study. Teaching and Learning in Nursing 2022 Oct;17(4):383-391. [doi: 10.1016/j.teln.2022.04.006]
18. Lanza M, Abajas-Bustillo R, Aparicio-Sanz M, Ortego-Maté C. A mixed methods study examining a socioemotional skills training program based on active learning methodologies in undergraduate nursing students. Nurse Educ Today 2025 Apr;147:106593. [doi: 10.1016/j.nedt.2025.106593] [Medline: 39908892]
19. Widad A, Abdellah G. Strategies used to teach soft skills in undergraduate nursing education: a scoping review. J Prof Nurs 2022;42:209-218. [doi: 10.1016/j.profnurs.2022.07.010] [Medline: 36150863]
20. Aggar C, Shinners L, Penman O, et al. Evaluation of a digital application to support internationally qualified nurses' communication and leadership skills. Nurse Educ Today 2021 Sep;104:104982. [doi: 10.1016/j.nedt.2021.104982] [Medline: 34237627]
21. Horton AG, Gibson KB, Curington AM. Exploring reflective journaling as a learning tool: an interdisciplinary approach. Arch Psychiatr Nurs 2021 Apr;35(2):195-199. [doi: 10.1016/j.apnu.2020.09.009] [Medline: 33781400]
22. Paliadelis P, Wood P. Learning from clinical placement experience: analysing nursing students' final reflections in a digital storytelling activity. Nurse Educ Pract 2016 Sep;20:39-44. [doi: 10.1016/j.nepr.2016.06.005] [Medline: 27428802]
23. Naber J, Markley L. A guide to nursing students' written reflections for students and educators. Nurse Educ Pract 2017 Jul;25:1-4. [doi: 10.1016/j.nepr.2017.04.004] [Medline: 28437690]
24. Miraglia R, Asselin ME. Reflection as an educational strategy in nursing professional development: an integrative review. J Nurses Prof Dev 2015;31(2):62-72. [doi: 10.1097/NND.0000000000000151] [Medline: 25790356]
25. Silva AD, Pinto C, Vieira RM. Mobile devices for developing nursing students' professional skills: scoping review. Rev Lat Am Enfermagem 2024;32:e4371. [doi: 10.1590/1518-8345.7383.4371] [Medline: 39319895]
26. Chen B, Wang Y, Xiao L, et al. Effects of mobile learning for nursing students in clinical education: a meta-analysis. Nurse Educ Today 2021 Feb;97:104706. [doi: 10.1016/j.nedt.2020.104706] [Medline: 33360012]
27. Lai CY, Wu CC. Promoting nursing students' clinical learning through a mobile e-portfolio. Comput Inform Nurs 2016 Nov;34(11):535-543. [doi: 10.1097/CIN.0000000000000263] [Medline: 27315365]
28. Lai CY, Yen YC. Using mobile devices to support cognitive apprenticeship in clinical nursing practice – a case study. ITSE 2018 Dec 10;15(4):348-362. [doi: 10.1108/ITSE-02-2018-0008]
29. Madden K, Collins E, Lander P. Nursing students' perspectives on ePortfolios: themes and preferences compared with paper-based experiences. Int J EPortfolio 2019;9(2):87-96 [FREE Full text]
30. O'Connor S, Andrews T. Smartphones and mobile applications (apps) in clinical nursing education: a student perspective. Nurse Educ Today 2018 Oct;69:172-178. [doi: 10.1016/j.nedt.2018.07.013] [Medline: 30096510]
31. Slade C, Downer T. Students' conceptual understanding and attitudes towards technology and user experience before and after use of an ePortfolio. J Comput High Educ 2020 Dec;32(3):529-552. [doi: 10.1007/s12528-019-09245-8]
32. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development: FT Press; 2014.
33. Zimmerman BJ, Moylan AR. Self-Regulation: Where Metacognition and Motivation Intersect: Handb Metacognition Educ Routledge; 2009.
34. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model and system features. In: Routledge Handbook of Policy Design: Routledge; 2018. [doi: 10.4324/9781351252928-6]
35. Gilheaney Ó, Quigley D. The enablers and barriers to facilitating the development of reflective practice skills of third level allied health professional students through technology: a scoping review. ACS 2022;24(2):125-145. [doi: 10.3233/ACS-210024]
36. Bembridge E, Crowfoot G, Wilson R, Reis J. Digital reflective practice among new graduate nurses - a scoping review. Nurse Educ Today 2025 Jul;150:106688. [doi: 10.1016/j.nedt.2025.106688] [Medline: 40132414]
37. Chan C, Lee K. Reflection literacy: a multilevel perspective on the challenges of using reflections in higher education through a comprehensive literature review. Educational Research Review 2021 Feb;32:100376. [doi: 10.1016/j.edurev.2020.100376]
38. O'Connell TS, Dyment JE. The case of reflective journals: is the jury still out? Reflective Practice 2011 Feb;12(1):47-59. [doi: 10.1080/14623943.2011.541093]
39. Schunk DH. Goal setting and self-efficacy during self-regulated learning. Educ Psychol 1990 Jan 1;25(1):71-86. [doi: 10.1207/s15326985ep2501_6]
40. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Research, 3rd edition: Sage; 2017.
41. Miller D, Bustamante C. Drawing mixed methods procedural diagrams second lang assess mix methods research. In: Moeller AJ, Creswell JW, Saville N, editors. Second Language Assessment and Mixed Methods Research: Cambridge University Press Cambridge; 2016:84-118.

XSL•FO
RenderX

42. Smith ML. Mixing and matching: methods and models. New Drctns Evaluation 1997 Jun;1997(74):73-85. [doi: 10.1002/ev.1073]
43. Azevedo R. Theoretical, conceptual, methodological, and instructional issues in research on metacognition and self-regulated learning: a discussion. Metacognition Learning 2009 Apr;4(1):87-95. [doi: 10.1007/s11409-009-9035-7]
44. Winne PH, Hadwin AF, Gress C. The learning kit project: software tools for supporting and researching regulation of collaborative learning. Comput Human Behav 2010 Sep;26(5):787-793. [doi: 10.1016/j.chb.2007.09.009]
45. Korthagen FAJ, Kessels J, Koster B, Lagerwerf B, Wubbels T. Linking Practice and Theory: The Pedagogy of Realistic Teacher Education: Routledge; 2001. [doi: 10.4324/9781410600523]
46. Garrett BM, MacPhee M, Jackson C. Evaluation of an eportfolio for the assessment of clinical competence in a baccalaureate nursing program. Nurse Educ Today 2013 Oct;33(10):1207-1213. [doi: 10.1016/j.nedt.2012.06.015] [Medline: 22789875]
47. Greviana N, Mustika R, Soemantri D. Development of e-portfolio in undergraduate clinical dentistry: how trainees select and reflect on evidence. Eur J Dent Educ 2020 May;24(2):320-327. [doi: 10.1111/eje.12502] [Medline: 31981383]
48. Roca J, Reguant M, Tort G, Canet O. Developing reflective competence between simulation and clinical practice through a learning transference model: a qualitative study. Nurse Educ Today 2020 Sep;92:104520. [doi: 10.1016/j.nedt.2020.104520] [Medline: 32653810]
49. Pai HC. An integrated model for the effects of self-reflection and clinical experiential learning on clinical nursing performance in nursing students: a longitudinal study. Nurse Educ Today 2016 Oct;45:156-162. [doi: 10.1016/j.nedt.2016.07.011] [Medline: 27518178]
50. Singer E, Ye C. The use and effects of incentives in surveys. Ann Am Acad Pol Soc Sci 2013 Jan;645(1):112-141. [doi: 10.1177/0002716212458082]

Abbreviations

SEC: socioemotional competencies

Edited by E Borycki; submitted 08.Jan.2026; peer-reviewed by B Dovdon, Q Zhou; revised version received 21.May.2026; accepted 17.Jun.2026; published 05.Aug.2026.

Please cite as:
Bril-Broers I, Degens N, Fleer J
Exploring the Potential for Smartphone-Based Reflection to Support Socioemotional Competence Development in Student Nurses' Clinical Placement: Mixed Methods Field Trial
JMIR Nursing 2026;9:e91083
URL: https://nursing.jmir.org/2026/1/e91083
doi:10.2196/91083

© Ivo Bril-Broers, Nick Degens, Joke Fleer. Originally published in JMIR Nursing (https://nursing.jmir.org), 5.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perceptions and Intentions of Nursing Students Regarding Digital Health: Cross-Sectional Study

Alexandre Castonguay[1], PhD; Sandrine Hegg-Deloye[1], PhD; Guy Paré[2,3], PhD; Faustin Armel Etindele Sosso[2,4], DBA

[1]Faculté des sciences infirmières, Université de Montréal, 2375 chemin de la Côte-Sainte-Catherine, Montreal, QC, Canada

[2]Département de technologie de l'information, HEC Montréal, Montreal, QC, Canada

[3]Research Chair in Digital Health, HEC Montréal, Montreal, QC, Canada

[4]Université de Montréal, Montreal, QC, Canada

Corresponding Author:
Alexandre Castonguay, PhD
Faculté des sciences infirmières, Université de Montréal, 2375 chemin de la Côte-Sainte-Catherine, Montreal, QC, Canada

Related Article:

This is a corrected version. See correction statement: https://nursing.jmir.org/2026/1/e98552

Abstract

Background: The integration of digital health technologies (DHTs) in clinical practice is accelerating, creating a need for nursing students to develop digital competencies aligned with professional expectations. In Quebec, curricular reforms aim to enhance digital health literacy, but data are limited on students' preparedness.

Objective: This study aimed to assess nursing students' perceptions, self-reported competencies, and willingness to engage with DHTs across different academic years.

Methods: A cross-sectional descriptive survey assessing self-reported digital health competencies, attitudes, perceived training coverage, and intentions was conducted using an online questionnaire administered through Qualtrics. Participants (N=136) were recruited from 3 cohorts: first-year (group 1; n=58, 42.6%), second-year (group 2; n=55, 40.4%), and third-year (group 3; n=23, 16.9%) nursing students. Data were analyzed using descriptive statistics and ANOVAs, with post hoc analyses performed via SPSS.

Results: Significant differences were observed among cohorts concerning digital competencies and access to digital tools. Compared with first-year students (group 1), third-year students (group 3) showed higher proficiency with electronic medical records (group 3: mean 3.29, SD 1.31; group 1: mean 2.59, SD 1.32; P=.01), virtual reality (group 3: mean 4.53, SD 1.11; group 1: mean 2.90, SD 1.44; P<.001), and clinical databases (group 3: mean 4.59, SD 1.00; group 1: mean 3.21, SD 1.55; P<.001). Despite positive attitudes toward DHTs across all groups, training coverage for most digital tools was perceived as low, with the highest levels reported for clinical databases (mean 2.97, SD 1.1). This underscored a substantial gap between institutional expectations and actual digital training across all cohorts.

Conclusions: This study highlights critical gaps in digital health training among nursing students, emphasizing the need for targeted curricular reforms such as the one currently underway at the Université de Montréal. These efforts represent a promising opportunity to better align educational content with the evolving demands of health care systems. Today, preparing students in digital competencies is no longer just advantageous but may soon become essential for the next generation of nurses to navigate and lead within technology-driven care environments.

(JMIR Nursing 2026;9:e77051) doi:10.2196/77051

KEYWORDS

health IT; digital health education; eHealth competencies; technology adoption in nursing; informatics competency; nursing education; health information systems

Introduction

Macro-Level Context: Digital Shift in Health Care

The digital transformation of health care systems is rapidly redefining how care is delivered, documented, and managed worldwide. Across clinical environments, digital health technologies (DHTs)—including electronic medical records (EMRs), virtual care platforms, and artificial intelligence—are increasingly deployed to enhance care coordination, improve efficiency, and respond to growing service demands [1,2]. In

XSL•FO
RenderX

Quebec, this shift is occurring amid a persistent nursing shortage that places additional pressure on health care institutions and personnel [3]. Strategic investments in digital infrastructure, such as those outlined in the province's 2022 to 2025 technological modernization plan and the federal Pan-Canadian Health Data Strategy, have positioned digital tools as key enablers of more responsive, equitable, and sustainable care delivery [4,5].

Professional Standards and Competencies

To ensure that nurses are prepared to operate within these digitized environments, regulatory and educational bodies have issued guidelines calling for the integration of informatics into nursing practice. The Ordre des infirmières et infirmiers du Québec (OIIQ) emphasizes the importance of digital competencies for safe, ethical, and efficient care across both in-person and virtual settings [6,7]. At the national level, the Canadian Association of Schools of Nursing (CASN) has updated its National Nursing Education Framework to include explicit expectations for digital health literacy among nursing graduates [8]. Together, these standards promote the incorporation of information and communications technologies into nursing curricula as both a professional requirement and a pedagogical priority.

Educational Challenges in Nursing Programs

Despite these mandates, gaps remain in the actual delivery of digital health training across Canadian nursing programs. While students often demonstrate general digital literacy, studies show that they are underprepared to use clinical information systems, navigate electronic records, or interpret data generated by digital tools in real-world practice [9-11]. Furthermore, many programs offer inconsistent or superficial exposure to health ITs (HITs), contributing to low confidence and proficiency among graduates [11]. These limitations risk widening the gap between institutional expectations and actual clinical readiness, particularly as the pace of digitalization continues to accelerate. Previous studies indicate that, although nursing students tend to have a positive attitude toward digital health, their clinical readiness remains limited due to inconsistent exposure to health information systems. Competency frameworks such as those from the CASN and OIIQ underscore the need for structured, progressive integration of digital training into nursing education [7,8]. These limitations contribute to low confidence and perceived readiness among graduates entering technology-rich clinical environments.

Study Rationale in the Quebec Context

Although several studies have examined nurses' perceptions of digital health tools, few have focused on nursing students in Quebec, where the educational structure, regulatory environment, and pace of curricular reform differ from those of other Canadian provinces [12]. Given recent efforts at the Université de Montréal (UdeM) to revise its undergraduate nursing program to better align with CASN and OIIQ guidelines, it is crucial to understand students' current perceptions, competencies, and intentions regarding DHTs. To our knowledge, this is the first study in Quebec to use a cohort-comparative design to examine digital health preparedness among nursing students. Grounded in the technology acceptance model and the OIIQ digital competency framework, this study aimed to identify year-based differences in perceptions and training gaps, thereby informing targeted curricular reforms to support Quebec's digital health transformation [12,13].

This study was conceptually inspired by the behavioral framework by Paré et al [14] examining medical students' intention to integrate DHTs, which combines elements from the theory of interpersonal behavior by Triandis [15] and the technology acceptance model [16]. The survey content was further adapted to the nursing context using insights from the 2020 National Survey of Canadian Nurses on the use and impact of DHTs in practice. The final instrument was reviewed and validated in collaboration with experienced nursing professors from the Faculty of Nursing (Faculté des sciences infirmières in French; FSI) at UdeM to ensure relevance and clarity for nursing students.

Research Objectives

This study had the following objectives:

- Explore the perceptions and intentions of nursing students regarding their knowledge, experiences, and training related to HITs and their beliefs, attitudes, and intentions concerning the use of HITs in their future professional practice
- Compare these perceptions and competencies based on students' year of study to identify potential differences across cohorts
- Provide recommendations for improving HIT-related education in response to the identified needs

Methods

Design, Setting, and Participants

This study used a cross-sectional descriptive design. It was conducted at the FSI at UdeM in May 2023. The target population included 1274 students enrolled in the FSI at UdeM. A target sample size of 300 was chosen based on recommended minimum thresholds for descriptive surveys in large student populations [17] to ensure reasonable representation as no formal power calculation was required for this exploratory design. Eligible participants were all students currently enrolled in an undergraduate nursing program who provided informed consent.

It is important to note that participant distribution across academic years was uneven. This reflects actual enrollment sizes within the program during the data collection period.

Ethical Considerations

This study was reviewed and approved by UdeM's Research Ethics Committee in Sciences and Health (2023-4300) and was supported by the FSI. All participants provided informed consent electronically before starting the survey and could withdraw at any time without consequence. Data were collected anonymously to ensure participant privacy and confidentiality. No financial or other compensation was provided for participation.

XSL•FO
RenderX

Recruitment

UdeM was selected as the research site due to its status as a major higher education institution in Quebec and the diversity of its undergraduate nursing student population. Recruitment was conducted electronically through the FSI, which emailed an initial invitation and one reminder to all eligible students via their institutional addresses. Interested participants accessed the study by clicking a secure link leading to the consent form and the online questionnaire. Participation was voluntary and anonymous. Although IP addresses were not collected to ensure confidentiality, the use of institutional email distribution and limited reminders minimized the risk of multiple entries.

Data Collection

Data were collected through an online self-administered questionnaire designed specifically for this study administered via Qualtrics (Qualtrics International Inc), a platform designed for web-based research that offers basic descriptive statistics tools while ensuring secure data handling.

The questionnaire took approximately 15 minutes to complete and included an open-ended question at the end, allowing participants to share additional comments. Although this open-ended question was included at the end of the questionnaire to capture additional student perspectives, these responses were not systematically analyzed in this study. An exploratory review of these responses was conducted to inductively identify recurring themes and enrich the interpretation of the quantitative findings by reflecting participants' lived experiences and unprompted perspectives.

The survey was developed to operationalize key constructs from the conceptual framework described above. Specific items were adapted from the study by Paré et al [14] and the 2020 National Survey of Canadian Nurses to align with the nursing context and relevant DHTs. The draft instrument was then reviewed and refined in collaboration with experienced nursing faculty at UdeM to ensure content validity and clarity. The final version included 15 items (Table S1 in Multimedia Appendix 1) and is available in Multimedia Appendix 2.

Variables

This study assessed nursing students' perceptions, competencies, and training needs related to HITs across 4 main categories of variables.

Sociodemographic and Academic Variables

Sociodemographic and academic variables included the year and type of academic program; the participants' age, sex, and gender; languages of interaction; daily use of digital devices; previous professional experience in the health care sector; and current or intended professional setting after graduation.

HIT Competencies and Training

Using Likert-type scales, participants assessed their level of proficiency with various technologies (1="none," 2="very low," 3="low," 4="moderate," 5="high," and 6="very high"), the extent to which HITs were covered in their nursing education program (1="not covered at all," 2="poorly covered," 3="moderately covered," 4="well covered," and 5="very well covered"), and the level of expertise they believed was necessary for clinical practice (1="no training needed," 2="basic training," 3="functional or intermediate training," and 4="expert or specialized training).

Perceived Impacts of HITs

Participants were asked to rate the perceived impact of DHTs on key aspects of the nursing profession, such as quality of care and work life. The question was phrased as follows: "Do you believe that the use of digital health technologies has, or will have, an impact on the following dimensions of the nursing profession?" Responses were collected on a Likert scale (1="very negative impact," 2="negative impact," 3="neutral impact," 4="positive impact," and 5="very positive impact"). An open-ended question followed, allowing respondents to elaborate or provide additional comments.

Statistical Analysis

Students were grouped into 3 categories based on their year of study: first year of the integrated diploma of college studies–bachelor's degree (*diplôme d'études collégiales–baccalauréat* in French; DEC-BAC) and regular bachelor's degree (*baccalauréat* in French; BAC) programs (group 1), second year of the DEC-BAC and BAC programs (group 2), and third year of the DEC-BAC and BAC programs (group 3). All Likert scale variables were treated as continuous variables and analyzed using means and SDs.

For categorical variables, frequencies and percentages were reported for each group as well as for the total sample. Group comparisons were conducted using exact *P* values from chi-square tests. For continuous variables, valid sample sizes, means, and SDs were presented. Group means for Likert scale variables were compared using one-way ANOVA followed by Tukey post hoc tests where appropriate.

An exception was made for the variable "training coverage—scheduling software," where the Welch correction for unequal variances was applied and post hoc comparisons were conducted using Games-Howell tests. All statistical analyses were performed using the SPSS software (version 28; IBM Corp), with a significance level set at 5%.

Results

Sociodemographic and Academic Variables

Overview

A total of 173 students accessed the questionnaire. The response rate for the item identifying participants' academic year within the nursing program was 78.6% (136/173). After grouping, group 1 comprised 42.6% (58/136) first-year students, group 2 comprised 40.4% (55/136) second-year students, and group 3 comprised 16.9% (23/136) third-year students.

Participant Profile

Across all groups, more than 67% (92/136) of participants were aged between 18 and 25 years, with group 3 standing out at 80% (19/23) in this age range (*P*=.03). This group also reported significantly higher access to a laptop (*P*=.02). Use of digital tools for specific activities was also significantly greater among

XSL•FO
RenderX

groups 2 and 3 (*P*=.01). Additionally, a higher proportion of group 3 participants reported working in the provincial public health care sector (*P*=.003; Table S2 in Multimedia Appendix 1).

HIT Competencies and Training

HIT Proficiency Levels

Proficiency levels across all work-related HITs were generally rated from low to moderate. Group 2 stood out in terms of proficiency with EMRs (*P*=.01), whereas group 3 demonstrated higher proficiency with virtual reality (*P*=.001) and clinical databases (*P*=.001; Table S3 in Multimedia Appendix 1).

Training Coverage by Group

Regardless of the specific HIT in question, coverage within the curriculum was generally perceived as very limited to minimal. Group 1 reported greater coverage in the use of scheduling software (*P*=.007), whereas group 3 reported more extensive coverage in database training (*P*=.004; Table S4 in Multimedia Appendix 1).

Required Expertise for Professional Use of HITs

Participants generally believed that most HITs required little to no specific expertise for professional use. An exception was noted for group 2, which reported a significantly higher need for basic or intermediate training in the use of health care robots (*P*=.02; Table S5 in Multimedia Appendix 1).

Perceived Impacts of HITs in the Nursing Profession

Perceptions of the impact of HITs on the nursing profession were consistent across all groups. Most students believed that DHTs have a positive or very positive impact on key aspects such as quality of care, therapeutic relationships, and productivity (Table S6 in Multimedia Appendix 1).

Narrative Comments

Of the 136 respondents, 31 (22.8%) provided written responses to the open-ended question. Most comments reflected key patterns observed in the quantitative results, particularly regarding the perceived lack of formal training in DHTs. Several students reported learning to use digital tools only during clinical placements, consistent with the limited curriculum coverage ratings found across cohorts. One respondent noted that "[w]e only learn to use digital tools during internships," whereas another remarked that "[a]t university, we have no training on the software we encounter in practice."

Others described barriers to accessing technologies during clinical placements due to their student status:

> *As a nurse it's fine, but as a student, it's complicated when we don't have the access.*

These comments aligned with cohort-based differences observed in digital tool use and confidence levels.

However, the responses also surfaced dimensions not captured in the structured survey. Several participants mentioned psychosocial or physical impacts of digital tool use, including "less sleep, eye strain, and a more sedentary lifestyle" and "technology sometimes limits social interaction, which affects mental health." A few students proposed improvements, such as earlier exposure to clinical systems or greater consistency in tools across health care settings:

> *It would help to learn some of the clinical software in advance.*
>
> *Please standardize the digital documents, it would simplify things when moving between institutions.*

While broadly aligned with the quantitative trends, these narrative comments introduced experiential nuances, such as well-being, autonomy, and implementation barriers, that were not accessible through the Likert scale items alone.

Discussion

Principal Findings

Both the OIIQ Comité jeunesse and the OIIQ [6,12] emphasize the importance of access to IT that supports nurses' clinical decision-making, optimizes care processes, and enhances patient safety, as well as maintaining professional and ethical standards.

Similarly, the CASN, through its National Nursing Education Framework [8], underscores the importance of integrating nursing informatics competencies into academic programs to better prepare future professionals for the increasing technological demands of clinical environments.

However, our findings indicate that students perceive their current training as insufficient. Their self-assessed proficiency with digital tools was limited, and coverage of key HITs was generally perceived as weak without a clear progression across academic years despite students showing an interest in these tools and their potential impact on nursing practice.

These observations suggest a concerning disconnect between the evolving expectations of the health care system and students' perceived preparedness despite the clear guidelines set forth by both the OIIQ and CASN. This aligns with findings by Kleib et al [11], who identified persistent gaps in digital health education that were exacerbated by pedagogical approaches that often overestimate students' technological competencies and fail to adequately prepare them for clinical realities.

The narrative comments reinforced this disconnect by providing concrete examples of students' perceived lack of preparation, particularly in relation to clinical software. The comments also highlighted barriers that were not captured in the structured survey, such as restricted access to digital tools during placements and the personal toll of technology use on well-being. These experiential accounts complement the quantitative findings and emphasize that perceived readiness is shaped not only by curricular content but also by the conditions of access, support, and emotional experience within clinical environments.

The fact that many participants in our study believed that high-level proficiency was unnecessary for the effective use of HITs may reflect a limited understanding of their relevance in nursing practice. Kleib et al [11] emphasize that, while students often demonstrate basic digital literacy, this does not necessarily translate into clinical competence. Consequently, students may

XSL•FO
RenderX

underestimate the complexity and significance of digital tools in real-world settings. These findings highlight the pivotal role of nursing educators in bridging this gap by explicitly integrating digital competencies into nursing curricula and grounding them within authentic clinical contexts to foster both awareness and applied proficiency.

Despite these challenges, the surveyed students remained generally optimistic about the role of HITs, with most believing that these technologies will have a positive or even very positive impact on the nursing profession. This optimism reflects a growing awareness of the digital shift in health care and aligns with the strategic directions promoted by the OIIQ and CASN, who advocate for stronger digital health training in nursing education [7,8].

Beyond these overarching findings, notable differences emerged between academic years in terms of access to digital tools, proficiency with specific technologies, and perceived curriculum coverage. This study identified several statistically significant differences in how students from each cohort perceived and interacted with HITs.

The 18- to 25-year age group predominated across all cohorts, particularly in group 3, where it represented 83% (19/23) of the participants. Group 3 students also reported greater access to digital tools, especially laptops, and more advanced use of EMRs, virtual reality, and databases. These findings are consistent with the broader and more in-depth training coverage that this group reported for these tools.

The observed differences in competencies and training exposure likely reflect curricular progression, with third-year students benefiting from more extensive clinical placements, capstone projects, and practical assignments that integrate digital tools. In addition, group 3 students may have greater opportunities for independent learning and part-time work in health care settings, reinforcing their proficiency with EMRs and clinical databases. In contrast, group 1 students are typically at the beginning of their program, focused on foundational theoretical content with limited hands-on exposure to applied digital health systems.

Group 2 students stood out for their more frequent use of digital tools and a higher reported need for advanced competencies in health care robotics. Meanwhile, group 1 students reported better coverage related to scheduling software, likely reflecting academic needs specific to first-year studies.

Taken together, these findings reveal a persistent and concerning gap between current digital health needs and the perceived preparedness of nursing students, echoing the concerns expressed by the OIIQ [7]. They underscore the importance of continuing efforts to integrate digital health competencies into nursing education in a structured and strategic manner.

Limitations and Future Research

Methodological and Contextual Limitations

Of the 1274 targeted students, 136 (10.7%) completed the questionnaire, representing 45.3% (136/300) of the sample size required to ensure statistical representation. While the target of 300 respondents was chosen based on practical guidelines for descriptive surveys, the absence of a formal power calculation means that the study may be underpowered to detect smaller effect sizes or subtle group differences, particularly given the uneven distribution across cohorts.

The unequal sample sizes between academic years may also have limited the statistical power for some comparisons. However, this limitation mirrors real enrollment patterns and was accounted for using appropriate statistical analyses.

Although no data were available to confirm a nonresponse bias, the relatively low participation rate raises the possibility of such a risk. Therefore, it is possible that respondents had characteristics or perceptions that differed from those of nonrespondents, potentially limiting the generalizability of the findings.

Additionally, the self-administered nature of the questionnaire introduces the possibility of self-selection and self-assessment bias. Students who chose to participate may have been more comfortable with technology or more aware of digital health issues, potentially skewing responses toward more favorable or more critical perceptions than would be typical. Moreover, as the study was conducted at a single institution, UdeM, the transferability of the findings to other educational, institutional, or cultural contexts is limited. The uneven distribution of participants across groups, particularly the low number of participants in group 3, further reduces the robustness of statistical comparisons and may have hindered the detection of significant effects in certain analyses.

In addition, while this study included an open-ended question that generated valuable narrative insights, these data were reviewed only through an exploratory, unsystematic process and were not coded or analyzed using a formal qualitative methodology. As such, they cannot be considered representative or exhaustive. Even so, the added value of these responses, particularly in revealing experiential and emotional dimensions absent from the structured survey, suggests that future research would benefit from an intentionally designed mixed methods approach. Incorporating qualitative inquiry from the outset guided by a clear theoretical framework would allow for a more comprehensive understanding of how nursing students engage with digital technologies across educational and clinical contexts.

Future Research Directions

The FSI at UdeM is currently in the process of implementing a revised version of its undergraduate program in response to recent recommendations from the CASN that emphasize the importance of developing digital competencies aligned with current clinical demands and realities.

In this context, a pretest-posttest design with independent cohorts comparing students who completed the old curriculum with those trained under the revised program would offer valuable insights into the impact of the reform on digital health literacy, integration, and appropriation. Additionally, consultations with key stakeholders within the faculty, including teaching staff, program directors, educational advisors, and student representatives from each cohort, could provide critical institutional perspectives on the validity of the findings

XSL•FO
RenderX

presented in this study. As part of a complementary qualitative component, this triangulated approach would not only contextualize the results but also help assess how curriculum reform may influence digital health teaching practices.

Several additional methodological avenues could be explored to build on this work and deepen understanding of students' perceptions and competencies regarding HITs. First, replicating the study across multiple nursing schools through a multisite comparative design would help assess the transferability of findings beyond the UdeM context. Second, adopting a mixed methods approach to combine quantitative questionnaires with qualitative interviews could yield a more nuanced understanding of student expectations, perceived barriers, and enabling conditions for meaningful engagement with digital tools. Finally, longitudinal studies conducted throughout the academic journey would allow for documentation of digital competency development over time in relation to different education milestones and ongoing curricular adjustments.

Conclusions, Implications, and Recommendations

This study reveals a significant gap between institutional expectations and perceptions of digital health training by nursing students, particularly in their proficiency with essential tools such as EMRs and clinical databases. Despite recognizing the value of digital technologies, students reported limited curricular coverage and confidence in using them.

These findings call for targeted reforms in nursing education, including the integration of hands-on training, digital simulations, and interactive modules. To address these gaps more concretely, nursing programs could implement several specific measures: (1) dedicated simulation laboratories focused on EMRs and clinical software; (2) mandatory digital health literacy modules that introduce students to national and provincial informatics frameworks; and (3) interprofessional digital health workshops that mirror real-world collaborative settings involving nurses, physicians, and allied health professionals. These recommendations are further supported by students' own accounts, which reflect both a desire for earlier, more practical exposure to digital tools and a recognition of the limitations in their current training environment.

The current curriculum revision at UdeM offers a timely opportunity to evaluate the implementation outcomes of such changes. Future studies using a clearly defined pretest-posttest design, with precise cohort tracking to distinguish students enrolled before, during, or after curriculum reforms and collecting detailed data on external exposure to HITs (eg, work settings) could help assess the true impact of these reforms.

Broader research across institutions and through mixed methods or longitudinal approaches will be crucial to strengthening the evidence base literature. In an increasingly technology-driven health care landscape, equipping the next generation of nurses with robust digital competencies is no longer optional and has become a professional imperative.

Acknowledgments

The authors thank the Statistical Consulting Service at the University of Montreal for their assistance with the study design and data analysis. They also appreciate the support of the Faculty of Nursing in coordinating the project and facilitating participant recruitment. During the preparation of this manuscript, GPT-4o, a generative artificial intelligence (AI) language model developed by OpenAI, was used via the ChatGPT interface to assist with translation, linguistic revision, structural refinement, and clarity enhancement. The authors also used ChatGPT to support narrative refinement and thematic clarity during manuscript preparation, consistent with emerging evidence on the constructive role of generative AI in academic meaning making and reflective synthesis [18]. All outputs generated through AI were carefully reviewed, edited, and approved by professionals and the authors to ensure accuracy, appropriateness, and alignment with the responsible use of generative AI in academic writing [19]. The final version of the manuscript was then reviewed by a professional proofreader to ensure clarity and consistency.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sector.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy and institutional restrictions but are available from the corresponding author on reasonable request.

Authors' Contributions

AC conceptualized the study and led project administration. AC and GP developed the methodology and supervised the study. SH-D curated the data and performed the formal analysis. FAES conducted validation. AC and SH-D wrote the original draft of the manuscript. All authors (AC, SH-D, GP, and FAES) reviewed and edited the manuscript and approved the final version.

Conflicts of Interest

AC serves as an associate editor for the *Journal of Medical Internet Research*. The other authors declare no conflicts of interest.

Multimedia Appendix 1

XSL•FO
RenderX

Descriptive and comparative tables summarizing participant characteristics, technology proficiency, training coverage, and perceived expertise requirements across student groups.
[DOCX File, 37 KB - nursing_v9i1e77051_app1.docx]

Multimedia Appendix 2
Study questionnaire assessing nursing students' perceptions, experiences, and attitudes toward digital health technologies.
[DOCX File, 50 KB - nursing_v9i1e77051_app2.docx]

References

1. Kraus S, Schiavone F, Pluzhnikova A, Invernizzi AC. Digital transformation in healthcare: analyzing the current state-of-research. J Bus Res 2021 Feb;123:557-567. [doi: 10.1016/j.jbusres.2020.10.030]
2. Conseil international des infirmières, transformation numérique de la santé et pratique infirmière. Conseil International des Infirmières. 2023. URL: https://www.icn.ch/sites/default/files/2023-08/ICN%20Position%20Statement%20Digital%20Health%20FINAL-30.6_FR.pdf [accessed 2025-05-25]
3. Rapport du groupe de travail national sur les effectifs infirmiers. Ministère de la Santé et des Services sociaux. 2022. URL: https://publications.msss.gouv.qc.ca/msss/document-003387/ [accessed 2025-05-25]
4. Plan de modernisation technologique 2022-2025 du réseau de la santé et des services sociaux. Ministère de la Santé et des Services sociaux. 2022. URL: https://msss.gouv.qc.ca/ministere/salle-de-presse/sujet-323/ [accessed 2025-05-25]
5. La stratégie pancanadienne de données sur la santé: rapports et sommaires des comité consultatif d'experts. Agence de la santé publique du Canada. 2022. URL: https://www.canada.ca/fr/sante-publique/organisation/mandat/a-propos-agence/organismes-consultatifs-externes/liste/strategie-pancanadienne-sante-rapports-sommaires.html [accessed 2025-05-25]
6. Technologies de l'information et de la communication (TIC) et télésanté: les mêmes devoirs et obligations professionnels et déontologiques s'appliquent, que l'on intervienne en personne ou à distance. Ordre des infirmières et infirmiers du Québec. 2023. URL: https://www.oiiq.org/pratique-professionnelle/encadrement-de-la-pratique/sante-numerique [accessed 2025-05-25]
7. Ordre des infirmières et infirmiers du québec, rapport annuel 2023-2024. Ordre des infirmières et infirmiers du Québec. 2024. URL: https://www.oiiq.org/documents/20147/237836/RA-2023-24-web.pdf [accessed 2025-05-25]
8. Cadre national de la formation infirmière. Canadian Association of Schools of Nursing. 2022. URL: https://www.casn.ca/fr/competences-et-indicateurs/cadre-national-de-lacesi-sur-la-formation-infirmiere/ [accessed 2025-05-25]
9. Sondage national des infirmières et infirmiers du canada 2020: utilisation des technologies de santé numériques au travail. Inforoute Santé du Canada. 2020. URL: https://www.infoway-inforoute.ca/fr/component/edocman/ressources/rapports/evaluation-des-avantages/3813-sondage-national-des-infirmieres-et-infirmiers-du-canada-2020-utilisation-des-technologies-de-sante-numeriques-au-travail [accessed 2025-05-25]
10. Ali S, Kleib M, Paul P, Petrovskaya O, Kennedy M. Compassionate nursing care and the use of digital health technologies: a scoping review. Int J Nurs Stud 2022 Mar;127:104161. [doi: 10.1016/j.ijnurstu.2021.104161] [Medline: 35032743]
11. Kleib M, Arnaert A, Nagle LM, et al. Digital health education and training for undergraduate and graduate nursing students: scoping review. JMIR Nurs 2024 Jul 17;7:e58170. [doi: 10.2196/58170] [Medline: 39018092]
12. La santé numérique: regard critique de la relève infirmière. Ordre des infirmières et infirmiers du Québec. 2024. URL: https://www.oiiq.org/documents/20147/12996513/avis-cj-sante-numerique-20240118-vf.pdf [accessed 2025-05-25]
13. Rouidi M, Elouadi AE, Hamdoune A, Choujtani K, Chati A. TAM-UTAUT and the acceptance of remote healthcare technologies by healthcare professionals: a systematic review. Inform Med Unlocked 2022;32:101008. [doi: 10.1016/j.imu.2022.101008]
14. Paré G, Raymond L, Pomey MP, Grégoire G, Castonguay A, Ouimet AG. Medical students' intention to integrate digital health into their medical practice: a pre-peri COVID-19 survey study in Canada. Digit Health 2022;8:20552076221114195. [doi: 10.1177/20552076221114195] [Medline: 35898286]
15. Triandis HC. Values, attitudes, and interpersonal behavior. Nebr Symp Motiv 1980;27:195-259. [Medline: 7242748]
16. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
17. Firdion JM. Construire un échantillon. In: Quadrige SP, editor. L'enquête Sociologique: Presses Universitaires de France; 2012:69-92. [doi: 10.3917/puf.paug.2012.01.0069]
18. Blyler AP, Seligman ME. Personal narrative and stream of consciousness: an AI approach. J Posit Psychol 2024 Jul 3;19(4):592-598. [doi: 10.1080/17439760.2023.2257666]
19. Leung TI, de Azevedo Cardoso T, Mavragani A, Eysenbach G. Best practices for using AI tools as an author, peer reviewer, or editor. J Med Internet Res 2023 Aug 31;25:e51584. [doi: 10.2196/51584] [Medline: 37651164]

Abbreviations

BAC: baccalauréat

XSL•FO
RenderX

CASN: Canadian Association of Schools of Nursing
DEC-BAC: diplôme d'études collégiales–baccalauréat
DHT: digital health technology
EMR: electronic medical record
FSI: Faculty of Nursing
HIT: health information technology
OIIQ: Ordre des infirmières et infirmiers du Québec
UdeM: University of Montreal

Edited by K Cato; submitted 06.May.2025; peer-reviewed by G Davlyatov, SA Mukhtar; revised version received 27.Jun.2025; accepted 20.Jan.2026; published 05.Mar.2026.

Please cite as:
Castonguay A, Hegg-Deloye S, Paré G, Etindele Sosso FA
Perceptions and Intentions of Nursing Students Regarding Digital Health: Cross-Sectional Study
JMIR Nursing 2026;9:e77051
URL: https://nursing.jmir.org/2026/1/e77051
doi:10.2196/77051

© Alexandre Castonguay, Sandrine Hegg-Deloye, Guy Paré, Faustin Armel Etindele Sosso. Originally published in JMIR Nursing (https://nursing.jmir.org), 5.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Psychometric Evaluation of the Canadian Nurse Informatics Competency Assessment Scale and the Digital-Technology Self-Efficacy Scale Among Saudi Nursing Students: Cross-Sectional Study

Nader Alnomasy[1*], PhD; Habib Alrashedi[2*], PhD; Sharifah Alsayed[3*], PhD; Petelyne Pangket[4*], PhD; Ebtsam Abou Hashish[3,5,6,7*], PhD; Razan Alsayed[8*], MSN; Romeo Jr Mostoles[9*], PhD

[1]Medical Surgical Nursing Department, College of Nursing, University of Ha'il, 3871, Hail, Ha'il Region, Saudi Arabia

[2]Medical Surgical Nursing Department, University of Ha'il, Hail, Ha'il Region, Saudi Arabia

[3]College of Nursing, King Saud bin Abdulaziz University for Health Sciences, Jeddah, Mecca Region, Saudi Arabia

[4]Medical Surgical Department, College of Nursing, Taif University, Taif, Mecca Region, Saudi Arabia

[5]King Abdullah International Medical Research Center, Jeddah, Mecca Region, Saudi Arabia

[6]Ministry of National Guard Health Affairs, Jeddah, Mecca Region, Saudi Arabia

[7]Faculty of Nursing, Alexandria University, Alexandria, Egypt

[8]Nursing Education Department, King Fahd Armed Forces Hospital, Jeddah, Mecca Region, Saudi Arabia

[9]Department of Psychiatric and Mental Health Nursing, College of Nursing, University of Ha'il, Hail, Ha'il Region, Saudi Arabia

[*]all authors contributed equally

Corresponding Author:
Nader Alnomasy, PhD
Medical Surgical Nursing Department, College of Nursing, University of Ha'il, 3871, Hail, Ha'il Region, Saudi Arabia

Abstract

Background: The integration of digital health technologies into nursing education in Saudi Arabia requires reliable tools to assess nursing informatics competency and digital technology self-efficacy among students.

Objective: This study aimed to evaluate the reliability and validity of the Canadian Nursing Informatics Competency Assessment Scale (C-NICAS) and Digital Technology Self-Efficacy (DT-SE) scale among undergraduate nursing students at a Saudi university.

Methods: A descriptive cross-sectional survey of 243 undergraduate nursing students at the University of Ha'il was conducted using the C-NICAS and DT-SE. Internal consistency was examined using Cronbach α, and construct validity was assessed using exploratory and confirmatory factor analyses.

Results: A total of 243 students participated (mean C-NICAS score 54.0, SD 16.9; mean DT-SE score 2.7, SD 0.56). Both scales showed good internal consistency (C-NICAS total α=0.90; DT-SE α=0.80). C-NICAS demonstrated a multidimensional factor structure with an acceptable model fit (comparative fit index=1.00; root mean square error of approximation=0.081), whereas DT-SE showed a 3-factor structure with a suboptimal confirmatory model fit (comparative fit index=0.76, root mean square error of approximation=0.146).

Conclusions: The C-NICAS and the DT-SE are suitable for assessing informatics competency and digital self-efficacy among undergraduate nursing students at this institution, although further refinement of the DT-SE may improve model fit. These validated tools can inform curriculum reform at this and similar institutions in Saudi Arabia and support the digital health goals of Saudi Vision 2030.

(JMIR Nursing 2026;9:e88075) doi:10.2196/88075

KEYWORDS

nursing informatics competency; digital self-efficacy; nursing students; Saudi Vision 2030; health care digital transformation

Introduction

Health care providers worldwide are urgently challenged to integrate digital health technologies into their practices, requiring them to acquire new competencies, especially nurses, the largest single profession in health care systems [1,2]. Electronic health records (EHRs), decision support systems, and mobile health apps are increasingly used to provide safe, efficient, and patient-centered care [3]. Informatics competencies are essential professional requirements for all practicing nurses

XSL•FO
RenderX

today and have been established through international initiatives, including the Technology Informatics Guiding Education Reform (TIGER) initiative. TIGER seeks to prepare new nursing graduates to be digitally literate to enhance evidence-based practice and collaboration among health care disciplines [4]. Additionally, the International Medical Informatics Association (IMIA) has published global educational recommendations that define core biomedical and health informatics competencies and provide a framework for designing and accrediting health and nursing informatics curricula across different professional roles and levels of specialization [5,6]. In parallel, the World Health Organization's *Global Strategy on Digital Health 2020–2025* [7] emphasizes that strengthening the digital competencies of the health workforce, including nurses, through pre-service and in-service education is a key component of achieving safe, effective, and equitable digital health implementation. The World Health Organization guidance on digital education further highlights that integrating digital health content and e-learning into health professional curricula is essential for building workforce capacity and supporting lifelong learning in rapidly evolving digital health systems [8].

Saudi Vision 2030 represents Saudi Arabia's national vision for economic development and social reform, with a focus on digital transformation and the adoption of eHealth as a mechanism to cultivate a modernized and technologically prepared workforce to address future digital challenges in health care [9]. Investments in national EHR and health information systems demonstrate this commitment to digital transformation [10]. Although significant progress has been made toward achieving the digital transformation of health care services in Saudi Arabia, barriers continue to exist among nursing students; newly graduated nurses from Saudi Arabia continue to indicate that they possess modest levels of informatics competency and low confidence in their ability to apply digital technology in both academic and clinical settings [2,11]. The results of these studies clearly illustrate an important disconnect between the national goal of a fully developed digital health care system and the preparedness of the future nursing workforce to deliver services in this system [12].

Several international studies support the use of the Canadian Nursing Informatics Competency Assessment Scale (C-NICAS) in both high- and low-income countries. Internationally, studies validating the tool have provided evidence for its reliability and validity across multiple nursing populations in Canada [13,14] and have demonstrated high levels of reliability during adaptation efforts in Australia [15]. The Digital Technology Self-Efficacy (DT-SE) scale, which measures self-efficacy in relation to digital technology, has been used in Europe, Asia, and the Arab world. These studies have demonstrated consistent reliability across regions [16,17]. Systematic review results show that when these tools are transferred to other countries, they do not always perform similarly in terms of psychometric properties. Therefore, it is essential to conduct local validation to account for variations in nursing education and practice settings.

However, no prior study has rigorously validated the C-NICAS and DT-SE scales for Saudi undergraduate nursing students, which represents a significant gap in the field. The lack of culturally adapted and psychometrically sound measures for assessing digital competency in undergraduate nursing students in Saudi Arabia does not allow educators to reliably measure baseline digital competency, assess the effectiveness of their curricula, or make evidence-based decisions for reform, as required by the digital health priorities of Saudi Vision 2030.

Therefore, this study aimed to evaluate the reliability and validity of the C-NICAS and DT-SE scales among undergraduate nursing students in Saudi Arabia to provide robust, locally validated measures of informatics competency and digital technology self-efficacy.

This study also enables a reliable measurement of students' digital readiness, which bridges the gap between knowledge and practice, supports the design of educational interventions, and provides benchmarks for educational institutions. Finally, it offers practical insights into developing a digitally competent nursing workforce and advancing cross-cultural scholarship in nursing informatics for nurse educators, curriculum developers, and policymakers. The validation of the C-NICAS and DT-SE tools in Saudi Arabia will further facilitate the ongoing evaluation of curriculum changes in digital health education through repeated assessments and targeted integration of simulation learning and hands-on experience using electronic medical records while supporting investment in digital resources and informing workforce modernization strategies for national health care transformation.

Methods

Study Design

This descriptive cross-sectional study evaluated the psychometric properties of the C-NICAS and DT-SE scale among Saudi undergraduate nursing students.

Setting, Participants, and Sampling

Students attending the College of Nursing at the University of Ha'il were recruited between February and June 2025 for the study. A priori power calculation was conducted to identify how many participants would be needed to support a one-way ANOVA, using G*Power software (version 3.1.9.7) [18] with a small-to-medium effect size (f=0.25), an α level of .05, and a statistical power of 0.80. Two hundred students were required to achieve these parameters.

The most practical and efficient method for recruiting a sample within the confines of the project timeline and budget was convenience sampling. This sampling strategy allowed for the recruitment of students across multiple academic years/levels and facilitated a broader representation of student participation. However, the use of convenience sampling also introduced a risk of selection bias, in that students who were more readily available and/or engaged may have been overrepresented. Additionally, limiting recruitment to a single institution restricted the generalizability of the findings, as they may not have accurately reflected the larger population of Saudi nursing students in the country. Therefore, future studies should use multi-institutional or probability-based sampling designs to increase the representativeness and external validity of their findings.

XSL•FO
RenderX

The inclusion criteria were as follows: active students in the nursing program who were aged 18 years or older, were fluent in English, and provided informed consent. Students in their preparatory year, non-nursing majors, students on leave, those who withdrew, and visiting students not affiliated with the college were excluded. An additional 20% of the originally calculated number of participants was included to allow for attrition and missing data in participant responses.

Rationale for Instrument Versions

Two valid instruments were used in this study. The first was C-NICAS version 2, the most recent peer-reviewed version of an instrument that represents the current informatics standards accepted throughout international competency frameworks. Research has supported the reliability and relevance of the domains of this instrument for measuring nursing informatics competency worldwide; therefore, it was determined to be appropriate for establishing benchmarks in nursing informatics competencies for undergraduate nursing education in Saudi Arabia. The second instrument was the DT-SE scale [19], which has been shown to have high psychometric properties (α>.90) and has been successfully adapted in several cultures. This scale can measure and compare self-efficacy for digital health tasks among nursing students with some degree of reliability and comparability. Both instruments provided evidence-based and internationally comparable assessments while providing locally validated measures for use within the Saudi context.

Survey Instrument

The survey tool was divided into two components: part 1 contained demographic questions regarding age, gender, level of education, prior experience with EHRs, and individual self-perception of digital readiness. Part 2 contained the C-NICAS version 2 (4 subscales with 26 items), a tool designed to measure competency in 4 areas: foundational computer skills, information and knowledge management, professional and regulatory accountability, and use of computers in patient care. The DT-SE (17 items) measures individual self-efficacy for completing digital health tasks (Likert scale: 1="strongly disagree" to 4="strongly agree"; some items were reverse scored); higher scores represent greater digital self-efficacy.

Data Collection

Data were collected from the participants using anonymous online surveys administered through a secure web-based platform. Withdrawal from the study was permissible at any time . All items were completed in a self-report format, and the participants submitted their responses electronically upon completing the survey.

Data Analysis

Following data collection, all responses were exported to SPSS Statistics for Windows (version 29; IBM Corp) for data cleaning and analysis. Descriptive statistics (mean, median, SD, and frequency distribution) were computed to summarize the sociodemographic characteristics and overall scores on the C-NICAS and DT-SE. Reliability analyses were conducted to estimate the internal consistency coefficients for each scale and subscale, and exploratory and confirmatory factor analyses were performed to evaluate the underlying factor structures and model fit of the C-NICAS and DT-SE. Mean imputation was applied to address missing values for continuous variables, whereas mode imputation was used for missing categorical variables. Python was used to generate supplementary data visualizations to support the interpretation and presentation of the findings.

Ethical Considerations

This study was approved by the Institutional Review Board of the University of Ha'il (approval H-2024 - 437). Online informed consent was obtained from each participant before the survey was initiated. The responses were anonymous and stored in a secure location on a password-protected server that only the research team could access. Participation in the study was voluntary and did not result in any penalties. No incentives were provided for participation.

Results

The demographic information of the 243 Saudi nursing undergraduates is presented in Table 1. Of the 243 participants, 110 (45.5%) were aged 20 years or less, 116 (47.7%) were aged between 21 and 25 years, and 17 (7%) were aged 26 years or more. The majority of participants were women (159/243, 65.4%) and a minority were men (84/243, 34.6%). There were 69 (28.4%) first-year students, 72 (29.6%) second-year students, 39 (16%) third-year students, and 63 (25.9%) fourth-year students. Of the 243 participants, 124 (51%) reported that they had completed some type of informatics-related education prior to participating in this study, and 119 (49%) stated that they had not completed any type of education related to informatics.

XSL•FO
RenderX

Table . Sociodemographic characteristics and experience/attitudes regarding technology among respondents (N=243).

		Participants, n (%)
Age		
	≤20 years	110 (45.5)
	21 - 25 years	116 (47.7)
	≥26 years	17 (7.0)
Gender		
	Male	84 (34.6)
	Female	159 (65.4)
Year		
	First year	69 (28.4)
	Second year	72 (29.6)
	Third year	39 (16.0)
	Fourth year	63 (25.9)
Informatics training		
	No	119 (49.0)
	Yes	124 (51.0)
Comfort with digital technology		
	Very uncomfortable	47 (19.3)
	Uncomfortable	53 (21.8)
	Neutral	54 (22.2)
	Comfortable	89 (36.6)
Electronic health record exposure		
	No experience	72 (29.6)
	Occasional	98 (40.3)
	Frequent	73 (30.0)
PC/laptop access		
	No	49 (20.2)
	Yes	194 (79.8)
Technology for education		
	Monthly	11 (4.5)
	Weekly	32 (13.2)
	Daily	200 (82.3)
Home internet		
	Fair	19 (8.7)
	Good	66 (27.2)
	Excellent	158 (65.0)
Career area		
	Clinical care	130 (53.5)
	Academic/teaching	35 (14.4)
	Public health/community	17 (7.0)
	Administration/leadership	14 (5.8)
	Informatics	14 (5.8)
	Undecided	33 (13.6)

XSL•FO
RenderX

Of the 243 participants, 47 (19.3%) stated that they felt very uncomfortable using digital technology, 53 (21.8%) felt uncomfortable, 54 (22.2%) felt neutral about using digital technology, and 89 (36.6%) felt comfortable. Of the participants, 72 (29.6%) stated that they had never seen an EHR before, 98 (40.3%) reported seeing EHRs monthly or less, and 73 (30%) reported using EHRs every week or more.

A large proportion of participants (194/243, 79.8%) either owned their own computers/laptops or had regular access to them. However, 49 (20.2%) participants indicated that they did not have regular access to a personal computer/laptop. Regarding the use of technology for educational purposes, 11 (4.5%) reported that they used this technology monthly, 32 (13.2%) used it weekly, and 200 (82.3%) reported that they used it daily.

In terms of accessing the internet at home, 19 (8.7%) of the participants rated their access as fair, 66 (27.2%) rated their access as good, and 158 (65%) rated their access as excellent. When asked what areas of nursing they planned to pursue after graduation, 130 (53.5%) of the participants planned to pursue roles in direct patient care in hospitals or clinics, 35 (14.4%) wanted to pursue roles in academia or as educators, 17 (7%) wanted to pursue roles in public health/community nursing, 14 (5.8%) wanted to pursue administrative/leadership roles, 14 (5.8%) were interested in pursuing roles in nursing informatics, and 33 (13.6%) were undecided about what area of nursing they would pursue.

Table 2 presents descriptive statistics for the C-NICAS and DT-SE scales. Mean scores across the C-NICAS domains were relatively consistent, including information and communication technology (ICT) skills (mean 2.6, SD 0.93), knowledge (mean 2.5, SD 0.87), accountability (mean 2.6, SD 0.89), and use (mean 2.6, SD 0.85). The overall mean score for the C-NICAS was 54 (SD 16.9), suggesting a moderate level of informatics competency among the students. The mean DT-SE score was 2.7 (SD 0.56), indicating moderate self-efficacy in using digital technology.

Table . Descriptive statistics of informatics competency and digital self-efficacy.

Variable	Scores, mean (SD)
Foundational information and communication technology skills	2.6 (0.93)
Information and knowledge management	2.5 (0.87)
Professional and regulatory accountability	2.6 (0.89)
Use of information and communication technology in patient care	2.6 (0.85)
Overall C-NICAS[a]	54 (16.9)
DT-SE[b]	2.7 (0.56)

[a]Canadian Nursing Informatics Competency Assessment Scale.

[b]Digital Technology Self-Efficacy.

Table 3 summarizes the reliability testing results for the C-NICAS. Cronbach α values indicated good to excellent internal consistency for all domains: foundational ICT skills (α=0.70), information and knowledge management (α=0.80), professional and regulatory accountability (α=0.82), and use of ICT in patient care (α=0.90). The overall the C-NICAS demonstrated excellent reliability (α=0.90). The DT-SE scale showed good reliability (α=0.80).

Table . Internal consistency reliability of the Canadian Nursing Informatics Competency Assessment Scale (C-NICAS) and Digital Technology Self-Efficacy (DT-SE) scale.

Scale	Items, n	Cronbach α	Level of internal consistency
Foundational information and communication technology skills	2	0.70	Good
Information and knowledge management	6	0.80	Good
Professional and regulatory accountability	6	0.82	Good
Use of information and communication technology in patient care	10	0.90	Excellent
C-NICAS	26	0.90	Excellent
DT-SE	17	0.80	Very good

XSL•FO
RenderX

Factor Structure of C-NICAS

Exploratory factor analysis supported a multidimensional structure for the C-NICAS, with item commonalities in an acceptable range and excellent sampling adequacy (Kaiser-Meyer-Olkin=0.976). Confirmatory factor analysis indicated an acceptable model fit for the 4-factor model (comparative fit index=1.000; incremental fit index=1.000; root mean square error of approximation=0.081; χ^2_1=2.02). The full factor loading matrix and related statistics are presented in Table 1 in Multimedia Appendix 1, with the corresponding factor loading plot in Figure 1 in Multimedia Appendix 1 and detailed CFA fit indices in Table 2 in Multimedia Appendix 1.

Factor Structure of DT-SE

For the DT-SE, sampling adequacy was strong (Kaiser-Meyer-Olkin=0.930), and 3 factors emerged in the exploratory factor analysis, explaining 67.7% of the total variance. The rotated factor loadings and variance explained are shown in Tables 3 and 4 in Multimedia Appendix 1, and the factor-loading plot is shown in Figure 2 in Multimedia Appendix 1. Confirmatory factor analysis of the DT-SE indicated a suboptimal model fit (χ^2_{119}=877.10; comparative fit index=0.76; root mean square error of approximation=0.146), suggesting the need for further refinement of the scale in this context. The detailed CFA indices are summarized in Tables 5 and 6 and Figure 3 in Multimedia Appendix 1.

Discussion

In this cross-sectional study, the C-NICAS demonstrated excellent internal consistency and a stable multidimensional factor structure, confirming its suitability for assessing informatics competency among undergraduate nursing students at this Saudi institution. The DT-SE scale showed good internal consistency and an interpretable factor structure, although some model fit indices indicated that further refinement may enhance its performance. Overall, the participants reported moderate levels of informatics competency and digital technology self-efficacy, with stronger foundational ICT skills than advanced digital readiness.

Psychometric Properties of C-NICAS and DT-SE

This reliability/validity assessment found that the C-NICAS [20] and DT-SE scale were both highly reliable and valid and thus appropriate for assessing informatics competency and self-efficacy in an educational setting of this type. The strong psychometric properties of both scales indicate that items within each scale group appropriately represent the theoretically supported domains and thus provide a meaningful assessment of competency and self-efficacy among undergraduate nursing students at the University of Ha'il in Saudi Arabia.

The results support similar validation assessments in Canada [21] and demonstrate that the C-NICAS has excellent internal consistency and a valid factor structure across multiple cohorts of undergraduate nursing students. Another similar validation assessment [22] for the Arabic version of the Self-Assessment Nursing Informatics Competencies Scale, which showed high reliability when used to assess nursing informatics competency among Arab nursing students, further supports the use of these measures to assess informatics competency in the Saudi educational context. Together, these studies demonstrate that valid and reliable scales are important for measuring nursing informatics competency.

Conte et al [23] pointed out that there is a possibility of reduced reliability when adapting self-efficacy scales to different cultures. Furthermore, a systematic review of the literature by Al-Qudah et al [24] concluded that nursing informatics scales exhibit variable reliability and validity, depending on the specific population and setting. Both research groups agreed that these scales must be refined and validated over time if they are to be used effectively on an international scale.

Validated and reliable measures of nursing informatics competencies and self-efficacy will provide Saudi nursing educators with data to assess students' readiness to learn, identify strong and weak areas, and provide an evidence base for changes in nursing education programs. The use of valid tools also supports continuing education and contributes to the transformation of the Saudi Arabian health care system by linking nurses' skills to the national digital health objectives.

Although the scales have strengths, their two main limitations are generalizability and fit for culture (especially for the DT-SE). Studies using larger samples and multiple iterations of the scales would assist in further refining the tools, following the authors' recommendations, to ensure continued utility and reliability as standards for health care and informatics continue to evolve in the Kingdom of Saudi Arabia.

Informatics Competency and Digital Self-Efficacy Levels

This study indicates that the undergraduate nursing students at this institution have an average level of informatics competence and a moderate amount of digital self-confidence; thus, they have partially adopted the digital requirements of today's health care environment but demonstrate a gap between what they are taught about clinical informatics and how they apply it in the workplace. Students have digital skills and confidence in performing these tasks; however, their ability to apply and integrate these skills into more complex applications is less developed than their basic skills.

Compared with their global counterparts, nurses in Jordan demonstrated an almost identical profile of moderate-to-high levels of informatics competency as the participants in this study, including strong performance in general basic computer skills and much weaker performance in clinical informatics subdomain areas [25]. A similar pattern has been observed in Australia and several other Arab countries, as well as in many other parts of the world, where nursing students have shown evidence of being competent in routine digital use; however, they are often found to be lacking in specialized areas such as system management and decision-making [26-29]. The similarities among these findings support the notion that this issue is universal in nature.

Recent studies have also reported similar patterns in nursing informatics education. One recent study found that nursing students typically show higher competence in basic computer

XSL•FO
RenderX

skills than in clinical information management and decision support tasks [6]. A further study proposed an undergraduate informatics blueprint that stressed progressive development from foundational to advanced competencies and explicit alignment with IMIA and TIGER recommendations, supporting the need for structured, longitudinal integration of informatics in nursing programs [29].

The level of self-confidence exhibited by nurses regarding digital abilities was similarly modest and reflects current global trends, where confidence in using technology for everyday tasks is commonplace, yet proficiency is limited in advanced applications of informatics [13,30]. While early exposure to digital tools, along with some education, support, and resources, has contributed to moderate progress in improving digital self-efficacy among Jordanian nursing students, limitations remain owing to the lack of preparation in advanced informatics, varied faculty expertise, and inconsistency in the availability of adequate technology and digital systems to support digital readiness [3]. Generational changes and increased digital exposure are likely to improve the basic competencies of nursing students; however, instructional and systemic barriers persist.

The data indicate that nursing programs in Saudi Arabia, including the program at this institution, have made some advancements in digital health; however, further strategic investments are needed to fully support the curriculum design, faculty development, and experiential learning necessary to meet the needs of clinical practice. The use of validated assessments, simulations using artificial intelligence technology, and increased emphasis on interprofessional informatics education will be the key next steps [4,31]. Ultimately, while moderate competency and self-efficacy levels suggest an advancement toward meeting the goals of Saudi Vision 2030, the overall implication of this study is that there continues to be a need for strategic reforms at all educational levels to ensure that graduates can assume leadership roles in future digitally advanced health care systems.

Implications for Nursing Education and Curriculum Development

This research has provided many critical findings that are relevant to nursing education, nursing program assessment, and curriculum development here and potentially elsewhere in Saudi Arabia, in support of Saudi Vision 2030 and global digital transformation objectives in health care.

First, because the nursing students in this sample possessed average levels of informatics competence and digital self-efficacy, there is an imperative for nursing curricula to enhance students' learning experiences with respect to digital health and informatics. For example, rather than simply exposing students to digital health as part of their introductory courses, nursing programs should systematically integrate advanced informatics education, simulation-based learning, EHR training, and education on artificial intelligence and decision support systems into the 4-year undergraduate nursing education experience. By doing so, students will be able to progress from a beginner level of digital literacy to being capable of effectively using emerging technologies in the delivery of patient care.

Second, by using these assessments (the DT-SE and C-NICAS) systematically, educators will be empowered to use data-driven insights when making decisions about the ongoing evaluation of nursing programs and curriculum enhancements. Ongoing assessment of nursing students' informatics competence will empower educators to assess potential weaknesses in their students' abilities to engage in clinical practice using digital technologies, adapt to new clinical requirements, and modify the content of their curricula to be consistent with national and international priorities related to digital health.

Additionally, these enhancements will directly contribute to the goals of Saudi Vision 2030, specifically by providing a future-ready nursing workforce that can assist in promoting and implementing advancements in health care through engagement in national eHealth initiatives and the provision of technology-enabled patient care. Furthermore, to ensure the sustainability of the positive outcomes of this study and to continue to use best practices in both educational and clinical environments, it is important for faculty to continuously receive professional development and for the institution to invest in its digital infrastructure.

Finally, the results of this study further emphasize the importance of innovation in curriculum design and competency-based education in developing nurses with the ability to meet the challenges associated with the digital transformation of health systems globally. Through the alignment of nursing education in Saudi Arabia (particularly in institutions similar to the one in which this study was conducted) with these principles, the health care system will be able to develop professionals who can deliver high-quality, safe, technology-enhanced care to patients.

Study Limitations

This study had several important limitations. The generalizability of the findings is limited by the use of a convenience sample recruited from only one university, and the results should therefore be interpreted as reflecting the students at this institution rather than all Saudi nursing students. Since participants were asked to report their own behavior, there is a potential for both recall and social desirability biases due to the self-reporting nature of the items. Furthermore, the cross-sectional nature of the research design precludes making causal statements about the relationships observed between variables, and possible scale adaptation issues may limit the accuracy of the scales in representing the context being studied. There is a possibility that non-response (dropout) bias may have affected the findings as well, because less technologically savvy students may not have been represented.

Recommendations

To address the gaps identified in this study, we recommend that curriculum developers continue to develop and implement advanced informatics content, simulation-based learning, and practical EHR training across undergraduate nursing programs at this and other institutions in Saudi Arabia. In addition, opportunities for faculty development and long-term investments in digital infrastructure are required to support and enhance the digital health competencies of nursing faculty members.

XSL•FO
RenderX

Policymakers should also provide incentives for the continued implementation and development of standardized and psychometrically sound assessment tools, such as the C-NICAS and DT-SE, for the ongoing monitoring and evaluation of curricula at both the institutional and national levels. We believe that this will help Saudi nursing graduates in an increasingly digitalized health care environment.

Future studies should use multi-institutional, longitudinal research designs and incorporate culturally appropriate tools to increase the representativeness of findings and enable follow-up assessments of curricular impact over time. Future research should also include refinement of the DT-SE scale to improve model fit and identify targeted interventions (eg, simulation and AI-based informatics training) to address existing gaps in advanced digital preparedness among nursing students. Future research examining the association between nursing students' informatics competency and patient outcomes will help optimize nursing education to meet national and international priorities.

Conclusion

The validation of assessments that measure digital competency and self-efficacy is necessary to enhance nursing education in Saudi Arabia. This descriptive cross-sectional study administered the C-NICAS and DT-SE to undergraduate nursing students at a Saudi university, finding that both scales demonstrated strong psychometric properties. The results indicated that nursing students displayed moderate levels of informatics competency and digital technology self-efficacy; specifically, they demonstrated adequate basic-level skills but lacked advanced digital competence compared to international standards. The results indicate the urgent need for comprehensive curriculum reform, faculty development, and strategic investments in advanced technology and assessment techniques to ensure that future nursing graduates are prepared to meet the demands of digital health care delivery systems. The data support the alignment of education with the provisions of Saudi Vision 2030.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Exploratory and confirmatory factor analysis results for the Canadian Nursing Informatics Competency Assessment Scale and Digital Technology Self-Efficacy scale, including factor loadings, model fit indices, and variance explained (Tables 1-6; Figures 1-3).
[DOCX File, 173 KB - nursing_v9i1e88075_app1.docx]

References

1. Abou Hashish E, Alsayed S, Alqarni BHM, Alammari NM, Alsulami RO. Nurse managers' perspectives on digital transformation and informatics competencies in e-leadership: a qualitative study. J Nurs Manag 2025;2025(1):8178924. [doi: 10.1155/jonm/8178924] [Medline: 40726545]
2. Abou Hashish EA, Alnajjar H. Digital proficiency: assessing knowledge, attitudes, and skills in digital transformation, health literacy, and artificial intelligence among university nursing students. BMC Med Educ 2024 May 7;24(1):508. [doi: 10.1186/s12909-024-05482-3] [Medline: 38715005]
3. Alrasheeday AM, Alshammari B, Alkubati SA, Pasay-An E, Albloushi M, Alshammari AM. Nurses' attitudes and factors affecting use of electronic health record in Saudi Arabia. Healthcare (Basel) 2023 Aug 25;11(17):2393. [doi: 10.3390/healthcare11172393] [Medline: 37685427]
4. Hübner U, Shaw T, Thye J, et al. Technology Informatics Guiding Education Reform - TIGER. Methods Inf Med 2018 Jun;57(S 01):e30-e42. [doi: 10.3414/ME17-01-0155] [Medline: 29956297]
5. Mantas J, Hasman A. IMIA educational recommendations and nursing informatics. Stud Health Technol Inform 2017;232:20-30. [Medline: 28106578]
6. le Roux L, Bimerew M, Chipps J. Preparation of nursing students in nursing informatics competencies for the South African healthcare practice environment. BMC Nurs 2024 Dec 18;23(1):923. [doi: 10.1186/s12912-024-02614-4] [Medline: 39696363]
7. Global strategy on digital health 2020–2025. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2026-04-20]
8. Digital education for building health workforce capacity. World Health Organization. 2020. URL: https://apps.who.int/iris/handle/10665/331524 [accessed 2026-04-20]
9. MOH vision for "e-health". Ministry of Health of Saudi Arabia. 2019. URL: https://www.moh.gov.sa/en/Ministry/nehs/Pages/default.aspx [accessed 2025-10-20]

XSL•FO
RenderX

10. Al-Kahtani N, Alruwaie S, Al-Zahrani BM, et al. Digital health transformation in Saudi Arabia: A cross-sectional analysis using Healthcare Information and Management Systems Society' digital health indicators. Digit HEALTH 2022;8(1):20552076221117742. [doi: 10.1177/20552076221117742] [Medline: 35959196]
11. Al-Olaimat K, Salameh B, Abdulhalim Alqadi R, et al. Readiness and acceptance of nursing students regarding AI-based health care technology on the training of nursing skills in Saudi Arabia: cross-sectional study. JMIR Nurs 2025 Jul 30;8:e71653. [doi: 10.2196/71653] [Medline: 40737490]
12. Alotaibi N, Wilson CB, Traynor M. Enhancing digital readiness and capability in healthcare: a systematic review of interventions, barriers, and facilitators. BMC Health Serv Res 2025 Apr 4;25(1):500. [doi: 10.1186/s12913-025-12663-3] [Medline: 40186200]
13. Kleib M, Nagle L. Development of the Canadian Nurse Informatics Competency Assessment Scale and evaluation of Alberta's registered nurses' self-perceived informatics competencies. Comput Inform Nurs 2018 Jul;36(7):350-358. [doi: 10.1097/CIN.0000000000000435] [Medline: 29668498]
14. Lo B, Nagle LM, White P, Kleib M, Kennedy MA, Strudwick G. Digital and informatics competencies: requirements for nursing leaders in Canada. Healthc Manage Forum 2021 Nov;34(6):320-325. [doi: 10.1177/08404704211015428] [Medline: 34018421]
15. Raghunathan K, McKenna L, Peddle M. Baseline evaluation of nursing students' informatics competency for digital health practice: a descriptive exploratory study. Digit Health 2023;9:20552076231179051. [doi: 10.1177/20552076231179051] [Medline: 37274371]
16. Conte G, Arrigoni C, Magon A, et al. Unlocking the determinants of digital and technological self-efficacy: insights from a cross-sectional study among nurses and nursing students. Healthcare (Basel) 2025 Sep 3;13(17):2208. [doi: 10.3390/healthcare13172208] [Medline: 40941558]
17. Paredes-Aguirre M, Campoverde Aguirre R, Hernandez-Pozas O, Ayala Y, Barriga Medina H. The Digital Self-Efficacy Scale: adaptation and validation of Its Spanish version. Human Behavior and Emerging Technologies 2024 Jan 5;2024:1-11. [doi: 10.1155/2024/3952946]
18. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
19. Hughes JE. Descriptive indicators of future teachers' technology integration in the PK-12 classroom: trends from a laptop-infused teacher education program. J Educ Comput Res 2013 Jun;48(4):491-516. [doi: 10.2190/EC.48.4.e]
20. Kleib M, Nagle L. Factors associated with Canadian nurses' informatics competency. Comput Inform Nurs 2018 Aug;36(8):406-415. [doi: 10.1097/CIN.0000000000000434] [Medline: 29596068]
21. Etoom M, Akhu-Zaheya L, Alnjadat R. Psychometric properties of the Arabic version of the Self-assessment Nursing Informatics Competency Scale. Jordan J Nurs Res 2024;3(4). [doi: 10.14525/JJNR.v3i4.08]
22. Sarac E, Yildiz E. A cultural adaptation study of Nursing Informatics Competency Scale. Nurse Educ Pract 2025 Feb;83:104271. [doi: 10.1016/j.nepr.2025.104271] [Medline: 39848106]
23. Conte G, Arrigoni C, Magon A, Pittella F, Stievano A, Caruso R. A stepwise methodological approach to develop and validate the Digital and Technological Self-Efficacy scale for healthcare workers among nurses and nursing students. Nurse Educ Pract 2023 Oct;72:103778. [doi: 10.1016/j.nepr.2023.103778] [Medline: 37716310]
24. Al-Qudah MM, Oweidat IA, Omar SA, et al. The relationship between nursing informatics competencies and clinical decision-making among registered nurses working at tertiary hospitals. Digit Health 2025;11:20552076251346015. [doi: 10.1177/20552076251346015] [Medline: 40433304]
25. Alnajjar R, Zyoud A, Al-Shamaly H, Tamimi A, AbedRabbu M. Perceptions of nursing informatics on quality of nursing practice and competency among nurses in Jordan. Front Digit Health 2025;7:1585603. [doi: 10.3389/fdgth.2025.1585603] [Medline: 40697444]
26. Salari N, Mohammadi M, Fallahi AR, Seif M, Abdolmaleki A. 24 years of nursing education; challenges and opportunities. A systematic review study. Teach Learn Nurs 2025 Apr;20(2):e383-e388. [doi: 10.1016/j.teln.2024.11.029]
27. Martzoukou K, Luders ES, Mair J, et al. A cross-sectional study of discipline-based self-perceived digital literacy competencies of nursing students. J Adv Nurs 2024 Feb;80(2):656-672. [doi: 10.1111/jan.15801] [Medline: 37489586]
28. Abaddi S. Digital skills and entrepreneurial intentions for final-year undergraduates: entrepreneurship education as a moderator and entrepreneurial alertness as a mediator. Manag Sustain Arab Rev 2024 Aug 13;3(3):298-321. [doi: 10.1108/MSAR-06-2023-0028]
29. Harerimana A, Wicking K, Biedermann N, Yates K. A blueprint for integrating nursing informatics into undergraduate nursing programs: a multiple-case study in Australia and South Africa. Int J Nurs Sci 2026 Jan;13(1):77-87. [doi: 10.1016/j.ijnss.2025.12.004] [Medline: 41684615]
30. Lokmic-Tomkins Z, Raghunathan K, Almond H, et al. Perspectives on the implementation of health informatics curricula frameworks. Contemp Nurse 2024 Apr;60(2):178-191. [doi: 10.1080/10376178.2024.2343010] [Medline: 38662767]
31. Guraya SY. Transforming simulation in healthcare to enhance interprofessional collaboration leveraging big data analytics and artificial intelligence. BMC Med Educ 2024 Aug 28;24(1):941. [doi: 10.1186/s12909-024-05916-y] [Medline: 39198809]

XSL•FO
RenderX

Abbreviations

C-NICAS: Canadian Nursing Informatics Competency Assessment Scale
DT-SE: Digital Technology Self-Efficacy
EHR: electronic health record
ICT: information and communication technology
IMIA: International Medical Informatics Association
TIGER: Technology Informatics Guiding Education Reform

Edited by UM Kinnunen; submitted 19.Nov.2025; peer-reviewed by J Mantas; revised version received 04.Apr.2026; accepted 06.Apr.2026; published 05.May.2026.

Please cite as:
Alnomasy N, Alrashedi H, Alsayed S, Pangket P, Abou Hashish E, Alsayed R, Mostoles RJ
Psychometric Evaluation of the Canadian Nurse Informatics Competency Assessment Scale and the Digital-Technology Self-Efficacy Scale Among Saudi Nursing Students: Cross-Sectional Study
JMIR Nursing 2026;9:e88075
URL: https://nursing.jmir.org/2026/1/e88075
doi:10.2196/88075

© Nader Alnomasy, Habib Alrashedi, Sharifah Alsayed, Petelyne Pangket, Ebtsam Abou Hashish, Razan Alsayed, Romeo Jr Mostoles. Originally published in JMIR Nursing (https://nursing.jmir.org), 5.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Application of the Technology Acceptance Model to Predict Nursing Students' Intention to Use Informatics: Cross-Sectional Study

Habib Alrashedi[1*], PhD; Nader Alnomasy[1*], PhD; Romeo Jr Mostoles[2], PhD

[1]Medical-Surgical Nursing Department, College of Nursing, University of Hail, Baqaa Street, Hail, Saudi Arabia

[2]Mental Health Nursing Department, College of Nursing, University of Hail, Hail, Saudi Arabia

[*]these authors contributed equally

Corresponding Author:
Habib Alrashedi, PhD
Medical-Surgical Nursing Department, College of Nursing, University of Hail, Baqaa Street, Hail, Saudi Arabia

Abstract

Background: Nursing informatics is essential for digital health transformation; however, the technology acceptance of undergraduate nursing students in Saudi Arabia remains underexplored.

Objective: This study examined factors influencing nursing students' intention to use informatics technologies using the technology acceptance model.

Methods: A cross-sectional survey was conducted with 132 undergraduate nursing students. Data were analyzed using descriptive, correlational, and hierarchical regression analyses.

Results: Perceived usefulness (mean 3.68, SD 1.22) and perceived ease of use (mean 3.64, SD 1.32) were the strongest predictors of acceptance, together explaining 87% of the variance (R^2=0.87; β=0.323 for usefulness, P<.001; β=0.195 for ease of use, P=.032). Only 25.8% (n=34) of the students often used electronic health records, while 31.8% (n=42) had no electronic health record experience, indicating a clear gap in practical informatics exposure.

Conclusions: Nursing students' acceptance of informatics is primarily driven by its perceived usefulness and perceived ease of use. These findings highlight the urgent need to integrate practical, user-centered informatics training and clinical simulation into undergraduate nursing curricula to better prepare students for technology-based practice.

(JMIR Nursing 2026;9:e85385) doi:10.2196/85385

KEYWORDS

nursing informatics; technology acceptance model; TAM; digital health; nursing students; informatics

Introduction

The integration of nursing informatics has emerged as a key issue in improving health care delivery, particularly in rapidly changing contexts [1,2], such as Saudi Arabia. Nursing informatics, which serves to bridge information science, computer science, and nursing, can make substantial advances in nursing practice and management through advanced technology usage [3-5]. The effective adoption of systems such as electronic health records (EHRs), clinical decision support tools, and computerized physician order entry is vital for reducing medical errors [6], enhancing patient safety [7], improving care coordination [8], and directly supporting national quality goals. Considering the ambitious objectives of Saudi Arabia's Vision 2030 to digitize and transform its health care industry [9], the practicality of technologies used by the next generation of practitioners is important and perhaps even imperative. Despite the emerging national need, the acceptance and ability of undergraduate nursing students to engage in nursing informatics for implementation are quite concerning and warrant further research [8,10-12].

Recent studies have highlighted issues related to informatics education in Saudi Arabia. In general, nursing students are aware of the importance of nursing informatics; however, their practical acceptance and use are limited in scope for several related reasons, such as a lack of curricular content, limited clinical experience with informatics, and a lack of consistent faculty support [10,13]. In these attempts, there is a definite disconnect between the theory taught and the actual informatics competencies students need in the clinical environment. This highlights the need for the reform of education due to its limited examples of hands-on training and the development of curricular content that aligns more closely with the current informatics standards in health care delivery [14-17]. Thus, the primary research problem addressed by this study was the lack of empirical understanding of the factors that specifically predict

XSL•FO
RenderX

the adoption and use of nursing informatics among undergraduate nursing students in Saudi Arabia, which is a necessary step before effective educational interventions can be designed.

To provide an empirical investigation of this gap, this study uses the established technology acceptance model (TAM) and aims to precisely investigate how undergraduate nursing students' personal perceptions—namely, perceived usefulness (PU) and perceived ease of use (PEOU) of informatics systems—predict their behavioral intention to use those technologies. This study provides a link between identified gaps in the curriculum, experiences, and measurable psychological factors.

A compelling theoretical framework is required to determine the gaps in this study. TAM, developed by Davis [18], is one of the most important and widely used models for studying users' willingness to adopt new technologies. This framework helps us study their main perceptions, PU, and PEOU of nursing informatics systems. The current literature helps to report these general barriers, but a quantitative study using the published and validated TAM framework that focuses on and measures acceptance of these general perceptions of nursing informatics by Saudi Arabian nursing students is nonexistent.

This study extends the core TAM by incorporating contextual factors relevant to nursing informatics education. Social influence was conceptualized in line with TAM2 [19], reflecting perceived expectations from peers, faculty, and clinical environments. Engagement and Sustainability were included as contextual constructs, capturing students' interactions with informatics technologies and their perceptions of their long-term relevance in clinical practice. These variables were examined as supplementary variables rather than primary predictors of acceptance. Although more recent models, such as TAM3 and the unified theory of acceptance and use of technology, incorporate additional contextual and organizational factors, the original TAM was selected for its parsimony and proven applicability in educational and early-stage technology adoption contexts.

As such, this study used the TAM framework in a quantitative study on nursing students' acceptance of nursing informatics in Saudi Arabia. Using the TAM framework, this study is a cross-sectional survey focused on how PU and PEOU affect students' behavioral intention to adopt informatics and provides empirical research and direction for educational policy and curriculum development in alignment with Vision 2030. Despite the recognized importance of nursing informatics in supporting digital health transformation, undergraduate nursing students in Saudi Arabia continue to have limited practical exposure to and variable acceptance of informatics technologies. The factors influencing the intention to adopt and use these technologies remain insufficiently understood, particularly within a theory-driven framework.

Therefore, this study aimed to identify factors that predict undergraduate nursing students' acceptance of nursing informatics technologies using the TAM. The following research questions guided this study: What are undergraduate nursing students' perceptions of nursing informatics in terms of PU, ease of use, and acceptance? What relationships exist between the core TAM constructs and nursing students' acceptance of informatics technologies? Which TAM-related factors significantly predict undergraduate nursing students' acceptance of nursing informatics technologies?

Methods

Research Design

This study used a quantitative, cross-sectional survey to investigate the factors affecting the intention to use informatics technologies, using the TAM as a framework. A cross-sectional survey design was selected for its practicality and feasibility in capturing undergraduate nursing students' perceptions and acceptance of informatics technologies at a single point in time. This design is appropriate for examining the relationships among TAM constructs and identifying key predictors of acceptance without the resource demands associated with longitudinal data collection. This study is reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement guidelines (Checklist 1) [20].

Population and Sampling

This study involved third- and fourth-year undergraduate nursing students from the University of Hail in the Hail Region, Saudi Arabia, from October 2025 to December 2025. Participants were recruited using convenience sampling from students who had successfully completed a nursing informatics course and had relevant clinical experience. All qualified students were contacted via a university communication system. This sampling strategy may limit the generalizability of the results; however, it also facilitates the recruitment of students with proper knowledge and clinical experience. Eligible participants were third- and fourth-year undergraduate nursing students who had successfully completed the nursing informatics and core clinical nursing courses, including supervised clinical training as part of the undergraduate curriculum. Completion of these courses ensured that participants had foundational theoretical knowledge of nursing informatics and prior exposure to clinical practice environments in which health information technologies were used.

A priori power analysis was conducted using G*Power software (Mac and Windows, University of Düsseldorf) for hierarchical multiple regression. The analysis assumed an α level of .05, statistical power of 0.80, and a medium effect size (f^2=0.15). The analysis indicated a minimum sample size of 118 participants. A total of 150 students were invited, of whom 132 completed the survey, yielding a response rate of 88%.

Instrument

This study is comprised of 2 main parts. The first part explored the demographic and technology profiles of undergraduate nursing students participating in the study. The variables of interest included age, sex, year of study, comfort level with digital technology, previous experience using EHRs, computing access, frequency of technology use for educational purposes, Internet access at home, and planned area of nursing work.

XSL•FO
RenderX

The second part used the TAM questionnaire, modified from Davis [18], to measure students' perceptions and behavioral intentions regarding their use of health informatics technologies. The instrument measured 3 core constructs: PU, PEOU, and behavioral intention (BI). Items that further explored issues such as subjective norms (SNs; social influence), sustainability, engagement, and willingness to learn new technologies were also assessed. The TAM items were adapted from a study by Davis [18] and expanded to include domains relevant to contemporary digital health education, namely sustainability, engagement, and willingness to learn new technologies. The final instrument consisted of 31 items distributed across 7 constructs: PU (6 items), PEOU (6 items), SNs (2 items), sustainability (4 items), engagement (7 items), and learning new technologies (6 items). Participants rated their responses on a 5-point Likert scale (from 1="Strongly Disagree" to 5="Strongly Agree"), and the total scores ranged from 12 to 60. Higher ratings indicated that nursing students had more positive opinions about the usability of artificial intelligence in health care. The questionnaire was administered in English, which is the official language of instruction in the nursing program, and all participants were proficient in English. Therefore, no translation or back-translation procedure was required.

Minor wording adjustments were made solely to contextualize the items to nursing informatics and EHR use (eg, replacing generic references to "technology" with "nursing informatics systems"), without changing the meaning or theoretical intent of any item. The underlying structure and theoretical integrity of the original TAM instrument were maintained.

Prior to the main data collection, the instrument was pilot-tested with 20 second-year undergraduate nursing students to assess its clarity, reliability, and internal consistency. These students were excluded from the final sample. Feedback from the pilot test resulted in minor wording refinements to improve item clarity, while the overall structure of the instrument remained unchanged. Internal consistency demonstrated good-to-excellent reliability across constructs in the pilot test, with Cronbach α values of PU=0.953, PE=0.900, SN=0.929, sustainability=0.917, engagement=0.932, and learning new technologies=0.911, indicating satisfactory psychometric properties of the instrument.

Ethical Considerations

After obtaining institutional review board approval from the University of Hail (H-2024 - 437 on September 16, 2024), an ethical recruitment method was established. The purpose of the study was shared with the nursing students in the classroom, and the students were invited to participate. A link to the survey tool was distributed to the students via WhatsApp numbers and university email addresses. Data were collected electronically using Google Forms. Informed consent was obtained electronically through Google Forms. The anonymity and confidentiality of the participants were maintained throughout the study. Moreover, participants were informed that there were no incentives or compensations for participating in this study. Google Forms was chosen for its ease of access and reliable data storage capabilities, providing an efficient process for collecting information on undergraduate nursing students' informatics competencies.

The survey began with an introductory page outlining the purpose of the study, assurances of anonymity, and details of voluntary participation. Participation was voluntary, and the students were informed that they could withdraw at any time without penalty or impact on their academic standing. Data were collected anonymously using Google Forms, and no identifying information was obtained. Access to the data was restricted to authorized members of the research team to ensure confidentiality and data security. The study involved minimal risk and adhered to ethical principles throughout the recruitment and data-collection processes.

Two reminders (1 email and 1 in-class announcement) were used to increase the response rates.

Statistical Analysis

The data were analyzed using IBM SPSS Statistics (version 27). Descriptive statistics were used to summarize participants' demographic characteristics and study variables. Pearson correlation coefficients were calculated to examine the relationships among the TAM constructs. Hierarchical multiple regression analysis was conducted to identify the predictors of nursing students' acceptance of informatics technologies. In model 1, demographic and technology-related control variables (age, gender, year level, prior informatics training, exposure to EHRs, comfort with technology, frequency of technology use, internet accessibility, and intended future nursing career) were entered. In model 2, the core TAM variables—PU and PEOU—were added. Model 3 included extended TAM-related variables (SN, sustainability, engagement, and willingness to learn new technologies).

Before the analysis, the assumptions of multiple regression were examined. The normality of residuals and linearity were assessed through visual inspection of the residual plots. Multicollinearity was evaluated using variance inflation factor values, all of which were within acceptable limits. Statistical significance was set at $P<.05$ for all analyses.

Results

Most participants were female (n=113, 85.6%) and aged 21 to 25 years (n=96, 72.7%). Although access to technology was high, 76.5% (n=101) reported access to a personal computer, and 79.5% (n=105) used technology daily for educational purposes. Comfort with digital technology varied, and clinical informatics exposure remained limited. Only 25.8% (n=34) of the students reported frequent use of EHRs, while 31.8% (n=42) had no prior EHR experience. Clinical care was the most common career pathway (n=80, 60.6%; Table 1).

XSL•FO
RenderX

Table . Demographic and technological profile (N=132).

Category and subcategory	Value, n (%)
Age (y)	
≤20	31 (23.5)
21-25	96 (72.7)
≥26	5 (3.8)
Sex	
Male	19 (14.4)
Female	113 (85.6)
Year level	
Third year	55 (41.7)
Fourth year	77 (58.3)
Rate your comfort level with digital technology (eg, computers, smartphones, and EHRs[a])	
Very uncomfortable	18 (13.6)
Uncomfortable	33 (25)
Neutral	36 (27.3)
Comfortable	45 (34.1)
Previous exposure to EHRs	
No experience	42 (31.8)
Yes, occasionally (monthly or less)	56 (42.4)
Yes, frequently (weekly or daily)	34 (25.8)
Do you own or regularly have access to a personal computer or laptop?	
No	31 (23.5)
Yes	101 (76.5)
How often do you use technology (smartphones, tablets, and computers) for educational purposes?	
Monthly	8 (6.1)
Weekly	19 (14.4)
Daily	105 (79.5)
Internet accessibility at home	
Fair	7 (5.3)
Good	36 (27.3)
Excellent	89 (67.4)
Intended future nursing career area	
Undecided	19 (14.4)
Administration or leadership role	7 (5.3)
Academic or teaching role	11 (8.3)
Nursing informatics	8 (6.1)
Public health or community nursing	7 (5.3)
Clinical care (hospital or clinical setting)	80 (60.6)

[a]EHR: electronic health record.

The descriptive statistics for the study variables are presented in Table 2. Among the TAM constructs, PU (mean 3.68, SD 1.22) and PEOU (mean 3.64, SD 1.32) had the highest mean scores, indicating generally positive perceptions of nursing informatics. Learning new technologies (mean 3.61, SD 1.13) and overall acceptance (mean 3.60, SD 1.24) were also rated positively. In contrast, SN had the lowest mean score (3.21, SD

XSL•FO
RenderX

1.22), suggesting that social influence played a less prominent role in students' perceptions than individual cognitive factors.

Table . Descriptive statistics of study variables.

Variable	Value, mean (SD)
Acceptance	3.60 (1.24)
Perceived usefulness	3.68 (1.22)
Perceived ease of use	3.64 (1.32)
SN[a]	3.21 (1.22)
Sustainability	3.50 (1.20)
Engagement	3.35 (1.15)
Learning new technologies	3.61 (1.13)

[a]SN: subjective norm.

As illustrated in Table 3, the acceptance of nursing informatics demonstrated strong positive correlations with PU (r=0.83; P<.01), PEOU (r=0.81; P<.01), and learning new technologies (r=0.83; P<.01). Additionally, PU and PEOU were highly interrelated (r=0.82; P<.01), supporting the theoretical assumptions of TAM. SNs showed weaker, though still significant, correlations with acceptance (r=0.45; P<.01), reinforcing the comparatively limited role of social influence in this sample. Despite these strong correlations, multicollinearity diagnostics remained within acceptable limits (variance inflation factor<6), suggesting that the predictors could be reliably included in the regression models.

Table . Intercorrelations between study variables.

Variable	Acceptance	PU[a]	PEOU[b]	SN[c]	Sustainability	Engagement	LN[d]
Acceptance	—[e]	.83	.81	.45	.79	.69	.83
PU	.83[f]	—	.82	.43	.75	.60	.78
PEOU	.81[f]	.82[f]	—	.43	.77	.61	.78
SN	.45[f]	.43[f]	.43[f]	—	.50	.56	.47
Sustainability	.79[f]	.75[f]	.77[f]	.50[f]	—	.77	.84
Engagement	.69[f]	.60[f]	.61[f]	.56[f]	.77[f]	—	.77
Learning new technologies	.83[f]	.78[f]	.78[f]	.47[f]	.84[f]	.77[f]	—

[a]PU: perceived usefulness.

[b]PEOU: perceived ease of use.

[c]SN: subjective norm.

[d]LN: learning new technologies.

[e]Not applicable.

[f]P<.01 (2-tailed).

The hierarchical regression for predicting the acceptance of nursing informatics (Table 4) revealed that, in model 1, only comfort with technology was a significant predictor with a negative effect (β=−0.363; P=.05). Model 2 included PU and PEOU, and both variables were significant positive predictors (β=0.379; P<.001). The final model indicated that PU (β=0.323; P<.001), PEOU (β=0.195; P=.03), and learning new technologies (β=0.260; P=.008) were significant positive predictors.

XSL•FO
RenderX

Table . Hierarchical regression predicting acceptance with 95% CIs.

Predictor	B	SE	β	*t* test (*df*)	*P* value	95% CI	VIF[a]
Model 1							
Age (y)	−0.136	0.217	−0.053	−0.626 (122)	.53	−0.566 to 0.294	1.11
Gender	−0.080	0.295	−0.023	−0.270 (122)	.79	−0.663 to 0.504	1.07
Informatics training	−0.019	0.213	−0.008	−0.089 (122)	.93	−0.441 to 0.403	1.13
Exposure to EHRs[b]	0.199	0.135	0.122	1.472 (122)	.14	−0.069 to 0.467	1.05
Year level	0.387	0.219	0.154	1.766 (122)	.08	−0.047 to 0.820	1.16
Comfort with technology	−0.427	0.099	−0.363	−4.324 (122)	<.001[c]	−0.623 to −0.232	1.08
Often using technology	−0.250	0.183	−0.113	−1.366 (122)	.18	−0.611 to 0.110	1.13
Internet accessibility	0.145	0.208	0.065	0.695 (122)	.49	−0.267 to 0.557	1.18
Future nursing career	0.032	0.087	0.026	0.365 (122)	.72	−0.140 to 0.203	1.07
Model 2							
PU[d]	0.357	0.077	0.379	4.632 (120)	<.001[c]	0.205 to 0.510	3.16
PEOU[e]	0.347	0.077	0.379	4.632 (120)	<.001[c]	0.205 to 0.510	3.16
Model 3							
PU	0.330	0.083	0.323	3.980 (116)	<.001[c]	0.166 to 0.494	3.94
PEOU	0.207	0.084	0.195	2.481 (116)	.03[f]	0.018 to 0.395	3.99
SN[g]	−0.006	0.054	−0.007	−0.131 (116)	.89	−0.113 to 0.101	1.61
Sustainability	0.092	0.071	0.113	1.294 (116)	.20	−0.048 to 0.233	4.12
Engagement	0.192	0.085	0.185	1.094 (116)	.28	−0.155 to 0.276	3.62
Learning new technologies	0.286	0.105	0.260	2.721 (116)	.008[c]	0.078 to 0.493	5.46

[a]VIF: variance inflation factor.

[b]EHR: electronic health record.

[c]**P<.01.

[d]PU: perceived usefulness.

[e]PEOU: perceived ease of use.

[f]*P<.05.

[g]SN: subjective norm.

Table 5 summarizes the proportion of variance in the acceptance of nursing informatics explained by predictors across each step of the hierarchical regression model. In model 1, the demographic and technology-related control variables explained 45% of the variance (R^2=0.45), with an adjusted R^2 of 0.20. In model 2, when PU and PEOU were entered, the model explained 87% of the variance (R^2=0.87), with an adjusted R^2 of 0.75. In model 3, when SN, sustainability, engagement, and learning new technologies were added, the final model explained 90% of the variance (R^2=0.90), with an adjusted R^2 of 0.80. The increase in R^2 from model 1 (0.45) to model 2 (0.87) is striking; therefore, PU and PEOU are the primary predictors, and the final model has a very good overall fit.

XSL•FO
RenderX

Table . Hierarchical multiple regression.

Model	R^2	Adjusted R^2
1	0.45	0.20
2	0.87	0.75
3	0.90	0.80

Table 6 displays hierarchical regression, revealing that all models predicting acceptance of nursing informatics were statistically significant. In model 1, which included demographic and technological controls, the regression was significant ($F_{9,122}$=3.503; P<.001), accounting for a small amount of variance in acceptance. Model 2, which included the core TAM constructs (PU and PEOU), showed a substantial increase in the variance explained ($F_{11,120}$=32.442; P<.001), confirming the primary role of these 2 constructs in acceptance behavior. Model 3 also showed a strong statistically significant outcome ($F_{15,116}$=32.024; P<.001), as it included all 7 study variables and accounted for the largest amount of variance explained. The increasing values of the F test and the consistent statistical significance across the models indicate that the contextual variables and TAM constructs work in succession to predict students' acceptance of nursing informatics.

Table . ANOVA summary for hierarchical regression predicting acceptance[a].

Model	Sum of squares (regression)	Sum of squares (residual)	Mean square (regression)	Mean square (residual)	F test (df)	P value
1	41.527	160.702	4.614	1.317	3.503 (9, 122)	<.001[b]
2	151.339	50.890	13.758	0.424	32.442 (11, 120)	<.001[b]
3	162.893	39.336	10.860	0.339	32.024 (15, 116)	<.001[b]

[a]Analyses were performed using IBM SPSS Statistics (version 27; IBM Corp).

[b]P<.01.

Discussion

Perceptions and Acceptance of Nursing Informatics

Nursing students have a strong, positive perception of nursing informatics, recognizing these technological tools as practical resources that contribute to the quality and safety of patient care, which supports national modernization strategies for digital health in the sector. Consistent with the present findings, Al-Olaimat et al [21] reported that perceived usefulness is a strong predictor of nursing students' acceptance of health information technologies, underscoring the importance of clinical relevance. Similarly, Alnajjar et al [22] found that perceived ease of use significantly influenced technology adoption among nursing students, aligning with our results that usability remained central to informatics acceptance.

However, positive perspectives on nursing informatics and the use of technology are not universal and have exposed serious challenges to implementation. Studies have shown negative or mixed perspectives on nursing informatics and technology use, especially among students who have received little or no training. Given this gap between positive and negative experiences, nursing curricula should incorporate core informatics competencies across all training levels.

Correlation Between TAM Domains

The high positive correlations found between factors influencing nursing students' acceptance of informatics technologies and their perceptions of usefulness, ease of use, and learning opportunities indicate that these personal factors are the primary drivers of adoption. This is not surprising, as students are inherently drawn to technologies that optimize their work, are easy to navigate, and directly advance the acquisition of important professional skills; these direct personal advantages ultimately fulfill their pragmatic and professional needs.

Conversely, the study results indicate that external social norms or pressures have significantly less influence. In other words, students rely on their own attitudes and experiences rather than peer or institutional pressure when making adoption decisions.

This finding is consistent with other studies showing that personal direct advantages are significant predictors of technology acceptance by nursing students and clinicians. In a 2025 cross-sectional study by Jallad et al [23], which used TAM as a research framework, students' personal factors (usefulness, ease of use, human, and technology) were highly positively correlated with informatics adoption. Moreover, their findings suggest that personal advantages still matter even in the presence of institutional support. Similarly, Aldosari et al [24] found a highly positive correlation between PU and PEOU, which led to positive acceptance of health information technologies by nurses. Nguyen et al [25] likewise found that, especially for clinicians, PU and PEOU had an exceptionally influential effect on information and communication technologies acceptance, typically outweighing social influence.

Although these data likely emphasize personal factors, other studies indicate the opposite premise, where external social influence is stronger, at least comparable to, or even better than personal benefits, indicating that context affects the degree of influence. For example, Warshawski [26] found that social influence, performance expectancy, and facilitating conditions

XSL•FO
RenderX

were important variables that positively contributed to nursing students' use of technology in Israel. Furthermore, Garavand et al's [27] systematic review made similar assertions regarding the importance of social impacts (eg, peer and institutional support), noting that they were not always weighted as heavily as personal factors.

Factors Influencing the Acceptance of Nursing Informatics

This study explored the factors influencing the prospective acceptance of nursing informatics by Saudi undergraduate nursing students. Overall, the findings highlight individuals' central and predominant personal perceptions as predictors of technology acceptance in this population, which is congruent with the underlying theoretical principles of the TAM. The input of cognitive and social variables provides institutions with practical recommendations to support the integration of informatics into the nursing education curriculum, which is pertinent, as the Saudi health care system continues to progress through digital changes.

The final structural model provided solid empirical support for the fundamental TAM constructs, highlighting that PU, PEOU, and students' ability and willingness to learn new skills are statistically significant positive predictors of perceived technology acceptance. The findings demonstrated that PU and PEOU were related to a positive and significant degree of perceived acceptance and were consistent with the fundamental TAM assumptions and an extensive array of empirical literature in health care. PU and PEOU were strongly and positively correlated with perceived acceptance. This is consistent with the fundamental premise of TAM, which posits that students who perceive informatics systems to be useful in their education and future desired clinical experience (PU) and easy to understand (PEOU) will act to adopt the technology. The existing literature, including studies on Saudi nursing students and practicing nurses [24,28-30], suggests that of all the elements explored, only 2 personal cognitive constructs, PEOU and PU, are consistently shown to have the greatest significance. The literature also describes how PEOU and PU are often interrelated, thereby bolstering their interconnected predictive ability [31], which may have contributed to the predictive ability of the present model.

In contrast, the endogenous variables of SN, sustainability, and engagement were not statistically significant, suggesting that these were even less important than personal experience with PEOU and PU. This suggests that social influence (SN), long-term environmental sustainability, and active engagement with technology (sustainability and engagement) did not influence the initiation of acceptance in this student population. While SN being nonsignificant is consistent with some uses of TAM, where personal interest overcomes the social element, sustainability and engagement are not significant and inconsistent with an emerging corpus of published research in the e-learning and health informatics space that has identified that user engagement, perceived enjoyment, and organizational support (variables conceptually associated with sustainability and engagement) could become predictive, particularly if the use was measured in a temporally posterior fashion [32,33].

The initial observation that PU and PEOU are key drivers of technology acceptance, whereas SN, sustainability, and engagement are not statistically significant, has important context. This finding indicates that, for initial adoption, personal internal motivation (use and ease) is more important than social pressure. Similarly, SN, sustainability, and engagement not being significant were probably affected by students' exposure to informatics systems (eg, EHRs) not being large-scale, and dimensions regarding long-term and organizational support do not have enough influence on study acceptance at the beginning of use.

This is consistent with TAM in the health care sector. Reviews have always supported PU and PEOU as dominant predictors and found that SN and its related constructs were nonsignificant [31]. Health care studies have shown that internal cognitive processes outweigh social influences on clinicians and students in novice settings [24,25,29]. The lack of predictive ability for sustainability and engagement constructs also aligns with the results in the context of a lack of practical experience [34,35].

However, this inverse result was also demonstrated. Research in learning technology or mobile health usually reports engagement or sustainability properties (such as enjoyment or organizational support) as significant [32,33] when examined in a technology-rich or long-term context.

Thus, the differential results highlight an important caveat. Although PU and PEOU drive acceptance for this undergraduate sample, the lack of importance of the other constructs may be a contextual (temporal and situational) feature of the participants' inexperience when using the systems in their educational (and possibly first) experiences with such systems. Consequently, participants may regard SN, engagement, and sustainability as collectively gaining importance as predictors later (ie, when adults have graduated and worked for a while in actual practice with longitudinal interactions with complex systems). Organizational support and peer influence have a greater impact in an institutional environment and lay the groundwork for future longitudinal studies.

The nonsignificant effects of social influence, engagement, and sustainability may reflect students' limited and largely theoretical exposure to informatics technologies. At this early stage, acceptance appears to be driven primarily by PU and PEOU, whereas contextual and organizational factors may become more influential during prolonged or mandatory system use in clinical practice. Overall, our findings support the classic TAM perspective that usability and use dominate early acceptance decisions, whereas social factors may play a greater role in long-term use or in highly structured environments.

Incremental Value of TAM Predictors

The hierarchical regression analysis demonstrated the incremental value of TAM predictors in explaining nursing students' acceptance of informatics technologies. While demographic and background variables accounted for a moderate proportion of the variance, their explanatory power was substantially enhanced by the inclusion of PU and PEOU, which emerged as the dominant predictors. This finding underscores that acceptance among undergraduate nursing students is driven

XSL•FO
RenderX

primarily by individual cognitive evaluations rather than contextual or social factors.

The role of demographic characteristics, ecosystem characteristics, and other common predictors of adoption, such as social influence, is negligible in comparison, further lending credence to the notion that a strategy focused on usability and functionality will best advance informatics acceptance. This pattern is expected, as a substantial amount of evidence has emerged that supports the TAM as the reigning model in health care, consistently identifying usability and PEOU as the primary predictors for nurses' and other health professionals' adoption of health records and myriad other systems [24,36,37].

While the first 2 key TAM predictors are statistically reliable, the literature is not homogeneous. However, caution should be exercised when interpreting the extent to which the other variables of technology acceptance are equally or more robust. Some studies, especially those that have combined the unified theory of acceptance and use of technology model or the information system success model as a foundation, offer examples where social influence, engagement, organizational factors, and education are the strongest predictors of technology use, such as in mobile learning [38]. Moreover, reviews have reported that facilitating conditions, technostress, system quality, and satisfaction could be as impactful as, or more impactful than the primary TAM variables, especially in social, mandated, or organizational pressure contexts [35,38,39]. This suggests that PEOU may not always be a direct predictor, as its impact may be mediated by organizational culture, trust, or policy. Additionally, the direct impacts of TAM variables can be diminished or eliminated during periods of strong compliance or external pressure [35,39]. Future research should explore additional factors that may shape nursing informatics acceptance, including technostress, system quality, and organizational support, particularly in clinical settings where informatics systems are mandatory. Longitudinal and multi-institutional studies may help clarify how these variables interact with core TAM constructs over time and across different stages of professional development.

Limitations of the Study

This study had several limitations. First, the cross-sectional design limits causal inference; longitudinal studies are needed to examine changes in informatics acceptance over time. Second, the data were collected from a single nursing program within one university in Saudi Arabia, which may limit the generalizability of the findings to other institutions or countries. Third, the use of self-reported data may introduce response and social desirability biases, and potential nonresponse bias cannot be ruled out. Finally, participants' limited prior exposure to EHRs may have influenced their perceptions of informatics acceptance.

Implication

The results indicate that PU and PEOU are key predictors of the initial acceptance of nursing informatics, reflecting students' early cognitive evaluations rather than experiential judgment. Given the limited practical exposure to informatics systems among the participants, factors related to sustained use, such as social influence, engagement, and sustainability, were not significant. These findings suggest that while usability and perceived value drive early acceptance, continued use is more likely to depend on prolonged clinical exposure, institutional support, and repeated system interaction, underscoring the need for longitudinal practice-based informatics training in nursing education. Informatics competencies should be explicitly embedded into course learning outcomes to align education with digital health practice requirements. In addition, faculty development programs are needed to prepare educators as informatics "super users" who can effectively support students in technology-rich learning environments. Beyond education, the results are relevant to policymakers and hospital administrators who play a critical role in supporting digital transformation. Investment in standardized training platforms, collaboration between academic institutions and health care organizations, and the provision of supportive clinical informatics infrastructure can facilitate smoother transitions of nursing graduates into digitally enabled health care systems.

Conclusion

This study demonstrates that undergraduate nursing students' acceptance of nursing informatics is primarily driven by perceptions of its usefulness and ease of use, despite generally positive attitudes and high access to technology. These findings highlight a persistent gap between theoretical exposure and practical informatics experience, particularly in relation to EHR use. By prioritizing user-friendly, clinically relevant informatics education and simulation, nursing programs can prepare graduates to participate effectively in Saudi Arabia's ongoing digital health transformation.

Acknowledgments

The authors express their gratitude to all the nursing students who participated in the research. No artificial intelligence tools were used.

Funding

The authors declared no financial support was received for this work.

Data Availability

Upon request, the corresponding author will provide the data.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: HA
Data curation: NA, RJM
Formal analysis: RJM
Investigation: HA, NA
Methodology: HA, RJM
Validation: HA
Visualization: HA
Writing – original draft: HA
Writing – review and editing: HA, NA, RJM
After reading and revising, each author approved the final draft of the manuscript for publication.

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[DOCX File, 33 KB - nursing_v9i1e85385_app1.docx]

References

1. Aldogiher A, Halim YT. Customer engagement in digital health transformation as strategic change: evidence from Saudi Arabia's Vision 2030. Sustainability 2025;17(18):8468. [doi: 10.3390/su17188468]
2. Hess S, Alper C. Nursing informatics: the vital nursing link between technology and patient care. HCA Healthc J Med 2024;5(5):513-516. [doi: 10.36518/2689-0216.2003] [Medline: 39524944]
3. What is nursing informatics. American Nurses Association. 2023 Jul 5. URL: https://www.nursingworld.org/content-hub/resources/nursing-resources/nursing-informatics [accessed 2026-03-25]
4. Hamad NI, El-Ashry AM, Ibrahim IM, Hassan EA. Embracing the future: an insight into nursing students' attitude and perception towards the usability of artificial intelligence in healthcare. Teaching and Learning in Nursing 2025 Apr;20(2):e500-e508. [doi: 10.1016/j.teln.2024.12.017]
5. Hassan EA, El-Ashry AM. Leading with AI in critical care nursing: challenges, opportunities, and the human factor. BMC Nurs 2024 Oct 14;23(1):752. [doi: 10.1186/s12912-024-02363-4] [Medline: 39402609]
6. Chimbo B, Motsi L. The effects of electronic health records on medical error reduction: extension of the DeLone and McLean information system success model. JMIR Med Inform 2024 Oct 16;12:e54572. [doi: 10.2196/54572] [Medline: 39412857]
7. Gatiti P, Ndirangu E, Mwangi J, Mwanzu A, Ramadhani T. Enhancing healthcare quality in hospitals through electronic health records: a systematic review. J Health Inform Dev Ctries 2021;15(2) [FREE Full text]
8. Adeniyi AO, Arowoogun JO, Chidi R, Okolo CA, Babawarun O. The impact of electronic health records on patient care and outcomes: a comprehensive review. World J Adv Res Rev 2024;21(2):1446-1455. [doi: 10.30574/wjarr.2024.21.2.0592]
9. Suleiman AK, Ming LC. Transforming healthcare: Saudi Arabia's Vision 2030 healthcare model. J Pharm Policy Pract 2025;18(1):2449051. [doi: 10.1080/20523211.2024.2449051] [Medline: 39845746]
10. Almarwani AM, Yacoub MI. Examining the effect of an educational program on nursing students' informatics competencies. Nurs Educ Perspect 2023;44(6):E59-E61. [doi: 10.1097/01.NEP.0000000000001106] [Medline: 36792354]
11. Ali AS, Alhirsan SM, Elshazly M, et al. Artificial intelligence in physiotherapy education: a multi‐country Middle East survey of AI acceptance and barriers among university students. Euro J of Educ 2026 Mar;61(1):e70476. [doi: 10.1111/ejed.70476]
12. El-Ashry AM, Hermis AH, Al-Salih SSH, et al. Attitudes toward artificial intelligence and impostor phenomenon among nursing students: a five-country cross-sectional study. Nurse Educ Today 2026 Jun;161:107011. [doi: 10.1016/j.nedt.2026.107011] [Medline: 41653898]
13. Alrashedi H, Alderaan SM, Alnomasy N, Lamine H, Saleh KA, Alkubati SA. Insights into factors affecting nurses' knowledge of and attitudes toward AI and implications for successful AI integration in critical care: cross-sectional study. JMIR Nurs 2026 Jan 16;9:e85649. [doi: 10.2196/85649] [Medline: 41544103]
14. Bergren MD, Maughan ED, VanDenBrink R, Foster BE, Carveth L. Nursing informatics and school nursing: specialists wanted. NASN Sch Nurse 2020 Jul;35(4):208-210. [doi: 10.1177/1942602X20928347] [Medline: 32468905]
15. Mostafa Younes B, Mohamed El-Ashry A, Abdelwahab Ibrahim A, Mohamed Farghaly S, A Alkubati S, G Abou MA. Investigating how digital competency mediates the relationship between administrative competencies and professionalism among nurse managers. Egypt J Health Care 2025 Sep 1;16(3):693-706. [doi: 10.21608/ejhc.2025.452083]
16. Alnawafleh KA, Almagharbeh WT, Alfanash HA, et al. Exploring the ethical dimensions of AI integration in nursing practice: a systematic review. J Nurs Regul 2025 Oct;16(3):228-237. [doi: 10.1016/j.jnr.2025.08.001]

XSL•FO
RenderX

17. Nashwan AJ, Cabrega JA, Othman MI, et al. The evolving role of nursing informatics in the era of artificial intelligence. Int Nurs Rev 2025 Mar;72(1):e13084. [doi: 10.1111/inr.13084] [Medline: 39794874]
18. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
19. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
20. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 2007 Oct 16;4(10):e296. [doi: 10.1371/journal.pmed.0040296] [Medline: 17941714]
21. Al-Olaimat K, Salameh B, Alqadi RA, et al. Readiness and acceptance of nursing students regarding AI-based health care technology on the training of nursing skills in Saudi Arabia: cross-sectional study. JMIR Nurs 2025 Jul 30;8:e71653. [doi: 10.2196/71653] [Medline: 40737490]
22. Alnajjar R, Zyoud A, Al-Shamaly H, Tamimi A, AbedRabbu M. Perceptions of nursing informatics on quality of nursing practice and competency among nurses in Jordan. Front Digit Health 2025;7:1585603. [doi: 10.3389/fdgth.2025.1585603] [Medline: 40697444]
23. Jallad ST, Sawalha R, Awamleh FT. Factors influencing acceptance of nursing informatics system among nursing students in nursing education: a cross-sectional study. Comput Inform Nurs 2025 Sep 1;43(9):e01330. [doi: 10.1097/CIN.0000000000001330] [Medline: 40440716]
24. Aldosari B, Al-Mansour S, Aldosari H, Alanazi A. Assessment of factors influencing nurses acceptance of electronic medical record in a Saudi Arabia hospital. Inform Med Unlocked 2018;10:82-88. [doi: 10.1016/j.imu.2017.12.007]
25. Nguyen MH, Sari NPWP, Li D, Vuong QH. Impacts of social influence, social media usage, and classmate connections on Moroccan nursing students' ICT using intention. Teach Learn Nurs 2025 Jan;20(1):e106-e117. [doi: 10.1016/j.teln.2024.08.014]
26. Warshawski S. Israeli nursing students' acceptance of information and communication technologies in clinical placements. J Prof Nurs 2020;36(6):543-550. [doi: 10.1016/j.profnurs.2020.08.005] [Medline: 33308554]
27. Garavand A, Mohseni M, Asadi H, Etemadi M, Moradi-Joo M, Moosavi A. Factors influencing the adoption of health information technologies: a systematic review. Electron Physician 2016 Aug;8(8):2713-2718. [doi: 10.19082/2713] [Medline: 27757179]
28. Ng L, Osborne S, Eley R, Tuckett A, Walker J. Exploring nursing students' perceptions on usefulness, ease of use, and acceptability of using a simulated electronic medical record: a descriptive study. Collegian 2024 Apr;31(2):120-127. [doi: 10.1016/j.colegn.2023.12.006]
29. Barzekar H, Ebrahimzadeh F, Luo J, Karami M, Robati Z, Goodarzi P. Adoption of hospital information system among nurses: a technology acceptance model approach. Acta Inform Med 2019 Dec;27(5):305-310. [doi: 10.5455/aim.2019.27.305-310] [Medline: 32210497]
30. El-Ashry AM, Al Saleh NS, AlOtaibi NG, et al. The impact of artificial intelligence attitudes and acceptance on critical thinking motivation among nursing students in Saudi Arabia. SAGE Open Nurs 2025;11:23779608251369564. [doi: 10.1177/23779608251369564] [Medline: 40809301]
31. Rahimi B, Nadri H, Afshar HL, Timpka T. A systematic review of the technology acceptance model in health informatics. Appl Clin Inform 2018 Jul;9(3):604-634. [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
32. Ho KF, Chang PC, Kurniasari MD, Susanty S, Chung MH. Determining factors affecting nurses' acceptance of a care plan system using a modified Technology Acceptance Model 3: structural equation model with cross-sectional data. JMIR Med Inform 2020 May 5;8(5):e15686. [doi: 10.2196/15686] [Medline: 32369033]
33. Neves C, Oliveira T, Cruz-Jesus F, Venkatesh V. Extending the unified theory of acceptance and use of technology for sustainable technologies context. Int J Inf Manage 2025 Feb;80:102838. [doi: 10.1016/j.ijinfomgt.2024.102838]
34. Papagiannidis S, Marikyan D. Environmental sustainability: a technology acceptance perspective. Int J Inf Manage 2022 Apr;63:102445. [doi: 10.1016/j.ijinfomgt.2021.102445]
35. Lee AT, Ramasamy RK, Subbarao A. Understanding psychosocial barriers to healthcare technology adoption: a review of TAM Technology Acceptance Model and Unified Theory of Acceptance and Use of Technology and UTAUT frameworks. Healthcare (Basel) 2025 Jan 27;13(3):250. [doi: 10.3390/healthcare13030250] [Medline: 39942440]
36. Hussain A, Zhiqiang M, Li M, et al. The mediating effects of perceived usefulness and perceived ease of use on nurses' intentions to adopt advanced technology. BMC Nurs 2025 Jan 9;24(1):33. [doi: 10.1186/s12912-024-02648-8] [Medline: 39789568]
37. Alsyouf A, Lutfi A, Alsubahi N, et al. The use of a technology acceptance model (TAM) to predict patients' usage of a personal health record system: the role of security, privacy, and usability. Int J Environ Res Public Health 2023 Jan 11;20(2):1347. [doi: 10.3390/ijerph20021347] [Medline: 36674105]
38. Su CY, Chao CM. Investigating factors influencing nurses' behavioral intention to use mobile learning: using a modified unified theory of acceptance and use of technology model. Front Psychol 2022;13:673350. [doi: 10.3389/fpsyg.2022.673350] [Medline: 35651564]

XSL•FO
RenderX

39. Vidal-Silva C, Sánchez-Ortiz A, Serrano-Malebrán J, et al. Social influence, performance expectancy, and price value as determinants of telemedicine services acceptance in Chile. Heliyon 2024 Mar 15;10(5):e27067. [doi: 10.1016/j.heliyon.2024.e27067] [Medline: 38562504]

Abbreviations

EHR: electronic health record
PEOU: perceived ease of use
PU: perceived usefulness
SN: subjective norm
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TAM: technology acceptance model

Edited by A Virani; submitted 06.Oct.2025; peer-reviewed by AM El-Ashry, M Galani; revised version received 10.Feb.2026; accepted 04.Mar.2026; published 13.Apr.2026.

Please cite as:
Alrashedi H, Alnomasy N, Mostoles RJ
Application of the Technology Acceptance Model to Predict Nursing Students' Intention to Use Informatics: Cross-Sectional Study
JMIR Nursing 2026;9:e85385
URL: https://nursing.jmir.org/2026/1/e85385
doi:10.2196/85385

© Habib Alrashedi, Nader Alnomasy, Romeo Jr Mostoles. Originally published in JMIR Nursing (https://nursing.jmir.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Nursing Students' Experiences in a Brief Virtual Reality–Enhanced Workshop: Cross-Sectional Exploratory Study

Niklas Ferdinand Carlsson[1], MLiS; Leonie Klompstra[2], PhD; Mats Westas[2], PhD

[1]University Library, Linköping University, Bredgatan 33, Norrköping, Sweden

[2]Department of Health, Medicine and Caring Sciences (HMV), Linköping University, Norrköping, Sweden

Corresponding Author:
Niklas Ferdinand Carlsson, MLiS
University Library, Linköping University, Bredgatan 33, Norrköping, Sweden

Abstract

Background: There is limited evidence on how brief, optional virtual reality (VR) experiences can be used with first-semester nursing students as experiential learning strategies to support understanding of foundational nursing concepts, outside of mandatory coursework or full-scale simulations. Additionally, little is known about students' and teachers' perceptions of VR as a low-stakes, supplemental learning strategy introduced early in nursing education. Examining these experiences can provide insight into the pedagogical value and scalability of VR-enhanced learning within the formal nursing curriculum.

Objective: This study explored students' and teachers' experiences of a brief, optional, VR-enhanced workshop offered outside mandatory coursework in first-semester nursing education and described students' perceptions of cognitive, social, and teaching presence.

Methods: This was a cross-sectional evaluation at a Swedish public university. A single-session workshop, co-designed by nursing teachers and the university library makerspace (implementation context), combined brief headset exposures (sympathetic arousal via a short roller coaster experience and parasympathetic engagement via a short guided meditation), peer vital-sign practice (instructional aid), small-group synthesis, and a guided debrief aligned with the community of inquiry (CoI) framework. Immediately after the session, students completed a demographics questionnaire, a 7-item workshop-specific VR-perception set, and the 34-item CoI instrument, plus 2 open-ended items; teachers provided short reflections. Analyses were descriptive for quantitative data and summative content analysis of open-ended responses. Participants included 11.9% (16/134) of the invited first-semester students (mean age 25 years, SD 5.1; 15/16, 93.8% women; 6/16, 37.5% with prior VR exposure) and 3 teachers.

Results: Most students agreed or strongly agreed that VR enhanced analysis and observation (12/16, 75%), exploration of phenomena (14/16, 87.5%), conceptual understanding and engagement (13/16, 81.3%), teacher support (13/16, 81.3%), and relevance to the session (14/16, 87.5%). CoI ratings indicated moderately positive perceptions (total mean 3.36, SD 0.44 on a 5-point scale), with cognitive presence rated the highest (mean 3.48, SD 0.41) and exploration being the top subdomain (mean 4.48, SD 0.49); design and organization and facilitation were similar (mean 3.42, SD 0.55 each), whereas direct instruction was rated lower (mean 2.88, SD 0.92). Open-ended remarks described links between theory and embodied experience and noted practical challenges.

Conclusions: This study used an early, optional format; the results showed that brief, contrastive VR exposures paired with scaffolded inquiry and a guided debrief were perceived as pedagogically valuable for exploring foundational physiological concepts, while also highlighting feasibility and logistical considerations for routine teaching. Findings are preliminary and reflect session-level perceptions from a small, self-selected sample; nevertheless, they suggest that structured, low-stakes VR may serve as a feasible supplemental strategy in first-semester nursing education, with implications for potential scalability.

(JMIR Nursing 2026;9:e85780) doi:10.2196/85780

KEYWORDS

virtual reality; education; community of inquiry; cross-sectional studies; experiential learning; education 4.0; academic library makerspace; nursing

Introduction

Background

Contemporary educational paradigms increasingly emphasize the need for immersive, interactive learning environments that integrate theoretical instruction with practical, hands-on experiences, grounded in educational theory and supported by institutional pedagogical frameworks [1-4]. In undergraduate nursing, where foundational concepts about physiology and patient care must be internalized early, there is a growing interest in introducing virtual reality (VR) early in the curriculum as an

XSL•FO
RenderX

optional, low-stakes complement to core coursework to spark engagement and support conceptual understanding in the first semester [5].

A growing evidence base suggests that VR can enhance learning by providing immersive, repeatable, and safe experiences that bridge abstract concepts and embodied understanding. Meta-analyses in nursing education have reported positive effects on theoretical knowledge, technical skills, self-confidence, and learner satisfaction when VR is integrated with attention to design features and instructional support [6-8]. At the same time, findings are not uniform: outcomes vary by intervention design and outcome type, and some studies note that while presence increases, learning can suffer when cognitive load is high or scaffolding is insufficient [5,9,10]. Practical considerations, such as cybersickness risks influenced by exposure duration, content, and locomotion, also warrant attention when working with first-time users [11].

There is limited evidence on how brief, optional VR experiences can be used with first-semester nursing students as experiential learning strategies to support understanding of foundational nursing concepts, outside of mandatory coursework or full-scale simulations. Additionally, little is known about students' and teachers' perceptions of VR as a low-stakes, supplemental learning strategy introduced early in nursing education. Examining these experiences can provide important insight into the pedagogical value and scalability of VR-enhanced learning within the formal nursing curriculum. Reports frequently emphasize usability or novelty yet underdescribe how short VR exposures are pedagogically integrated to connect conceptual content (eg, autonomic physiology) with immediate experiences and collaborative sensemaking [5,10].

To guide a purposefully designed, concise activity, we drew on education 4.0 (learner-centered, experience-driven, and digitally supported learning); constructive alignment (coherence among intended learning outcomes, teaching and learning activities, and assessment); and the community of inquiry (CoI) framework (cognitive, social, and teaching presence) [12-18]. Framing VR as one tool within an aligned pedagogical design helps ensure that brief exposures are structured to foster inquiry, link embodied responses to concepts, and support facilitated reflection.

To address these gaps, we implemented and evaluated a brief, optional, VR-enhanced workshop offered outside mandatory coursework in first-semester nursing education. The model combined short headset exposures designed to elicit contrasting autonomic responses, peer practice of vital-sign measurement to support observation and discussion, small-group synthesis, and a guided debrief aligned to the CoI practical inquiry cycle. We describe the implementation and examine students' and teachers' experiences, focusing on students' perceptions of cognitive, social, and teaching presence as indicators of the perceived learning environment [15-17].

Objective

This study explores students' and teachers' experiences of a brief, optional VR-enhanced workshop offered outside mandatory coursework in first-semester nursing education and described students' perceptions of cognitive, social, and teaching presence.

Methods

Design and Reporting

This cross-sectional study evaluated an optional, VR-enhanced workshop using postworkshop surveys (Likert-type items and open-ended responses). No behavioral observation checklists were used. Reporting follows STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidance for cross-sectional studies [19].

Setting, Participants, and Recruitment

The activity took place at a Swedish public university (south central region). All first-semester nursing students (N=134) were invited to participate via announcements on the online classroom platform and brief in-class invitations. Registration remained open for one week, and 16 (11.9%) students who expressed interest received a confirmation message securing their place in the workshop. Participation was voluntary and outside mandatory coursework. In addition, 3 nursing teachers participated across all workshop phases and provided brief reflections. Library makerspace staff co-facilitated implementation (roles are detailed in subsequent sections) but were not included in the teacher feedback items.

Workshop

This optional, supplemental, VR-enhanced experiential learning activity (distinct from a clinical simulation scenario) was offered alongside the first-semester curriculum and was not included in required assessments or mandatory components. Students were divided into 2 groups of 8 for workflow and logistical reasons (nonrandom allocation). Each group completed one VR module, either a short roller-coaster exposure (sympathetic arousal) or a short guided meditation (parasympathetic engagement), followed by inquiry tasks and a guided debrief (Table 1). The workshop was implemented as an optional addition to existing teaching (ie, not part of the school's required simulation infrastructure) to explore perceived experiences with a concise, aligned model.

Table . Time structure of the workshop.

Component of the workshop	Time (min)
Virtual reality exposure and peer vital-sign practice	60
Working on answering the given questions and preparing a PowerPoint presentation	75
Group presentations and reflection	45

XSL•FO
RenderX

Implementation Roles (Contextual)

The workshop was co-designed and co-facilitated by 2 units: nursing teachers (pedagogical design, alignment with first-semester concepts, and facilitation of the debrief and links to physiology) and university library makerspace staff (readiness and comfort protocols for first-time VR use; headset onboarding; hygiene, sanitation, and logistics; and facilitation of information-seeking and source evaluation during the inquiry task). This cross-unit collaboration supported delivery; no evaluation of collaboration effects (eg, feasibility, process outcomes, or comparative effectiveness) was undertaken, and collaboration was not a study outcome.

VR Exposures and Rationale

We selected brief headset exposures to contrast autonomic activation states: (1) sympathetic activation via an approximately 4-minute Epic Roller Coasters track and (2) parasympathetic engagement via an approximately 5- to 6-minute TRIPP guided meditation, primarily to support conceptual understanding of autonomic physiology and secondarily to maintain comfort and accessibility for first-time users [20,21]. Evidence summarized in prior reviews indicates that cybersickness risk is associated with exposure duration, visual stimulation, and locomotion and content; comfort-oriented, shorter sessions are commonly recommended for first-time users [11,22].

Vital-Sign Practice Procedure

Before headset use, each student had resting baseline values recorded (heart rate, blood pressure, respiratory rate, oxygen saturation, and tympanic temperature) using standard training instruments in the student practice ward. To minimize movement artifact and measurement reactivity, spot-checks were obtained immediately after the VR exposure (not during). These measurements were intended to support observations, understanding of sympathetic and parasympathetic responses, and postexperience discussion, rather than research-grade physiological inference; no continuous monitoring was performed, and readings were not designed to isolate specific organ-system mechanisms.

Grouping and Sequence

Two groups of 8 proceeded in parallel (roller coaster *or* meditation). Assigning 1 module per group supported the single-session timebox, comfort for first-time users, and a contrastive learning design that allowed cross-group comparison during synthesis and debrief. Students could reference their own baseline and immediate postexposure readings (eg, heart rate, blood pressure, and respiratory rate) in slides and discussion to illustrate sympathetic vs parasympathetic patterns.

Inquiry Tasks and Presentation (CoI Aligned)

Small groups completed 3 tasks designed to link embodied responses to conceptual understanding and to support the CoI practical inquiry cycle [15,16].

First, explain observations with evidence—locate and use credible sources (textbooks, peer-reviewed literature, and institutional resources) to interpret observed changes consistent with sympathetic vs parasympathetic mechanisms.

Second, synthesize a brief minilesson—prepare a short slide deck that links observations and theory (autonomic physiology; cardiovascular regulation), applies source appraisal, and communicates clearly to peers. Students could optionally include their own heart rate, blood pressure, and respiratory rate comparisons; content on heart dynamics, signaling pathways, and energy metabolism was addressed via literature-based synthesis.

Third, present and debrief (guided)—deliver the minilesson and then participate in a guided debrief aligned with CoI prompts (eg, triggering event, exploration and integration, and bounded resolution). Facilitation emphasized guided reflection and feedback tailored to the goals of the workshop.

Scaffolds: Prompts and Rubrics

Instructors used structured supports to scaffold the activity. The observation log rubric included baseline vs immediate postexposure readings and salient sensations, along with a brief plain-language interpretation. The source evaluation rubric included authority, accuracy, relevance, and timeliness. The synthesis rubric included claim evidence reasoning, alignment to the learning goal, and clarity for peers. These scaffolds guided observation, evidence appraisal, and coherent explanation and were not graded assessments.

The source evaluation rubric included authority, accuracy, relevance, and timeliness. The synthesis rubric included claim evidence reasoning, alignment to the learning goal, and clarity for peers. These scaffolds guided observation, evidence appraisal, and coherent explanation and were not graded assessments.

Data Collection and Measures (Combined)

Students completed an anonymous postworkshop survey with 4 components.

First, the survey collected demographic data on age, gender, and prior VR headset experience.

Second, the survey included 7 VR-perception items that were workshop-specific and targeted perceived analysis and exploration, conceptual understanding and engagement, teacher support, relevance, and the value of learning new technology. These items were tailored to this workshop and were not intended as a validated general instrument; therefore, reliability statistics were not reported for this subscale.

Third, the CoI instrument (34 items) was used and its wording was adapted minimally to the in-person, single-session context (eg, replacing “course” with “workshop activity”) while preserving full coverage of the constructs (cognitive, social, and teaching presence). Given the small sample, we did not estimate reliability in this study; we relied on prior validation literature [15,16].

Fourth, the survey included 2 open-ended questions addressing challenges, barriers, and opportunities for learning with VR, including reflections that could reference sensations experienced. Teachers (nursing teachers) provided brief reflections, whereas makerspace staff co-facilitated implementation but did not complete teacher items.

XSL•FO
RenderX

Ethical Considerations

Participation was voluntary with written informed consent obtained from all participants. Responses were anonymized to protect privacy and confidentiality. Formal ethics review was not required for this type of educational research at our institution and no sensitive personal data were collected [23]. No compensation was provided to participants.

Analysis

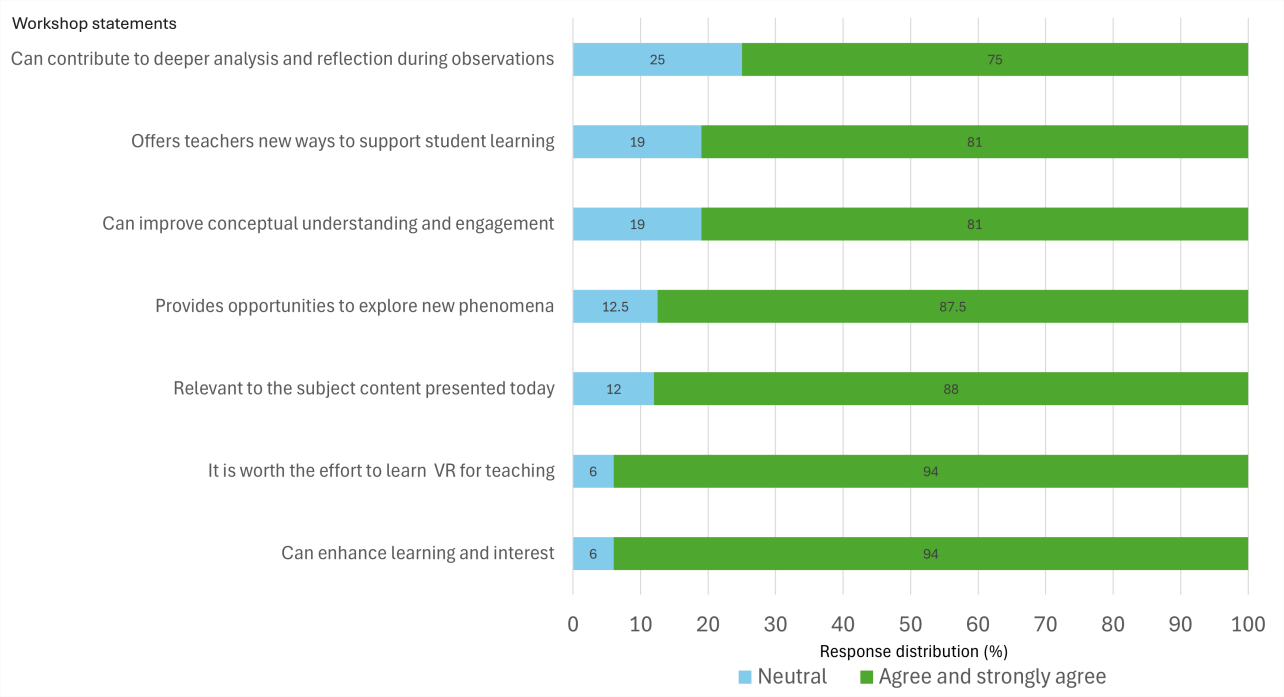
Quantitative data were summarized descriptively (no inferential statistics were performed). The 7-item workshop-specific perceptions were reported as distributions; no reliability statistics were calculated. For the CoI instrument, we reported subscale and total scores descriptively and relied on prior validation literature rather than estimating reliability in this small sample. Open-ended responses underwent summative content analysis (2 coders independently open coded the data and reconciled differences through discussion; categories were not assumed to be mutually exclusive) [24].

Results

Overview of Outcomes

Table 2 summarizes students' CoI scores (teaching, social, and cognitive presence and their subdomains); Figure 1 shows agree and strongly agree distributions for the 7-item VR-perception statements, sorted from highest to lowest.

Figure 1. Nursing student ratings concerning using virtual reality (VR) in education.



XSL•FO
RenderX

Table . Mean score of the community of inquiry (CoI) survey instrument (n=16).

CoI survey elements	Scores, mean (SD)
Teaching presence	3.28 (0.57)
Design and organization	3.42 (0.55)
Facilitation	3.42 (0.55)
Direct instruction	2.88 (0.92)
Social presence	3.26 (0.60)
Affective expression	3.04 (0.65)
Open communication	3.35 (0.68)
Group cohesion	3.38 (0.69)
Cognitive presence	3.48 (0.41)
Triggering event	3.58 (0.46)
Exploration	4.48 (0.49)
Integration	3.46 (0.46)
Resolution	3.42 (0.49)
Total score	3.36 (0.44)

Of the students at the nursing school (n=134), 16 (11.9%) signed up to participate. The study included 16 students (mean age 25, SD 5.1; range 20‐38 years), of whom 15 (93.8%) were women. Of the 16 students, 37.5% (n=6) had prior experience using VR headsets. Among these, 2 (12.5%) students had used VR headsets for gaming purposes and 3 (18.8%) students experienced motion sickness or stress during the use of VR. One (6.3%) student did not provide details about their prior experiences with VR. In addition to the students, 3 teachers (mean age 51; SD 10.0, range 43‐62 years; all male) participated in the workshop. Two teachers had previous experience with VR headsets: one had used them on a few occasions and the other had explored both recreational and pedagogical applications of VR.

Of the 16 students, most agreed or strongly agreed that the VR headsets could increase analysis and observation (n=12, 75%), could help with exploring a new phenomenon (n=14, 87.5%), could increase concept knowledge and engagement (n=13, 81.3%), could create new possibilities for teachers to support their students in learning (n=13, 81.3%), and were relevant to use for the topic of the workshop (n=14, 87.5%) and thought it was worth learning a new technology (n=15, 93.8%; Figure 1).

CoI Results

For interpretability, teaching presence reflects design and organization, facilitation, and direct instruction; social presence reflects affective expression, open communication, and group cohesion; and cognitive presence reflects triggering event, exploration, integration, and resolution. Participants' ratings indicated moderately positive perceptions of the workshop's learning environment (Table 2). Within teaching presence, design and organization and facilitation were similar (mean 3.42, SD 0.55 each), whereas direct instruction was lower (mean 2.88, SD 0.92). Social presence averaged 3.26 (SD 0.60), with subdomains ranging from affective expression (mean 3.04, SD 0.65) to group cohesion (mean 3.38, SD 0.69). Cognitive presence had the highest overall mean 3.48 (SD 0.41), with exploration rated highest among its subdomains (mean 4.48, SD 0.49), followed by triggering event (mean 3.58, SD 0.46), integration (mean 3.46, SD 0.46), and resolution (mean 3.42, SD 0.49). Students referenced their own baseline and immediate postexposure readings (eg, heart rate, blood pressure, and respiratory rate) descriptively to illustrate sympathetic versus parasympathetic patterns; these readings were not analyzed as outcomes.

Qualitative Results (Open-Ended Responses)

Overview

As outlined in the Methods section (summative content analysis), we analyzed anonymous, session-specific open-ended comments. A total of 15 (93.8%) of the 16 students and all 3 teachers provided responses (18/19, 94.7% contributors in total). The unit of analysis was the individual comment. Codes were merged into 3 descriptive categories, and the counts reported represent the number of unique contributors per category. Categories were not mutually exclusive, and the quotations included are illustrative rather than exhaustive [24].

Challenges to VR Use

Participants (9/18, 50%) noted cybersickness and physiological discomfort, headset comfort and time burden, access, logistical, and technical issues, and unfamiliarity and the learning curve.

> *What I can think of is that one can easily become nauseous, and that leads to difficulty concentrating.* [Participant 2402]
>
> *Overall experience can be time-consuming.* [Participant 2403]
>
> *Technical issues. Digital literacy is required in the teacher competency.* [Participant 2419]
>
> *Everyone is not familiar with the technology.* [Participant 2415]

XSL•FO
RenderX

Opportunities for Learning

Students and teachers (16/18, 88.9%) highlighted visualization and immersion to aid understanding and recall, stronger links between theory and observed autonomic responses, low-stakes practice, and a structured debrief that supported collaborative sensemaking.

> *I could experience the connection between theory and practice.* [Participant 2415]
>
> *It allows for your senses to be used in other ways with VR which allows for learning in new ways.* [Participant 2407]
>
> *To experience something and the symptoms that arise with it makes it easier to remember in the learning process.* [Participant 2416]
>
> *To mix learning methods gives you a greater understanding and a wider picture.* [Participant 2411]

Feasibility and Logistics

Practical considerations according to participants (5/18, 27.8%) included equipment availability, onboarding, hygiene, sanitation, and fit adjustments, occasional software and hardware restarts, and scheduling.

> *Technical aspects need to work, both software and hardware.* [Participant 2418]
>
> *Students need guidance to find the right levels where they do not get caught in technical details.* [Participant 2417]
>
> *There might be a need to educate the students in VR-basics.* [Participant 2406]

Immediate sensations (eg, nausea, dizziness, and stress) were self-reported by students during the guided debrief and/or the open-ended survey; vital-sign values were not displayed in-headset.

Interpretation

These brief comments help contextualize the session-level experiences but do not provide evidence of longer-term learning or engagement. The counts are descriptive, and no claims of thematic saturation are made, consistent with STROBE guidance for transparent reporting in cross-sectional studies [19].

Link to Cognitive Presence

The qualitative category “opportunities for learning” (16/18, 88.9%) aligns with several components of cognitive presence. Students’ noticing was primarily subjective during exposure and later referenced to postexposure spot-check readings in the debrief, which together reflect CoI triggering events. Their emphasis on visualization, immersion, and linking theory to physiological responses corresponds to exploration, which also received the highest quantitative subdomain score (mean 4.48, SD 0.49). References to small-group synthesis activities and the guided debrief illustrate integration, as students worked to make sense of information across sources. Comments referring to feasibility constraints and the brief scope of the workshop help explain why resolution was rated lower than exploration (mean 3.42, SD 0.49 vs mean 4.48, SD 0.49), although it remained moderate overall. Together, the open-ended responses complement and contextualize the quantitative profile of cognitive presence while remaining descriptive and limited to session-level experiences.

Instructor Observations

Instructors noted that the format enabled students to observe VR-elicited physiological responses and link these observations to autonomic physiology concepts during the group debrief. These impressions are limited to the session; given the brief, individual VR exposure and the supplemental nature of the workshop, the study does not provide evidence for sustained changes in active or reflective learning or overall engagement. Terms such as “personal” or “meaningful” engagement were not measured. Peer vital-sign measurements were used to facilitate discussion and illustrate concepts (not to evaluate learning outcomes).

Discussion

Principal Findings

This study explored students’ and teachers’ experiences with a brief, optional, VR-enhanced workshop introduced outside mandatory coursework in first-semester nursing education. Students perceived the activity as pedagogically valuable for linking autonomic physiology to embodied experience and for exploration within the learning environment. Interpreted through the CoI framework, the overall profile suggested moderately positive perceptions, with the strongest pattern in cognitive presence: exploration. That pattern is consistent with short, inquiry-driven designs that prompt noticing, questioning, and sensemaking during and after exposure [15,16]. In doing so, this study addresses the identified gap by showing how a brief, optional, low-stakes VR exposure can be positioned early in the curriculum and by documenting students’ and teachers’ perceptions of its pedagogical value and perceived feasibility, with implications for potential scalability as a supplemental strategy. These results reflect session-level perceptions during the workshop’s collaborative elements (peer readings, small-group synthesis and presentations, and guided debrief) and should be interpreted cautiously given the brief, optional, single-session design and small sample (n=16); the study did not assess longer-term changes in engagement or learning.

What the Findings Contribute

A central contribution of this work is the demonstration of a low-stakes, early-semester format in which students experience contrasting autonomic states themselves via a short roller-coaster exposure (sympathetic arousal) and a short guided meditation (parasympathetic engagement) and then interpret those experiences through peer discussion, literature-informed minilessons, and a guided debrief. This sequence appears well-suited to catalyzing curiosity (CoI “triggering event” to subsequent “exploration”) and to making abstract physiological concepts more tangible early in the curriculum [15,16]. The optional nature and the brief duration likely reduced performance pressure while preserving focus on conceptual understanding and engagement, which helps explain the salience of exploration even in a single session. Students’ noticing was primarily subjective during exposure (eg, feelings of arousal or

XSL•FO
RenderX

relaxation) and was later referenced to postexposure spot-check readings during the guided debrief and minilessons, providing concrete anchors for interpretation.

Positioned against the broader VR literature, these session-level perceptions align with findings that VR can enhance learning, particularly motivation, satisfaction, knowledge, and skills, when instructional design and scaffolding are explicit [5-8,25]. At the same time, they are consistent with evidence that presence alone does not guarantee learning and that insufficient scaffolding can increase cognitive load and dampen outcomes [5,9]. Our brief, contrastive exposure format with a structured debrief addresses that design dependence by pairing immersive moments with meaning-making activities rather than treating VR as a standalone novelty.

Interpretation in Context (Education 4.0, Constructive Alignment, and CoI)

The workshop design drew on education 4.0 as a high-level orientation to learner-centered, experience-driven, and digitally supported learning, and on constructive alignment to ensure coherence among intended learning outcomes, teaching and learning activities, and reflection [12-14,18]. Notably, vital sign values were not displayed in-headset; students' "triggering events" arose from subjective sensations during exposure and subsequent reference to postexposure spot checks in the debrief. In practical terms, alignment was enacted by contrastive exposures that instantiate target concepts (autonomic arousal vs relaxation), scaffolded inquiry that connects observations to theory (minilessons with source appraisal), and a brief guided debrief aligned with CoI (cognitive, social, and teaching presence) to support sensemaking and closure [15,16]. The study did not conduct competency assessments and does not claim competence gains; rather, it offers exploratory evidence that a concise, optional format can be perceived as pedagogically valuable and feasible for early supplemental use.

Implementation Considerations

Students' reflections and facilitator notes highlighted pragmatic factors important for routine teaching: headset fit and hygiene protocols, onboarding time for first-time users, and occasional software and hardware restarts. Such start-up frictions are common when integrating immersive devices and should be planned for in scheduling and staffing [5]. The brief contrastive design also served a comfort and accessibility function (limited headset time and seated options), which is advisable given that cybersickness risk is associated with exposure duration, visual stimulation, and locomotion; carefully shorter, comfort-oriented sessions are appropriate for novices [11,22].

Implications and Future Recommendations

The findings of this study suggest several practical implications and directions for future research. Positioning early integration of VR as an optional, low-stakes supplement may be beneficial. Offering brief, clearly scaffolded experiences in the first semester may help to prime curiosity and anchor foundational physiology in embodied experience [5,6].

In addition, using contrastive exposures to make mechanistic contrasts vivid (eg, sympathetic vs parasympathetic) and following these with a guided debrief that traces CoI's practical inquiry cycle from triggering events to bounded resolution may support learning [15,16].

Scaffolding with simple tools (observation logs and source evaluation and synthesis rubrics) allows students to produce evidence-informed explanations without adding grading burden [12,13].

Planning for logistics and comfort includes onboarding, sanitation, fit checks, and rapid troubleshooting, and exposures should be kept brief for first-time users [11,22].

Research next steps include moving beyond session-level perceptions to comparative and mixed methods designs that examine knowledge retention, reflective writing quality, and observational rubrics, as well as assessing accessibility (nonheadset alternatives) and scalability (equipment and time models) across cohorts [7,9,25].

Limitations

This study has several limitations that should be considered when interpreting the findings. First, the activity was optional, single-session, and conducted in a single institution with a small, self-selected sample (n=16), which limits generalizability and precludes inference about longer-term effects. Second, outcomes relied on postsession self-report (VR-perception items and the CoI instrument) without objective educational performance measures, comparators, or follow-up; findings therefore reflect session-level perceptions rather than effectiveness or competence.

Third, vital sign readings (eg, heart rate, blood pressure, and respiratory rate) were collected only as instructional aids to support observation and discussion. In this study, we recorded a resting baseline before exposure and spot-checks immediately after exposure; we did not collect data during exposure or conduct continuous physiological monitoring, and values were obtained with standard training instruments by peers rather than research-grade devices. Consequently, these readings are not suitable for physiological inference, characterization of within-session dynamics, or clean separation of phenomena such as anticipatory arousal. Immediate sensations (eg, nausea, dizziness, and stress) were self-reported during the guided debrief and/or in open-ended survey responses and were not systematically assessed with a validated instrument.

Fourth, the 7-item VR-perception set was tailored to this workshop and was not intended as a validated, generalizable scale; therefore, we do not report reliability for this set. For the CoI instrument, we adapted wording minimally to fit an in-person, single-session context, but given the small sample, we did not estimate reliability in this study and instead relied on prior validation evidence reported elsewhere. Finally, equipment and logistical factors and cross-unit collaboration supported implementation but were not evaluated as outcomes.

Taken together, these factors indicate that results should be interpreted as exploratory evidence about perceived learning environment and feasibility for an early, optional, low-stakes VR activity; they do not demonstrate objective learning gains or competency development. Directions for addressing these

XSL•FO
RenderX

constraints (eg, comparative designs, larger samples, objective outcomes, structured measures of discomfort, and continuous physiological monitoring) are outlined in the Implications and Future Recommendations section.

Conclusions

A brief optional VR workshop introduced early in nursing education can provide low-stakes, embodied experiences that students perceive as helpful for exploring foundational physiological concepts. When paired with scaffolded inquiry and a guided debrief aligned with CoI, short exposures can prompt curiosity and sensemaking within the formal curriculum, offering a practical, scalable supplement to the required coursework [5,15].

Acknowledgments

The authors would like to express their sincere gratitude to the staff at Clinicum, Campus Norrköping, for their invaluable support and assistance throughout the entire day of the workshop. The authors also acknowledge Peter Johansson, Ghassan Mourad, and Johan Lundgren for their valuable contributions in developing this study.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets used and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

MW, NFC, and LK contributed to the conceptualization of the pedagogical workshop. Data curation was performed by MW and LK. Methodology was developed by MW, NFC, and LK. Original draft preparation was conducted by NFC and MW. Writing, review, and editing were carried out by LK.

Conflicts of Interest

None declared.

References

1. Kolb DA. Experiential Learning: Experience as the Source of Learning and Development: Prentice-Hall; 1984. URL: https://books.google.co.in/books/about/Experiential_Learning.html?id=zXruAAAAMAAJ&redir_esc=y [accessed 2026-03-30]
2. Young GW, Stehle S, Walsh BY, Tiri E. Exploring virtual reality in the higher education classroom: using VR to build knowledge and understanding. J Univ Comput Sci 2020;26(8):904-928. [doi: 10.3897/jucs.2020.049]
3. Lowell VL, Yan W. Applying systems thinking for designing immersive virtual reality learning experiences in education. TechTrends 2024 Jan;68:149-160. [doi: 10.1007/s11528-023-00922-1]
4. Freeman S, Eddy SL, McDonough M, et al. Active learning increases student performance in science, engineering, and mathematics. Proc Natl Acad Sci U S A 2014 Jun 10;111(23):8410-8415. [doi: 10.1073/pnas.1319030111] [Medline: 24821756]
5. Radianti J, Majchrzak TA, Fromm J, Wohlgenannt I. A systematic review of immersive virtual reality applications for higher education: design elements, lessons learned, and research agenda. Comput Educ 2020 Apr;147:103778. [doi: 10.1016/j.compedu.2019.103778]
6. Liu K, Zhang W, Li W, Wang T, Zheng Y. Effectiveness of virtual reality in nursing education: a systematic review and meta-analysis. BMC Med Educ 2023 Sep 28;23(1):710. [doi: 10.1186/s12909-023-04662-x] [Medline: 37770884]
7. Ropponen P, Tomietto M, Pramila-Savukoski S, et al. Impacts of VR simulation on nursing students' competence, confidence, and satisfaction: a systematic review and meta-analysis of randomised controlled trials. Nurse Educ Today 2025 Sep;152:106756. [doi: 10.1016/j.nedt.2025.106756] [Medline: 40279771]
8. Cho MK, Kim MY. Enhancing nursing competency through virtual reality simulation among nursing students: a systematic review and meta-analysis. Front Med (Lausanne) 2024;11:1351300. [doi: 10.3389/fmed.2024.1351300] [Medline: 38774395]
9. Makransky G, Terkildsen TS, Mayer RE. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learn Instr 2019 Apr;60:225-236. [doi: 10.1016/j.learninstruc.2017.12.007]
10. Stracke CM, Bothe P, Adler S, et al. Immersive virtual reality in higher education: a systematic review of the scientific literature. Virtual Real 2025 Apr;29:64. [doi: 10.1007/s10055-025-01136-x]
11. Rebenitsch L, Owen C. Review on cybersickness in applications and visual displays. Virtual Real 2016 Jun;20(2):101-125. [doi: 10.1007/s10055-016-0285-9]
12. Biggs J. Enhancing teaching through constructive alignment. High Educ 1996 Oct;32:347-364. [doi: 10.1007/BF00138871]
13. Biggs J, Tang C. Teaching for Quality Learning at University, 4th edition: McGraw-Hill Education; 2011.

XSL•FO
RenderX

14. Defining Education 4.0: a taxonomy for the future of learning. World Economic Forum. 2023. URL: https://www3.weforum.org/docs/WEF_Defining_Education_4.0_2023.pdf [accessed 2026-03-25]
15. Arbaugh JB, Cleveland-Innes M, Diaz SR, et al. Developing a community of inquiry instrument: testing a measure of the Community of Inquiry framework using a multi-institutional sample. Internet High Educ 2008;11(3-4):133-136. [doi: 10.1016/j.iheduc.2008.06.003]
16. Ariati J, Pham T, Vogler JS. Constructivist learning environments: validating the community of inquiry survey for face-to-face contexts. Act Learn High Educ 2025 Mar;26(1):41-57. [doi: 10.1177/14697874231183642]
17. Hwang GJ, Chen YT, Chien SY. A concept map-based community of inquiry framework for virtual learning contexts to enhance students' earth science learning achievement and reflection tendency. Educ Inf Technol 2024;29(12):15147-15172. [doi: 10.1007/s10639-024-12454-z]
18. Mukul E, Büyüközkan G. Digital transformation in education: a systematic review of education 4.0. Technol Forecast Soc Change 2023 Sep;194(5):122664. [doi: 10.1016/j.techfore.2023.122664]
19. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
20. Epic Roller Coasters. B4T Games. URL: https://b4t.games/ [accessed 2026-03-25]
21. TRIPP: meditation, relaxation, sleep on Meta Quest. Meta. URL: https://www.meta.com/en-gb/experiences/tripp-meditation-relaxation-sleep/2173576192720129/?srsltid=AfmBOop-_ambh1Au4RKsp47xOCYCbdva0_A7wwfjSfN638gTvcrYMcWn [accessed 2026-03-25]
22. Saredakis D, Szpak A, Birckhead B, Keage HAD, Rizzo A, Loetscher T. Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis. Front Hum Neurosci 2020;14:96. [doi: 10.3389/fnhum.2020.00096] [Medline: 32300295]
23. Socialstyrelsens råd för etiska frågor. Socialstyrelsen. 2018. URL: https://www.socialstyrelsen.se/om-socialstyrelsen/organisation/rad-och-namnder/socialstyrelsens-rad-for-etiska-fragor [accessed 2025-10-09]
24. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
25. Park S, Shin HJ, Kwak H, Lee HJ. Effects of immersive technology-based education for undergraduate nursing students: systematic review and meta-analysis using the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach. J Med Internet Res 2024 Jul 24;26:e57566. [doi: 10.2196/57566] [Medline: 38978483]

Abbreviations

CoI: community of inquiry
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
VR: virtual reality

Edited by A Sowan; submitted 29.Oct.2025; peer-reviewed by FK Koo, J Afonso; revised version received 10.Feb.2026; accepted 17.Mar.2026; published 02.Apr.2026.

Please cite as:
Carlsson NF, Klompstra L, Westas M
Exploring Nursing Students' Experiences in a Brief Virtual Reality–Enhanced Workshop: Cross-Sectional Exploratory Study
JMIR Nursing 2026;9:e85780
URL: https://nursing.jmir.org/2026/1/e85780
doi:10.2196/85780

© Niklas Ferdinand Carlsson, Leonie Klompstra, Mats Westas. Originally published in JMIR Nursing (https://nursing.jmir.org), 2.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Role of Digital Health Literacy and Cultural Intelligence in Cultural Competence of Private Hospital Nurses: Cross-Sectional Study

Emad Shdaifat[1], PhD; Amira Alshowkan[1], PhD; Bilal Alsaaideh[2], MSN

[1]Community Nursing Department, College of Nursing, Imam Abdulrahman Bin Faisal University, King Faisal Road, Dammam, Eastern Province, Saudi Arabia

[2]Dr. Sulaiman Al Habib Medical Group, Al Khober, Saudi Arabia

Corresponding Author:
Amira Alshowkan, PhD
Community Nursing Department, College of Nursing, Imam Abdulrahman Bin Faisal University, King Faisal Road, Dammam, Eastern Province, Saudi Arabia

Abstract

Background: As patient populations become increasingly diverse, nurses must develop cultural competence to provide effective and equitable care.

Objective: This study aims to examine the predictors of cultural competence among nurses, specifically investigating the roles of digital health literacy (DHL) and cultural intelligence, while also assessing how demographic factors, competence levels, and DHL relate to culturally competent behaviors.

Methods: A cross-sectional study was conducted with 296 nurses at a private hospital in Saudi Arabia. Data were collected using validated instruments to assess DHL, cultural intelligence, and cultural competence. Descriptive statistics, independent *t* tests, Pearson correlations, and hierarchical multiple linear regression analyses were performed using SPSS (IBM Corp) to evaluate the associations and predictors of cultural competence.

Results: In a sample of 296 nurses, cultural intelligence was observed at a moderate level, present in 35.8% (n=106) of participants. Conversely, more than half of the participants (n=158, 53.4%) exhibited low DHL. The majority of the nurses (n=195, 65.9%) demonstrated moderate cultural competence. Education was the only demographic factor significantly associated with cultural competence, with nurses holding a Bachelor of Science degree scoring higher than those with a diploma (P=.003). Furthermore, cultural intelligence and DHL each showed a strong positive correlation with cultural competence (P<.001). Hierarchical regression analyses indicated that both cultural intelligence (β=.427; P<.001) and DHL (β=.315; P<.001) were significant predictors of cultural competence, jointly explaining 47.7% of the variance (adjusted R^2=0.460; P<.001). Nationality also emerged as a significant predictor (β=.105; P=.03), whereas other demographic variables were not statistically significant. Cultural intelligence accounted for the largest proportion of explained variance (ΔR^2=0.375), followed by DHL (ΔR^2=0.051).

Conclusions: Cultural competence among nurses is predominantly influenced by cultural intelligence and DHL, rather than by most demographic variables. Consequently, enhancing DHL and cultural intelligence may strengthen culturally competent practice. It is recommended that targeted educational and professional development interventions be implemented to elevate competence and support high-quality patient care across diverse health care settings.

(JMIR Nursing 2026;9:e97334) doi:10.2196/97334

KEYWORDS

cultural competency; cultural intelligence; digital health literacy; DHL; nurses; Saudi Arabia

Introduction

In an increasingly globalized health care environment, cultural competence has become an essential component of high-quality nursing practice. It enables nurses to deliver care that respects patients' cultural and religious values, beliefs, and practices, thereby improving patient outcomes, satisfaction, and health care equity [1]. In addition to ethnicity and language sensitivity, culturally competent nursing care requires sensitivity to patients' religious beliefs and practices, particularly in multicultural health care settings. Religious values may influence health care decisions regarding informed consent, end-of-life care, organ donation, reproductive health, dietary restrictions, and acceptance of specific medical treatments [2-4]. Previous studies have emphasized that culturally responsive care requires nurses to understand and respect patients' religious and spiritual beliefs, including those associated with Islam, Judaism, Christianity, and other faith traditions, while maintaining ethical and

XSL•FO
RenderX

patient-centered practice [2-4]. These findings further highlight that cultural competence extends beyond communication skills to include religious and spiritual awareness in clinical decision-making [2].

As health care systems become more culturally diverse, nurses must possess the knowledge, skills, and attitudes necessary to communicate effectively and provide culturally responsive care. This includes recognizing how cultural and religious diversity influences patients' preferences, expectations, and health care decisions. Recent international literature also highlights that workforce diversity, intercultural communication, and technology-enhanced health care are increasingly shaping culturally competent nursing practice [5].

Cultural competence is influenced by multiple individual and professional factors, including digital health literacy (DHL) and cultural intelligence [6]. DHL, defined as the ability to search for, evaluate, and apply health information using digital technologies, has become an essential competency for evidence-based nursing practice in technology-driven health care systems [7,8]. Emerging evidence further suggests that digital technologies, including AI-supported educational tools and culturally adapted digital resources, can strengthen nurses' clinical decision-making and promote culturally competent care. Likewise, cultural intelligence, the capability to function effectively across different cultural contexts, enables nurses to adapt their communication and care to the needs of culturally diverse patients [9,10].

International studies have demonstrated that structured educational interventions incorporating experiential learning, intercultural communication, reflective practice, and technology-enhanced learning improve both cultural intelligence and cultural competence among nursing students and practicing nurses [9,11,12]. In addition, nurses with higher DHL are more likely to use evidence-based clinical guidelines, patient education resources, and culturally appropriate digital tools to deliver individualized care [13,14]. Park and Nahm [15] further reported that nurses' use of point-of-care technologies is associated with greater confidence in clinical decision-making and participation in culturally sensitive evidence-based practice.

The importance of these competencies is particularly evident in Saudi Arabia, where a rapidly evolving health care system serves an increasingly multicultural patient population [5]. Previous studies conducted in Saudi Arabia have identified gaps in cultural competence among both nursing students and practicing nurses, largely attributed to limited exposure to transcultural education and culturally sensitive communication training [5,16]. Hamdan Alshammari and Alboliteeh [6] further emphasized that professional competence and cultural competence are closely interconnected, suggesting that both education and clinical experience contribute to culturally congruent nursing care.

Demographic characteristics such as age, educational level, and years of clinical experience may further influence cultural competence, DHL, and cultural intelligence [13,17]. Nurses with greater clinical experience often demonstrate stronger cultural competence due to prolonged exposure to diverse patient populations, whereas younger nurses may possess stronger DHL but require additional support in translating these digital skills into culturally responsive clinical practice [7,18-20].

Although previous studies have independently demonstrated that DHL and cultural intelligence contribute to cultural competence, most research has focused on educational interventions or specific dimensions of cultural competence, whereas limited empirical evidence has examined how DHL and cultural intelligence jointly predict cultural competence among practicing nurses, particularly within the Saudi health care context [1,8]. Furthermore, the extent to which demographic characteristics contribute to cultural competence alongside these factors remains insufficiently explored [6,14]. Addressing this gap may inform educational strategies and workforce development initiatives to strengthen culturally competent nursing care in increasingly diverse health care settings.

Therefore, the purpose of this study was to investigate the predictors of cultural competence among nurses by examining the roles of DHL and cultural intelligence while controlling for demographic variables and competence levels.

Methods

Design and Setting

A cross-sectional design was used for this study. The research was conducted in a private hospital situated in the Eastern region of Saudi Arabia. Registered nurses from various clinical units, including emergency, intensive care, medical-surgical, and outpatient departments, were invited to participate. These diverse clinical settings offered a suitable context for investigating DHL, cultural intelligence, and cultural competence within the framework of routine nursing practice. This study constitutes an in-depth analysis of an existing dataset previously used to test a structural model of culturally competent behavior. In the current analysis, a multiple regression approach was used to examine the individual predictive effects of DHL, cultural intelligence, and demographic variables on cultural competence behavior.

Participants and Sampling

A convenience sampling method was used to recruit eligible registered nurses. The inclusion criteria specified that participants must be registered nurses engaged in direct patient care with a minimum of 6 months of clinical experience. Nurses occupying primarily administrative or managerial positions were excluded from the study due to their limited direct patient interaction. An a priori power analysis conducted using G*Power (Heinrich Heine University Düsseldorf) determined that a minimum sample size of 222 participants was necessary to identify a small-to-moderate effect size (f^2=0.05) at a significance level of α=.05, with 80% statistical power in a multivariable regression model.

Measures

DHL was evaluated using the eHealth Literacy Scale, an 8-item instrument designed to measure individuals' perceived capacity to locate, assess, and apply online health information. Each item was rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) [21].

XSL•FO
RenderX

Cultural intelligence was measured using the 20-item Cultural Intelligence Scale, which encompasses metacognitive, cognitive, motivational, and behavioral domains. Responses were recorded on a 7-point Likert scale, with values ranging from 1 (strongly disagree) to 7 (strongly agree) [22].

Cultural competence was assessed through an 8-item behavioral scale derived from established instruments measuring cultural competence. The items targeted observable culturally responsive nursing behaviors and were rated on a 5-point scale, where 1 indicated "never" and 5 indicated "always" [23]. In the present study, all instruments exhibited excellent internal consistency reliability, with Cronbach α exceeding 0.95.

Demographic data included age, gender, marital status, education level (diploma vs Bachelor of Science [BSc] in Nursing), specialty, years of experience, and nationality. In Saudi Arabia, a nursing diploma is conferred upon completion of a 2 - to 3-year program at a health institute, whereas a BSc in Nursing requires successful completion of a 4- to 5-year university degree.

Data Collection Procedure

The survey was administered using QuestionPro (QuestionPro Inc) in February 2026. Access was restricted to invited participants via a unique link. Nurses were invited through hospital nursing departments via email and WhatsApp. Of the 411 nurses who viewed the survey, 296 completed it (participation rate of 72.0%). IP address checking and cookies were enabled to prevent duplicate submissions. Incomplete questionnaires were excluded from analysis. Responses completed in less than 1 minute were excluded to ensure data quality.

Ethical Considerations

Ethical approval was secured from the Institutional Review Board of the participating hospital, Dr Sulaiman Al Habib Medical Group (RC26.02.08). Participation in the study was voluntary, and informed consent was obtained prior to the commencement of data collection. No identifiable information was gathered and all responses were treated with confidentiality. Data were stored securely and were accessible solely to the research team. The study complied with the ethical principles delineated in the Declaration of Helsinki.

Data Analysis

Data were analyzed using SPSS (version 22; IBM Corp). Descriptive statistics were computed for demographic variables and study constructs. Categorical variables were summarized using frequencies and percentages, while continuous variables were reported as means and SDs.

Levels of cultural competence, cultural intelligence, and DHL were categorized using percentile-based cutoffs for descriptive purposes. Categories were established as low, moderate, and high through visual binning based on percentile distribution in SPSS. This methodology facilitated the empirical derivation of cutoffs, reflecting the natural distribution of scores within the sample and enabling meaningful interpretation and comparative analysis across constructs.

Independent samples *t* tests were used to explore differences in competence across various demographic groups. Pearson correlation analysis was performed to investigate relationships among the primary study variables. Hierarchical multiple linear regression was conducted to identify predictors of cultural competence. Demographic variables were entered in the first block, cultural intelligence in the second block, and DHL in the final block. Statistical significance was established at $P<.05$.

Prior to regression analysis, all assumptions were rigorously tested. Normality of residuals was assessed using the Shapiro-Wilk test and visual inspection of P-P plots. Linearity was confirmed through scatterplots of standardized residuals against predicted values. Homoscedasticity was assessed via visual inspection of residual plots, and independence of errors was confirmed using the Durbin-Watson statistic, which fell within an acceptable range. All assumptions were met. Multicollinearity was assessed using tolerance values. All tolerance values exceeded 0.10 (ranging from 0.511 to 0.944), indicating no serious multicollinearity among predictors.

Results

Internal consistency was evaluated using Cronbach α for all study scales (N=296). The Cultural Intelligence Scale, consisting of 20 items, exhibited excellent reliability (α=0.969). The eHealth Literacy Scale, comprising 8 items, similarly demonstrated high internal consistency (α=0.959). Likewise, the Cultural Competence Scale, which also included 8 items, reflected strong reliability (α=0.957). The corrected item-total correlations for all scales exceeded 0.65, and the Cronbach α values did not significantly increase upon the deletion of any item, indicating robust homogeneity and stability of the items within each construct.

The study's sample comprised 296 participants, with a predominance of females (n=285, 96.3%), while males constituted only 3.7% (n=11) of the cohort. More than half of the participants were single (n=171, 57.8%), whereas 42.2% (n=125) were married. The majority of respondents held a BSc degree (n=221, 74.7%), while 25.3% (n=75) possessed a diploma qualification. In terms of clinical specialty, the majority were employed in non–critical care unit settings (n=234, 79.1%), with 20.9% (n=62) working in critical care environments. Furthermore, most participants were non-Saudi (n=225, 76.0%), while 24.0% (n=71) identified as Saudi nationals.

The mean age of the participants was 30.36 (SD 5.8) years, with a median age of 29 (IQR 25-34) years and a range spanning 23 to 51 years, indicating a relatively young sample. The average years of professional experience was 4.97 (SD 4.8) years, with a median of 3 (IQR 1-7) years and a range of 1 to 24 years, suggesting moderate variability in clinical experience among participants (Table 1).

XSL•FO
RenderX

Table . Demographic characteristics of participants (N=296).

Variable and category	Participants
Gender	
Male, n (%)	11 (3.7)
Female, n (%)	285 (96.3)
Marital status	
Single, n (%)	171 (57.8)
Married, n (%)	125 (42.2)
Education	
Diploma, n (%)	75 (25.3)
BSc[a], n (%)	221 (74.7)
Specialty	
Critical care unit, n (%)	62 (20.9)
Non–critical care unit, n (%)	234 (79.1)
Nationality	
Saudi, n (%)	71 (24.0)
Non-Saudi, n (%)	225 (76.0)
Age (years)	
Mean (SD)	30.36 (5.8)
Median (IQR; range)	29 (25-34; 23-51)
Experience (years)	
Mean (SD)	4.97 (4.8)
Median (IQR; range)	3 (1-7; 1-24)

[a]BSc: Bachelor of Science.

The mean cultural competence score among participants was 28.62 (SD 4.2), with the majority, 65.9% (n=195), classified as possessing moderate competence and 34.1% (n=101) exhibiting low competence. Regarding cultural intelligence, the mean score was 85.87 (SD 14.0), with participants relatively evenly distributed across the competency levels: 35.8% (n=106) categorized as moderate, 34.1% (n=101) as low, and 30.1% (n=89) as high. The average DHL score was 33.16 (SD 5.9), indicating that over half of the participants (n=158, 53.4%) were classified as having low DHL, while 18.6% (n=55) fell into the moderate category and 28.0% (n=83) were classified as high. Overall, while cultural competence and cultural intelligence scores were predominantly moderate, a significant proportion of nurses demonstrated low DHL, underscoring a critical area that requires targeted professional development (Table 2).

Table . Levels of cultural competence, cultural intelligence, and digital health literacy of participants (N=296).

Variable and category	Score, mean (SD)	Participants, n (%)
Cultural competence	28.62 (4.2)	
Low (≤27)		101 (34.1)
Moderate (28–32)		195 (65.9)
Cultural intelligence	85.87 (14.0)	
Low (≤81)		101 (34.1)
Moderate (82-96)		106 (35.8)
High (>96)		89 (30.1)
Digital health literacy	33.16 (5.9)	
Low (≤32)		158 (53.4)
Moderate (33-37)		55 (18.6)
High (>37)		83 (28.0)

XSL•FO
RenderX

Independent samples *t* tests were conducted to assess differences in cultural competence across various demographic groups. A statistically significant difference was identified solely in relation to education level. Participants holding a BSc degree (mean 29.06, SD 4.14) exhibited significantly higher competence scores compared to those holding a diploma (mean 27.36, SD 4.40; t_{294}=−3.022; *P*=.003). No statistically significant differences in competence were detected based on gender, marital status, specialty, nationality, age category, or years of experience (all *P*>.05). These findings suggest that educational qualification is the only demographic factor associated with elevated cultural competence scores within this sample (Table 3).

Table . Differences in cultural competence across demographic groups.

Variable and group	Participants, n (%)	Cultural competence score, mean (SD)	*t* test (*df*)	*P* value
Gender			−0.425 (294)	.67
Male	11 (3.7)	28.09 (4.95)		
Female	285 (96.3)	28.65 (4.24)		
Marital status			−1.339 (294)	.18
Single	171 (57.8)	28.35 (4.38)		
Married	125 (42.2)	29.02 (4.09)		
Education			−3.022 (294)	.003
Diploma	75 (25.3)	27.36 (4.40)		
BSc[a]	221 (74.7)	29.06 (4.14)		
Specialty			0.805 (294)	.42
Critical care unit	62 (20.9)	29.02 (3.83)		
Non–critical care unit	234 (79.1)	28.53 (4.37)		
Nationality			0.490 (294)	.62
Saudi	71 (24.0)	28.85 (4.15)		
Non-Saudi	225 (76.0)	28.56 (4.31)		
Age (years)			−0.236 (294)	.81
≤29	157 (53.0)	28.57 (4.42)		
>29	139 (47.0)	28.69 (4.09)		
Experience (years)			−0.437 (294)	.66
≤3	156 (52.7)	28.53 (3.93)		
>3	140 (47.3)	28.74 (4.62)		

[a]BSc: Bachelor of Science.

Pearson correlation analysis revealed the presence of strong and statistically significant positive relationships among all study variables. Cultural intelligence exhibited a robust correlation with DHL (*r*=0.694; *P*<.001) and a moderate to strong correlation with cultural competence (*r*=0.637; *P*<.001). Furthermore, DHL showed a strong positive association with cultural competence (*r*=0.610; *P*<.001). These findings suggest that higher levels of cultural intelligence and DHL are associated with higher levels of cultural competence. All correlations were significant at the .01 level (2-tailed), indicating substantial interrelationships among the constructs (Table 4).

XSL•FO
RenderX

Table . Pearson correlations among cultural intelligence, digital health literacy, and cultural competence (N=296).

Variable	Cultural intelligence	Digital health literacy	Cultural competence
Cultural intelligence			
r	1	0.694	0.637
P value	—[a]	<.001	<.001
Digital health literacy			
r	0.694	1	0.610
P value	<.001	—	<.001
Cultural competence			
r	0.637	0.610	1
P value	<.001	<.001	—

[a]Not applicable.

The final model accounted for 47.7% of the variance in cultural competence (adjusted R^2=0.460; *P*<.001). After controlling for demographic variables, both cultural intelligence (β=.427; *P*<.001) and DHL (β=.315; *P*<.001) emerged as significant positive predictors of cultural competence. Nationality also reached significance (β=.105; *P*=.03), while age, gender, marital status, education, specialty, and experience did not reach significance in the final model. Cultural intelligence accounted for the largest proportion of explained variance (ΔR^2=0.375), followed by DHL (ΔR^2=0.051). See Table 5 for breakdown by predictor variable and Table 6 for models.

Table . Predictor variable results for hierarchical multiple linear regression predicting cultural competence (N=296).

Predictor	B	SE	β	*t* test (*df*)	*P* value
Constant	6.53	2.979	—[a]	2.192 (286)	.03
Age	−0.02	0.052	−.027	−0.376 (286)	.71
Gender	0.358	0.987	.016	0.363 (286)	.72
Marital status	0.259	0.441	.03	0.587 (286)	.56
Education	0.61	0.434	.062	1.408 (286)	.16
Specialty	−0.053	0.483	−.005	−0.109 (286)	.91
Experience	0.027	0.064	.03	0.413 (286)	.68
Nationality	1.051	0.48	.105	2.19 (286)	.03
Cultural intelligence	0.13	0.019	.427	6.978 (286)	<.001
Digital health literacy	0.226	0.043	.315	5.261 (286)	<.001

[a]Not applicable.

Table . Hierarchical multiple linear regression predicting cultural competence (N=296).

Model	*R*	R^2	Adjusted R^2	ΔR^2	*F* (*df*)	*P* value
(1) Demographics	0.226	0.051	0.028	—[a]	2.208 (7, 288)	.03
(2) + Cultural intelligence	0.653	0.426	0.41	0.375	26.642 (8, 287)	<.001
(3) + Digital health literacy	0.69	0.477	0.46	0.051	28.958 (9, 286)	<.001

[a]Not applicable.

Discussion

Principal Findings

This study examined predictors of cultural competence among nurses, focusing on DHL and cultural intelligence, and explored associations with demographic factors and competence levels. Cultural intelligence was identified at a moderate level. In contrast, over half of the participants demonstrated low DHL. Most nurses (n=195; 65.9%) displayed moderate cultural competence. Education was the only demographic factor significantly associated with cultural competence, with BSc holders scoring higher than diploma holders. Cultural

XSL•FO
RenderX

intelligence and DHL were positively correlated with competence, and hierarchical regression revealed that both significantly predicted cultural competence, explaining a substantial portion of its variance.

The results showed that nurses' cultural intelligence was moderate. This finding is consistent with previous studies reporting moderate levels of cultural intelligence among nursing students and health care professionals [9,10]. Higher levels have been observed among nurse educators, suggesting that professional experience and continuous professional development may enhance cultural intelligence [16]. The moderate levels observed in this study are likely due to a lack of formal education in cultural intelligence and limited opportunities for structured intercultural learning. In Saudi Arabia, nurses are often exposed to culturally diverse patients, which can lay the foundation for intercultural interactions; however, this exposure alone may not be sufficient to develop advanced cultural intelligence without targeted educational programs [1].

More than half of the participants had low DHL, suggesting difficulties accessing, appraising, and applying digital health information. This finding is consistent with previous studies reporting inadequate DHL among nursing students and nurses [7,8]. Other studies have also highlighted considerable variability in DHL and emphasized the need for systematic educational interventions [13]. Conversely, higher levels of DHL have been reported among nurses with higher educational qualifications and professional positions, indicating that educational and workplace factors can positively influence digital competency [14]. Low DHL levels in the current study may be explained by a lack of training in using digital platforms, limited institutional support, and factors such as low confidence, restricted access, and the rapid pace of technological change [15].

Most nurses had moderate levels of cultural competence, consistent with previous studies among practicing nurses and nursing students in Saudi Arabia [1,5]. Moderate levels of cultural competence have commonly been reported in settings with limited formal transcultural nursing education, whereas greater professional experience has been associated with higher levels of competence [6,16]. The findings of this study may be due to a combination of theoretical knowledge gained through education and practical experience with different patients. However, the absence of continuous, structured programs for cultural competence may limit further development toward higher levels of competence [11,12].

Among the demographic variables, education was the only factor that significantly predicted cultural competence, with BSc holders outperforming diploma holders. This finding supports previous evidence suggesting that higher educational attainment is associated with stronger professional and cultural competencies [6,14]. However, previous research has also indicated that clinical experience and continuing professional education may enhance cultural competence regardless of initial educational preparation [1]. The current study's results are likely a reflection of the fact that BSc courses involve more transcultural topics, evidence-based practice, intercultural communication, and digital literacy, which, when combined, improve cultural competence [7,9].

Both cultural intelligence and DHL were positively correlated with cultural competence, suggesting that nurses with high adaptability to diverse cultural settings and strong digital literacy skills exhibit greater cultural competence. This finding is consistent with previous research demonstrating that cognitive-cultural capabilities and digital competencies jointly contribute to culturally responsive nursing practice [9,10,13]. Conceptually, cultural intelligence and DHL represent complementary competencies. Cultural intelligence enables nurses to understand, interpret, and adapt to patients' cultural beliefs, values, and communication styles, whereas DHL equips nurses to locate, evaluate, and apply culturally appropriate health information and digital resources. Together, these competencies facilitate informed clinical decision-making, individualized patient care, and effective communication across diverse cultural contexts. Although technological competence is important, previous studies suggest that digital skills alone are insufficient to ensure culturally competent practice without the motivation and ability to adapt to patients' cultural needs [15]. The results therefore indicate that nurses with high cultural intelligence can decode patients' cultural messages and respond appropriately, while those with high DHL can effectively use digital resources to support culturally competent care [7,8].

Hierarchical regression analysis also showed that cultural intelligence and DHL were significant predictors of cultural competence, accounting for a large proportion of the variance in scores. This finding reinforces previous evidence that cognitive-cultural capabilities and digital competencies are both important determinants of culturally competent nursing practice [9,16]. More importantly, the results support the multidimensional nature of cultural competence, demonstrating that effective culturally competent care depends on both interpersonal adaptability and technology-related competencies. Cultural intelligence facilitates adaptive cross-cultural interactions, whereas DHL enables nurses to integrate evidence-based digital resources into culturally responsive care. Their combined contribution suggests that neither competency alone is sufficient; rather, culturally competent nursing practice emerges from integrating both human-centered and digital capabilities. Nurses with adaptive cultural intelligence and DHL are therefore more likely to implement evidence-based, culturally responsive interventions that improve the quality of patient care [12]. Although these individual competencies explained a substantial proportion of cultural competence, organizational factors such as leadership support, access to culturally appropriate digital resources, and continuing professional development may further strengthen culturally competent practice. Future research should examine these organizational influences using longitudinal and theory-driven study designs [1].

Recommendations

Based on the results of this study, the following recommendations can be made. Nursing education programs should integrate cultural intelligence and DHL into undergraduate curricula and continuing professional

XSL•FO
RenderX

development to strengthen nurses' ability to deliver culturally competent care. Simulation-based intercultural communication training and workshops focusing on culturally sensitive digital health services should also be incorporated to enhance both cultural adaptability and effective use of digital health technologies. Additionally, health care organizations should establish institutional policies that promote equitable access to culturally appropriate digital resources, provide ongoing technical and educational support, and foster culturally inclusive work environments through leadership initiatives. Efforts should also be directed toward improving digital accessibility by addressing barriers such as limited technological confidence, inadequate access to digital resources, and unavailability of linguistically and culturally appropriate digital health information for diverse patient populations.

As AI and digital health technologies become increasingly integrated into health care, future educational and organizational initiatives should explore their role in supporting culturally competent nursing practice, including clinical decision support, culturally tailored patient education, and communication across multicultural health care settings. Strengthening these competencies has the potential to improve health care quality, patient safety, and patient satisfaction. For future studies, longitudinal and mixed-methods research is recommended to provide a more comprehensive understanding of the development of cultural competence over time. Future studies should also investigate additional individual and organizational factors, such as leadership support, organizational culture, digital infrastructure, and access to professional development, which may influence cultural intelligence, DHL, and cultural competence among nurses in Saudi Arabia.

Limitations

Although this study identified significant predictors of cultural competence among nurses, it has some limitations. Data were collected from a single private setting, limiting generalizability, and the cross-sectional design prevents causal inferences. Additionally, reliance on self-reported questionnaires may introduce response bias. Future studies should use larger, more diverse samples, longitudinal designs, and objective measures to strengthen the findings.

Conclusions

This study identified DHL and cultural intelligence as key predictors of nurses' cultural competence. While cultural intelligence was moderate, over half of the participants had low DHL, and most showed moderate competence. Education was the only significant demographic factor, with BSc holders scoring higher. Both cultural intelligence and DHL significantly predicted competence, highlighting the need for health care institutions to provide supportive environments with culturally appropriate resources and technology to enhance culturally competent care.

Funding

The authors declared no financial support was received for this work.

Data Availability

Data are available from the corresponding author upon request.

Conflicts of Interest

None declared.

References

1. Saeed NA, Fayiz SM, Abdullah MM, et al. Cultural competence in nursing care for Saudi patients. Rev Diabet Stud 2024 Jun 10;20(S6):883-890. [doi: 10.70082/xbw3qa26]
2. Murgia C, Notarnicola I, Caruso R, De Maria M, Rocco G, Stievano A. Spirituality and religious diversity in nursing: a scoping review. Healthcare (Basel) 2022 Aug 31;10(9):1661. [doi: 10.3390/healthcare10091661] [Medline: 36141273]
3. Bokek-Cohen Y, Gonen LD, Tarabeih M. The ethical standards of Sunni Muslim physicians regarding fertility technologies that are religiously forbidden. J Relig Health 2022 Aug;61(4):2876-2904. [doi: 10.1007/s10943-022-01583-9] [Medline: 35616821]
4. Bokek-Cohen Y, Tarabeih M. The use of porcine-derived materials for medical purposes: what do Muslim and Jewish individuals know and opine about it? J Bioeth Inq 2022 Dec;19(4):599-612. [doi: 10.1007/s11673-022-10203-w] [Medline: 36063344]
5. Alharbi K, Bakarman SS, Alqahtani N, Bashatah A, Syed W. Exploring cultural competence knowledge, skills, and comfort among male nursing students in Riyadh, Saudi Arabia. BMC Med Educ 2025 Jul 24;25(1):1106. [doi: 10.1186/s12909-025-07666-x] [Medline: 40707933]
6. Hamdan Alshammari M, Alboliteeh M. Structural equation modeling of the association between professional and cultural competencies of nurses in Saudi Arabia. Nurse Educ Pract 2022 Aug;63:103382. [doi: 10.1016/j.nepr.2022.103382] [Medline: 35772307]
7. Abou Hashish EA, Alnajjar H. Digital proficiency: assessing knowledge, attitudes, and skills in digital transformation, health literacy, and artificial intelligence among university nursing students. BMC Med Educ 2024 May 7;24(1):508. [doi: 10.1186/s12909-024-05482-3] [Medline: 38715005]

XSL•FO
RenderX

8. Ramadan OME, Elsharkawy NB, Hafiz AH, et al. Neonatal nurses' e-health literacy and technology-mediated clinical practice: a cross-sectional analysis of digital health competencies and practice patterns. BMC Nurs 2025 Sep 26;24(1):1199. [doi: 10.1186/s12912-025-03839-7] [Medline: 41013614]
9. Majda A, Zalewska-Puchała J, Bodys-Cupak I, Kurowska A, Barzykowski K. Evaluating the effectiveness of cultural education training: cultural competence and cultural intelligence development among nursing students. Int J Environ Res Public Health 2021 Apr 11;18(8):4002. [doi: 10.3390/ijerph18084002] [Medline: 33920334]
10. Al-Otaibi N. Cultural intelligence and its relationship to digital citizenship skills among university students. Humanit Soc Sci Commun 2025 Nov 24;12(1):1816. [doi: 10.1057/s41599-025-06096-w]
11. O'Brien EM, O' Donnell C, Murphy J, O' Brien B, Markey K. Intercultural readiness of nursing students: an integrative review of evidence examining cultural competence educational interventions. Nurse Educ Pract 2021 Jan;50:102966. [doi: 10.1016/j.nepr.2021.102966] [Medline: 33454512]
12. Ozkara San E, Lim F. Enhancing cultural competence in nursing education through generative artificial intelligence tools. J Transcult Nurs 2026 Mar;37(2):343-348. [doi: 10.1177/10436596251372944] [Medline: 41186118]
13. Liu L, Fu M, Wu J, et al. Digital health literacy among undergraduate nursing students in China: associations with health lifestyles and psychological resilience. BMC Med Educ 2024 Oct 14;24(1):1139. [doi: 10.1186/s12909-024-06075-w] [Medline: 39402500]
14. Comparcini D, Simonetti V, Tomietto M, et al. The relationship between nurses' digital health literacy and their educational levels, professional roles, and digital attitudes: a cluster analysis based on a cross-sectional study. J Clin Nurs 2025 Jul;34(7):2885-2897. [doi: 10.1111/jocn.17484] [Medline: 39381920]
15. Park J, Nahm ES. Clinical nurses' point-of-care technology use and digital health competency: scoping review. Comput Inform Nurs 2026 Jul 1;44(7):e01465. [doi: 10.1097/CIN.0000000000001465] [Medline: 41562426]
16. Abou Hashish EA, Alnajjar H, Bajamal EZ, AlJohani HAR, Bahawi RA. Cultural competence and transcultural teaching behaviors and influencing factors in nursing academia: a mixed-methods study among nurse educators. SAGE Open Nurs 2025 Sep 4;11:23779608251375375. [doi: 10.1177/23779608251375375] [Medline: 40919285]
17. Zhao BY, Chen MR, Lin R, Yan YJ, Li H. Influence of information anxiety on core competency of registered nurses: mediating effect of digital health literacy. BMC Nurs 2024 Sep 6;23(1):626. [doi: 10.1186/s12912-024-02275-3] [Medline: 39243073]
18. Fitzpatrick PJ. Improving health literacy using the power of digital communications to achieve better health outcomes for patients and practitioners. Front Digit Health 2023 Nov 17;5:1264780. [doi: 10.3389/fdgth.2023.1264780] [Medline: 38046643]
19. Holt KA, Overgaard D, Engel LV, Kayser L. Health literacy, digital literacy and eHealth literacy in Danish nursing students at entry and graduate level: a cross sectional study. BMC Nurs 2020 Apr 10;19(1):22. [doi: 10.1186/s12912-020-00418-w] [Medline: 32308559]
20. Yang Y. Influences of digital literacy and moral sensitivity on artificial intelligence ethics awareness among nursing students. Healthcare (Basel) 2024 Oct 31;12(21):2172. [doi: 10.3390/healthcare12212172] [Medline: 39517384]
21. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
22. Van Dyne L, Ang S, Koh C. Cultural intelligence: measurement and scale development. In: Moodian MA, editor. Contemporary Leadership and Intercultural Competence: Exploring the Cross-Cultural Dynamics within Organizations: SAGE Publications; 2009:233-254. [doi: 10.4135/9781452274942.n18]
23. Perng SJ, Watson R. Construct validation of the Nurse Cultural Competence Scale: a hierarchy of abilities. J Clin Nurs 2012 Jun;21(11-12):1678-1684. [doi: 10.1111/j.1365-2702.2011.03933.x] [Medline: 22239136]

Abbreviations

BSc: Bachelor of Science
DHL: Digital Health Literacy

Edited by E Borycki; submitted 06.Apr.2026; peer-reviewed by LM Minshew, M Tarabeih; revised version received 30.Jul.2026; accepted 07.Sep.2026; published 22.Sep.2026.

Please cite as:
Shdaifat E, Alshowkan A, Alsaaideh B
The Role of Digital Health Literacy and Cultural Intelligence in Cultural Competence of Private Hospital Nurses: Cross-Sectional Study
JMIR Nursing 2026;9:e97334
URL: https://nursing.jmir.org/2026/1/e97334
doi:10.2196/97334

XSL•FO
RenderX

© Emad Shdaifat, Amira Alshowkan, Bilal Alsaaideh. Originally published in JMIR Nursing (https://nursing.jmir.org), 22.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Knowledge Acquisition, Case Discussion, and Engagement in Health Professional Students Using Interactive Virtual Patient Cases Versus Written Case Studies: Randomized Controlled Trial

Jeffrey J Borckardt[1,2,3], PhD; Debra Henninger-Borckardt[1,2], PhD; Kelly Barth[1], DO; Dusti Annan-Coultas[1], EdD; Lisa Langdale[1], MSN, EdD; Kimberly Kascak[1], MS; Christopher Pelic[1], MD; Michael A. de Arellano[1], PhD; Kathleen Brady[1], MD, PhD

[1]Department of Psychiatry and Behavioral Sciences, College of Medicine, Medical University of South Carolina, 67 President Street, Charleston, SC, United States

[2]Palmetto Innovative Education, LLC, Mount Pleasant, SC, United States

[3]Office of Interprofessional Initiatives, Medical University of South Carolina, Charleston, SC, United States

Corresponding Author:
Jeffrey J Borckardt, PhD
Department of Psychiatry and Behavioral Sciences, College of Medicine, Medical University of South Carolina, 67 President Street, Charleston, SC, United States

Abstract

Background: Effective pain and opioid management education remains a persistent challenge in health professions training. Static case studies and lecture-based instruction are widely used but may be insufficient to develop the clinical reasoning and communication skills required in practice. Interactive virtual patient simulations, including those powered by large language models, offer scalable alternatives, but head-to-head comparisons of interactive and static case formats using identical content among interprofessional learner populations remain limited.

Objective: This randomized controlled trial examined whether learner knowledge, quiz performance, and engagement differed when educational case studies were delivered through an interactive, AI-enabled virtual patient format versus a static written format. Because the formats differ in interactivity as well as in their use of AI, the trial was designed to test the effect of an "interactive virtual patient format" rather than to isolate any AI-specific effect.

Methods: Health professional students from medicine, nursing, dentistry, pharmacy, and allied health programs at a US academic health sciences university viewed a 12-minute video on pain management and were then randomized 1:1 to complete either 3 static written case studies or 3 interactive virtual patient case studies using the Figment Learning Labs Virtual Clinic. Case content was identical across conditions; only the delivery format was manipulated. The virtual patient system used an AI-enabled conversational interface (GPT-4o) delivering facilitator-authored case material, including a viewable medical record. Knowledge was assessed preintervention and postintervention using an 8-item test. Case discussion quiz accuracy and learner engagement (17 items) were also measured. Of 74 enrolled participants, 58 (78%) completed all study components and were included in this completer-based analysis. Analyses included a 2×2 mixed ANOVA and 2-tailed independent-sample *t* tests (α=.05).

Results: A significant time (pre-post)×group (interactive vs written) interaction was observed ($F_{1,56}$=4.39; P=.04; partial η^2=0.07), indicating greater knowledge gains in the interactive condition. Participants in the virtual patient group demonstrated higher case discussion quiz accuracy (mean 0.79, SD 0.15) than those in the written case group (mean 0.64, SD 0.11; t_{56}=4.48; P<.001). Agreement scores were significantly higher in the interactive group on 13 of 17 engagement items (P<.05), including items addressing realism, enjoyment, and perceived clinical relevance; because these were uncorrected item-level comparisons, they should be read as supportive rather than definitive.

Conclusions: In this preliminary trial, exposure to interactive virtual patient cases was associated with higher engagement and higher short-term knowledge scores than static written cases. The design cannot isolate the independent contribution of AI, conversational interactivity, or time-on-task; the findings therefore support the value of an interactive virtual patient "format" rather than a distinct AI-specific advantage. Because the analysis was completer-based, featured unexplained attrition, and involved 2 authors who are founders of the company commercializing the platform, independent replication in larger, multi-institutional, and longitudinal samples is needed.

Trial Registration: ClinicalTrials.gov NCT06624124; https://clinicaltrials.gov/study/NCT06624124

(JMIR Nursing 2026;9:e98700) doi:10.2196/98700

XSL•FO
RenderX

KEYWORDS

virtual patients; AI; large language models; ChatGPT; GPT-4o; interprofessional education; health professions education; simulation-based education; learner engagement; pain management education; randomized controlled trial

Introduction

Background

Effective pain management education, particularly regarding opioid risks, remains a challenge in health professions training. Despite initiatives such as the Pain Core Curriculum [1], critical knowledge gaps persist. Traditional approaches, including lectures, readings, and static case studies, often fall short of building the clinical reasoning skills required to address the complexities of patient-centered pain management [2]. Clinical rotations, while valuable, are constrained by availability, patient safety concerns, limited case exposure, and institutional barriers, leaving many students underprepared for nuanced decision-making in clinical practice [3,4].

Simulation-based learning (SBL) offers a solution, enabling practice, feedback, and skill development in complex encounters [5,6]. However, existing SBL formats involve distinct trade-offs in fidelity, cost, scalability, and standardization. Written cases lack interactivity [7], role-playing may feel artificial or inconsistent [8], laboratory-based simulations can be expensive [3], and standardized patient encounters, while realistic, are resource-intensive and limited in scope [9]. Textbox 1 summarizes the relative strengths and limitations of the major SBL modalities, situating interactive virtual patients within this landscape. Recent simulation research has emphasized the importance of deliberate practice, debriefing, fidelity alignment, cognitive scaffolding, and mastery progression to ensure that simulation produces meaningful learning gains [10-12]. There is a pressing need for innovative, scalable approaches that combine the benefits of simulation with greater efficiency and flexibility.

Textbox 1. Relative strengths and limitations of major simulation-based learning modalities.

Written/static case studies

- Strengths: Low cost, highly standardized, and easy to scale and distribute
- Limitations: No interactivity, limited fidelity, and do not require real-time data gathering or decision-making

Role-play with peers

- Strengths: Interactive and low cost
- Limitations: Variable realism and consistency, dependent on peers' skills and engagement

Manikin-/laboratory-based simulation

- Strengths: High physical fidelity and hands-on procedural practice
- Limitations: High cost and infrastructure requirements, limited availability, and faculty-intensive

Standardized (human) patients

- Strengths: High interpersonal fidelity and realistic communication practice
- Limitations: Resource-intensive, limited scalability, and scheduling constraints

Decision-tree virtual patients:

- Strengths: Standardized, scalable, and interactive within preset branches
- Limitations: Constrained/branching dialogue and can feel scripted

Large language model–based interactive virtual patients (this study)

- Strengths: Scalable, open-ended natural-language interaction, real-time medical record and tool use, and asynchronous access
- Limitations: Potential for model errors/hallucinations, requires guardrails and authoritative source content, and dependent on the quality of case-authoring

Cost and scalability remain critical considerations in simulation. Procedural mastery experiences, while effective, can require substantial infrastructure and faculty investment. For example, mastery learning via virtual reality ultrasound simulation reached competency standards but required an estimated cost exceeding US $600 per trainee [10]. Other work in health professional education demonstrates that simulation-based mastery learning facilitates self-efficacy, authenticity of practice, and transfer to clinical performance when supported by repetition and cognitive frameworks [11].

The distinction between active and passive learning is an important confounder when comparing teaching and learning modalities. Multiple experimental studies demonstrate that

XSL•FO
RenderX

interactive, problem-solving approaches improve learning outcomes and perceived learning more than passive, lecture-based instruction [13]. Students, particularly lower-achieving learners, benefit most from active engagement formats [13]. Earlier controlled trials similarly showed that fostering learner-to-learner interaction increased engagement and observable participation even when learning outcomes were equivalent [14]. Flipped classroom models likewise produce higher achievement, greater engagement, and improved retention compared to traditional didactic instruction [15], and gamified and game-based approaches have been associated with increased motivation and engagement in health professions learners [16]. These findings underscore the need to interpret comparative simulation studies through the lens of active-learning effects in addition to any technological capabilities.

Virtual Patients and Large Language Models

Advances in digital simulation technologies and conversational interfaces have expanded opportunities for scalable, interactive learning environments. Virtual patient simulations can be implemented using a range of technologies, from structured decision-tree architectures to newer conversational systems supported by large language models (LLMs). These systems allow learners to engage in real-time dialogue with simulated patients and explore clinical reasoning processes in ways that more closely resemble clinical encounters [17-19]. LLM-driven clinical cases are increasingly being studied across a variety of educational scenarios [20], and this trial is distinguished by its head-to-head randomized comparison of interactive versus static delivery using identical case content in an interprofessional sample.

The pedagogical promise of LLM-based virtual patients must, however, be weighed against well-documented limitations of the underlying technology. LLMs generate responses probabilistically and can produce fluent but incorrect or fabricated content (“hallucinations”), exhibit pattern-driven or stereotyped responses, and vary across runs [21]. Recent work also indicates that straightforward, LLM-crafted cases are not always as engaging or as cognitively demanding as more complex, real-world cases [21]. These properties have motivated calls for explainable and verifiable AI in clinical and educational applications [22,23] and underscore the need for design safeguards, such as facilitator-authored ground-truth case content, explicit guardrails, and a priori scoring keys, when LLMs are used to deliver instructional material. We describe the specific safeguards used in this study in the “Methods” section.

Aim and Hypotheses

Within this context, this randomized study examined whether learning outcomes differed when educational cases were delivered through interactive virtual patient encounters compared with static written case studies. We hypothesized that, relative to static written cases with identical content, the interactive virtual patient format would be associated with (1) greater pretraining to posttraining knowledge gains, (2) higher case discussion quiz accuracy, and (3) higher learner engagement among health professional students learning about pain and opioid management. Because the 2 formats differ in interactivity as well as in their use of AI, the study was explicitly designed to evaluate the “interactive virtual patient format” as a whole rather than to isolate the unique contribution of AI.

Methods

Study Design

This was a single-center, parallel-group, 1:1 randomized controlled trial comparing 2 case-delivery formats following a common didactic video. The trial is reported in accordance with the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist (Checklist 1).

Participants

Participants were health professional students enrolled at a US-based academic health sciences university, representing programs in medicine, dentistry, nursing, pharmacy, and allied health. Inclusion criteria required active enrollment in clinical or preclinical coursework. An announcement about the study opportunity was made during an interprofessional course in the spring semester. Participation was voluntary.

Ethics Approval

This study was approved by the Medical University of South Carolina’s Institutional Review Board for the Protection of Human Subjects (Pro00137370). All participants reviewed an online informed consent waiver prior to enrollment. No participant-identifying information was shared with the AI model or retained by the AI provider.

Sample Size and Power

This was a preliminary (pilot-scale) trial intended to estimate effect sizes and establish feasibility; an a priori power analysis was not performed, and the achieved sample size was determined by voluntary enrollment within a single course offering. To contextualize the analyzed sample, a sensitivity analysis (computed in G*Power; Heinrich Heine University Düsseldorf) indicated that 58 completers yielded approximately 80% power at α=.05 to detect a between-within interaction in the medium-to-large range. Results are therefore interpreted as preliminary and hypothesis-generating, and the limitations of the sample size for generalizability are addressed in the “Discussion” section.

Knowledge Assessment

A brief, custom 8-item multiple-choice test was developed by the investigators to assess performance on key learning objectives of the training activity: (1) distinguishing among pain types and definitions, (2) selecting evidence-based pain management strategies, (3) assessing opioid treatment considerations and risks, and (4) identifying disparities in pain management. Question stems were developed, and correct answers were determined by consensus of the investigative team; interrater reliability among the authors was not formally assessed. The test was administered pretraining and posttraining. A full copy of the knowledge assessment is included in Multimedia Appendix 1.

XSL•FO
RenderX

Training Materials

For the didactic portion of this study, participants watched a 12-minute video lecture developed by the authors. The video covered (1) pain definitions, types, and etiologies; (2) empirically supported treatments for each pain type; (3) challenges associated with opioid prescribing identified in the empirical literature; (4) disparities in opioid prescribing; (5) prevalence of chronic pain by race and ethnicity; (6) factors influencing disparities in chronic pain prevalence; and (7) strategies for talking to patients about pain and opioids. All participants watched the same didactic video and were unable to skip any video content.

Randomization and Case Conditions

Participants were randomized 1:1 by the software platform into 1 of 2 groups. Following the didactic video, participants completed either 3 written case studies or 3 interactive case studies. Case content was identical between groups; only the delivery method and the mode of interacting with the case material were manipulated (written cases vs virtual patient cases).

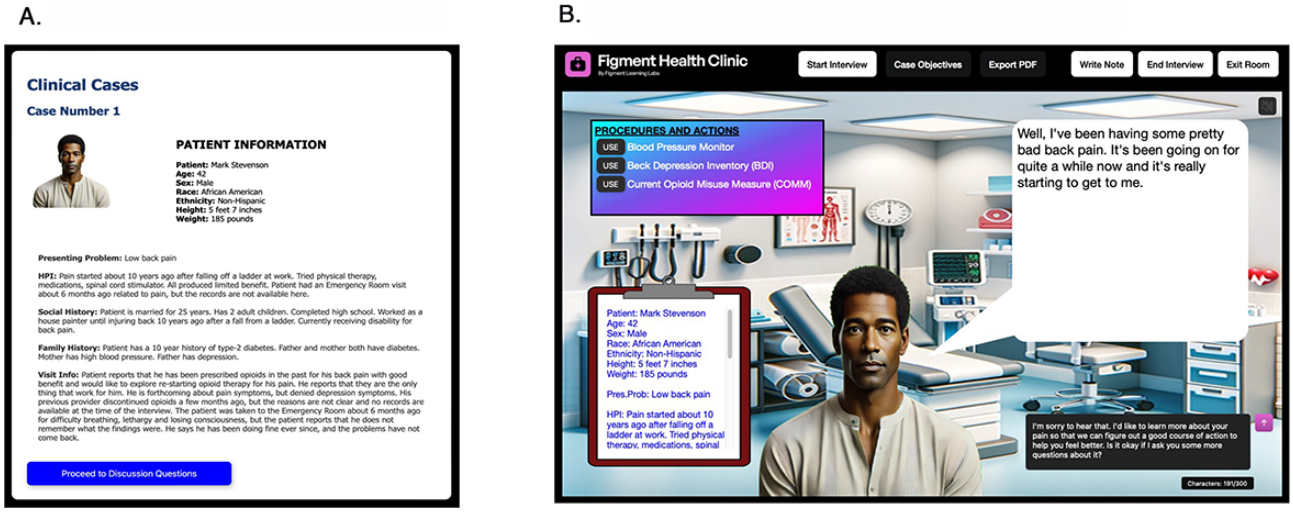
Case Discussion Quiz Questions

In both groups, after completing each of the 3 case studies, participants answered three multiple-choice discussion questions: (1) Is this patient's pain best categorized as nociceptive, neuropathic, both, or neither? (2) Is this patient's pain acute, chronic, or neither? (3) Is this patient a good candidate for opioid therapy? Correct answers were determined a priori by the investigators; the same 3 questions were presented after each case, but correct answers varied as a function of case content. Participants received 1 point for each correct answer and could not return to the case information to answer the questions.

Interactive Virtual Patients and Available Clinical Data

The Figment Learning Labs Virtual Clinic is an online system that allows course facilitators to develop AI-powered interactive cases that students can interview. Importantly, the platform is not a chat-only interface: each case includes a viewable electronic medical record and a set of facilitator-defined assessments, tools, or procedures (eg, vital signs, laboratory results, imaging, medication and allergy lists, and prior treatments) that learners can access during the encounter. Virtual patient avatars (selectable by the facilitator) are animated against a clinic backdrop; the interface includes the patient's medical record, an editable visit/progress note, and a full transcript available afterward for review (Figure 1).

Figure 1. (A) Screenshot of the static, text-based case presentation format. (B) The Figment Learning Labs AI-powered interactive case interface used in this study.



Instructors specify learning objectives for each case along with patient information: name, age, sex, race, ethnicity, height, weight, personality, mood, communication style, chief complaint, history of present illness, social history, family and medical histories, medication list, allergies, and other relevant clinical information. Facilitators also indicate the procedures, tools, or assessments that learners can order, along with the corresponding results that populate the medical record. The record updates in real time when a tool is used, and the AI avatar is instructed on how to respond to each tool. A hidden facilitator instruction field provides more comprehensive case direction that is not visible to the learner. Interviews are online, asynchronous, and interactive, and they can be conducted via voice or text.

Because the clinical "facts" of each case (history, examination findings, laboratory and imaging results, medications, allergies, and the a priori correct discussion answers) are authored by the facilitator rather than generated de novo by the model, the case content available to learners includes the supporting clinical data needed for reasoning, and the determination of correct/incorrect reasoning does not depend on AI-generated facts.

AI System Transparency, Accuracy, and Safeguards

At the beginning of a session, the facilitator-authored patient background information is sent to the OpenAI API as a system-level prompt, together with the activity context and explicit guardrail instructions that constrain the model to the role of the virtual patient and prevent it from supplying

XSL•FO
RenderX

out-of-case clinical information or breaking character. Each learner utterance or tool use is relayed to the API, and the returned response is rendered as the virtual patient's reply. The model also tracks a simulated emotional state, which drives the avatar's facial expression so that learners can respond to verbal and nonverbal cues.

Several design features were intended to support clinical fidelity and reliability. First, the medical record, tool results, and correct discussion answers were fixed and facilitator-authored, so the substantive clinical content and scoring were independent of model variability. Second, guardrail prompts restricted the model to in-case information. Third, no learner-identifying information was shared with or retained by the provider; the API is a stateless instance that does not retain information after a session. We did not implement automated, turn-by-turn verification of every conversational response against the source record; the potential for occasional model errors (eg, hallucinations) and approaches to mitigate them are addressed as a limitation in the "Discussion" section.

AI Model Information

Figment Clinic version 1.0 (Palmetto Innovative Education LLC, doing business as Figment Learning Labs) was used in conjunction with the OpenAI API and the ChatGPT (GPT-4o) model. The model parameters were as follows: temperature=0.7, top_*P*=0.9, frequency_penalty=0.8, presence_penalty=0.8, and max_tokens=500.

Case Study Information

The 3 cases were developed by the authors to represent common encounters with patients presenting with acute, chronic, neuropathic, and nociceptive pain, and each specified the patient's demographic and social context (including age, sex, race, ethnicity, and relevant social history) so that learners encountered contextualized rather than generic patients. All participants received the same 3 cases; only presentation format (written vs interactive) was randomized. Case study details are provided in Multimedia Appendix 1.

Learner Engagement Survey

After completing all training and knowledge tests, participants completed an online engagement survey developed by the authors, consisting of 17 items assessing interface usability and value, engagement and satisfaction, and applicability of the information to real-world scenarios. Items used a 5-point Likert scale (1="strongly disagree" to 5="strongly agree").

Procedures

Study enrollment opened on April 1, 2025, and closed on July 31, 2025. Participants visited the study website, reviewed the informed consent waiver online, completed a demographics form and a baseline knowledge assessment, and were then randomized in a 1:1 ratio by the online system to either written or interactive cases. All participants then viewed the 12-minute educational video on evidence-based pain and opioid management, with an emphasis on care for diverse populations. In the written condition, participants reviewed 3 text vignettes (each with a patient photograph) and answered the discussion questions. In the interactive condition, the same cases were presented as simulated telehealth consultations with AI-powered virtual patients, and participants conducted online interviews to elicit case details (rather than only reading the material in text format) before answering the same discussion questions. After the learning activity, all participants repeated the knowledge test and completed the engagement survey. All sessions were conducted remotely and asynchronously and lasted approximately 45 to 60 minutes.

Statistical Analysis

Knowledge gains were assessed via a 2×2 mixed ANOVA examining knowledge test scores (percent correct) pretraining and posttraining and between groups (written vs interactive). Case discussion accuracy and learner self-report ratings were compared between conditions using 2-tailed independent-sample *t* tests. Where Levene test indicated unequal variances, the Welch correction was applied and the adjusted degrees of freedom were reported. Internal consistency for each instrument was estimated with Cronbach α. Analyses were conducted in SPSS (version 27; IBM Corp), with the critical α set at .05. Because the 17 engagement items were analyzed individually without correction for multiple comparisons, item-level results are interpreted as supportive rather than confirmatory.

Results

Participants and Attrition

A total of 74 participants were enrolled, and 58 (78%) completed all study components and were included in the completer-based analysis. All 16 noncompleters (22%) discontinued before finishing the video training and did not begin the case study portion; reasons for discontinuation were not collected and are unknown. Baseline knowledge test data were available for 63 of the 74 enrolled participants (all 58 completers and 5 of the 16 noncompleters); the degrees of freedom for the baseline comparisons in Table 1 reflect this denominator. There were no significant differences between completers and noncompleters in sex, race, ethnicity, degree program, or age (Table 1), but because the reasons for attrition are unknown, the possibility that unmeasured differences biased the observed group differences cannot be excluded (see "Limitations" section). The overall study procedures are summarized in Figure 2.

XSL•FO
RenderX

Table . Breakdown of sample sizes (n), mean age, and mean pretest scores by participant sex, race, ethnicity, degree program, group (written cases vs interactive virtual patient cases), and status (study completers vs noncompleters).

Characteristics	Written cases[a]	Interactive cases	Significance		Completers	Noncompleters	Significance		Pretest score, mean (SD)[b]	Significance[c]		Total
			Statistic	*P* value			Statistic	*P* value		Statistic	*P* value	
Age (y), mean (SD)	27.16 (8.34)	27.44 (8.10)	0.15 (72)[d]	.88	26.07 (6.74)	31.75 (11.19)	1.94 (72)[d]	.07	27.30 (8.17)	—[e]		27.30 (8.17)
Sex, n			2.06 (1)[f]	.15			0.08 (1)[f]	.77		0.96 (61)[d]	.34	
Male	8	13			16	5			5.11 (1.02)			21
Female	30	23			42	11			5.42 (1.22)			53
Race, n			1.95 (3)[f]	.58			2.46 (3)[f]	.48		1.80 (3, 59)[g]	.16	
White	26	26			42	10			5.44 (1.16)			52
Asian	7	7			9	5			5.50 (.97)			14
Black	3	3			5	1			4.67 (1.21)			6
Multiracial	2	0			2	0			4.00 (1.41)			2
Ethnicity			1.19 (2)[f]	.55			2.47 (2)[f]	.29		0.75 (2, 60)[g]	.48	
Hispanic	1	3			4	0			5.00 (.82)			4
Non-Hispanic	35	31			50	16			5.40 (1.13)			66
Prefer not to say	2	2			4	0			4.75 (1.89)			4
Group			—	—			0.47 (1)[f]	.49		1.61 (61)[d]	.11	
Written cases	—	—			31	7			5.10 (1.14)			38
Interactive cases	—	—			27	9			5.56 (1.16)			36
Status			0.47 (1)[f]	.49			—	—		0.93 (61)[d]	.35	
Completers	31	27			—	—			5.29 (1.17)			58
Noncompleters	7	9			—	—			5.80 (1.01)			16
Degree			3.84 (5)[f]	.57			8.86 (5)	.12		1.08 (5, 57)[g]	.38	
Medicine	9	10			13	6			5.08 (1.04)			19

XSL•FO
RenderX

Characteristics	Written cases[a]	Interactive cases	Significance		Completers	Noncompleters	Significance		Pretest score, mean (SD)[b]	Significance[c]		Total
			Statistic	*P* value			Statistic	*P* value		Statistic	*P* value	
Dental medicine	1	0			1	0			6.00 (SD not estimable, n=1)			1
Pharmacy	2	6			7	1			4.57 (1.27)			8
Allied health professions	22	17			34	5			5.56 (1.12)			39
Graduate studies	3	2			2	3			5.40 (1.34)			5
Nursing	1	1			1	1			5.00 (SD not estimable, n=1)			2

[a]All 74 enrolled participants are represented (58 completers and 16 noncompleters). Group columns show the allocation of all enrolled participants, while the completers and noncompleters columns partition that same sample.

[b]Baseline (pretest) knowledge scores were available for 63 of the 74 enrolled participants (all 58 completers and 5 of the 16 noncompleters). The *df* for the baseline comparisons reflect this denominator.

[c]Comparisons of categorical variables were conducted using chi-square tests of independence. Comparisons of age and baseline knowledge scores by sex, group, and completion status were performed using independent-sample *t* tests. Comparisons of baseline knowledge scores across race, ethnicity, and degree program were conducted using one-way ANOVA.

[d]Value is reported as *t* test (*df*).

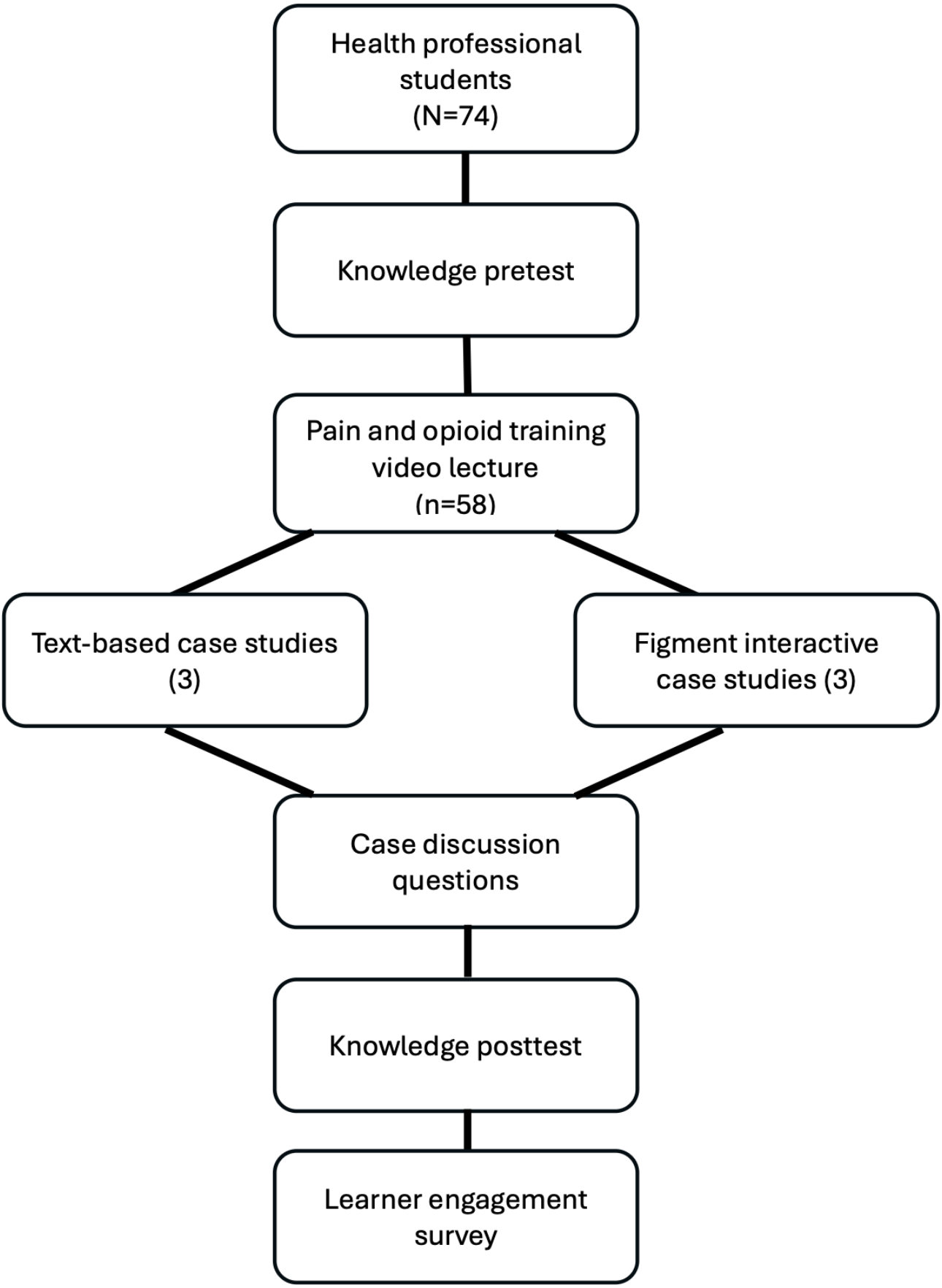
[e]N/A: not available.

[f]Value is reported as chi-square (*df*).

[g]Value is reported as *F* test (*df*).

XSL•FO
RenderX

Figure 2. Study procedure diagram.



Among the 58 completers, 16 (28%) were male and 42 (72%) were female; mean age was 26.07 (SD 6.74) years. Self-reported race was White (n=42, 72%), Asian (n=9, 16%), Black (n=5, 9%), and multiracial (n=2, 3%); ethnicity was Hispanic (n=4,

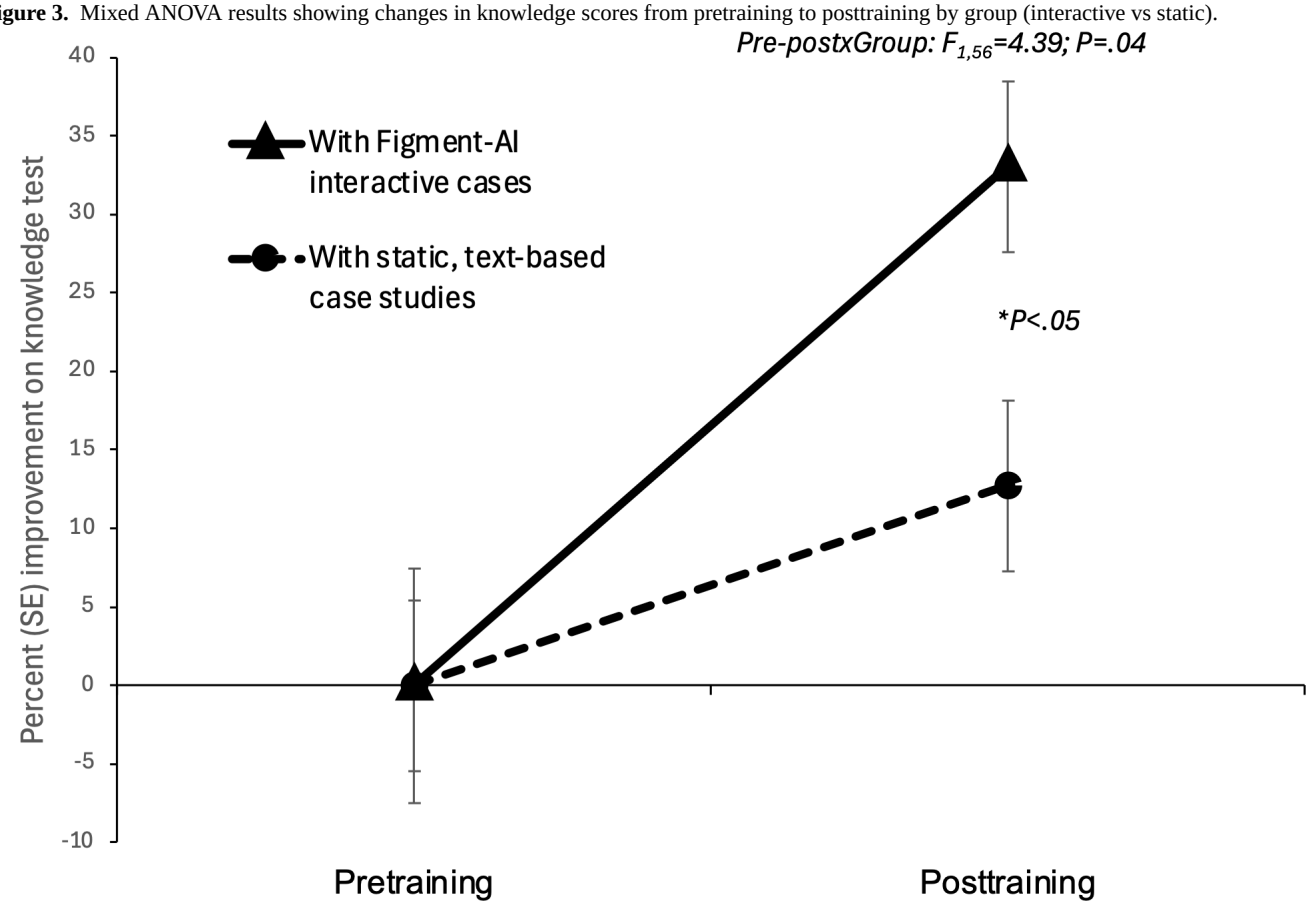
XSL•FO
RenderX

7%), non-Hispanic (n=50, 86%), and prefer-not-to-say (n=4, 7%). Degree programs were allied health professions (n=34, 59%), medicine (n=13, 22%), pharmacy (n=7, 12%), graduate studies (n=2, 3%), dental medicine (n=1, 2%), and nursing (n=1, 2%). Of the 74 enrolled participants, 38 (51%) were randomized to the written-case group and 36 (49%) to the interactive-case group; among the 58 completers, 31 (53%) were in the written-case group and 27 (47%) in the interactive-case group. Sample sizes, mean age, and mean pretest knowledge scores by category are shown in Table 1.

Knowledge Gains

Internal consistency of the 8-item knowledge test was modest (Cronbach α=0.56), as expected for a brief, custom instrument, and this value should be considered when interpreting the knowledge results. A 2×2 mixed ANOVA revealed a significant time (pretraining to posttraining)×group (written vs interactive) interaction ($F_{1,56}$=4.39; P=.04; partial η^2=0.07), indicating greater knowledge gains in the interactive condition; the posttest between-group difference was also significant (P=.03). Figure 3 shows the percent improvement from pretraining to posttraining by group.

Figure 3. Mixed ANOVA results showing changes in knowledge scores from pretraining to posttraining by group (interactive vs static).



Case Discussion Quiz Accuracy

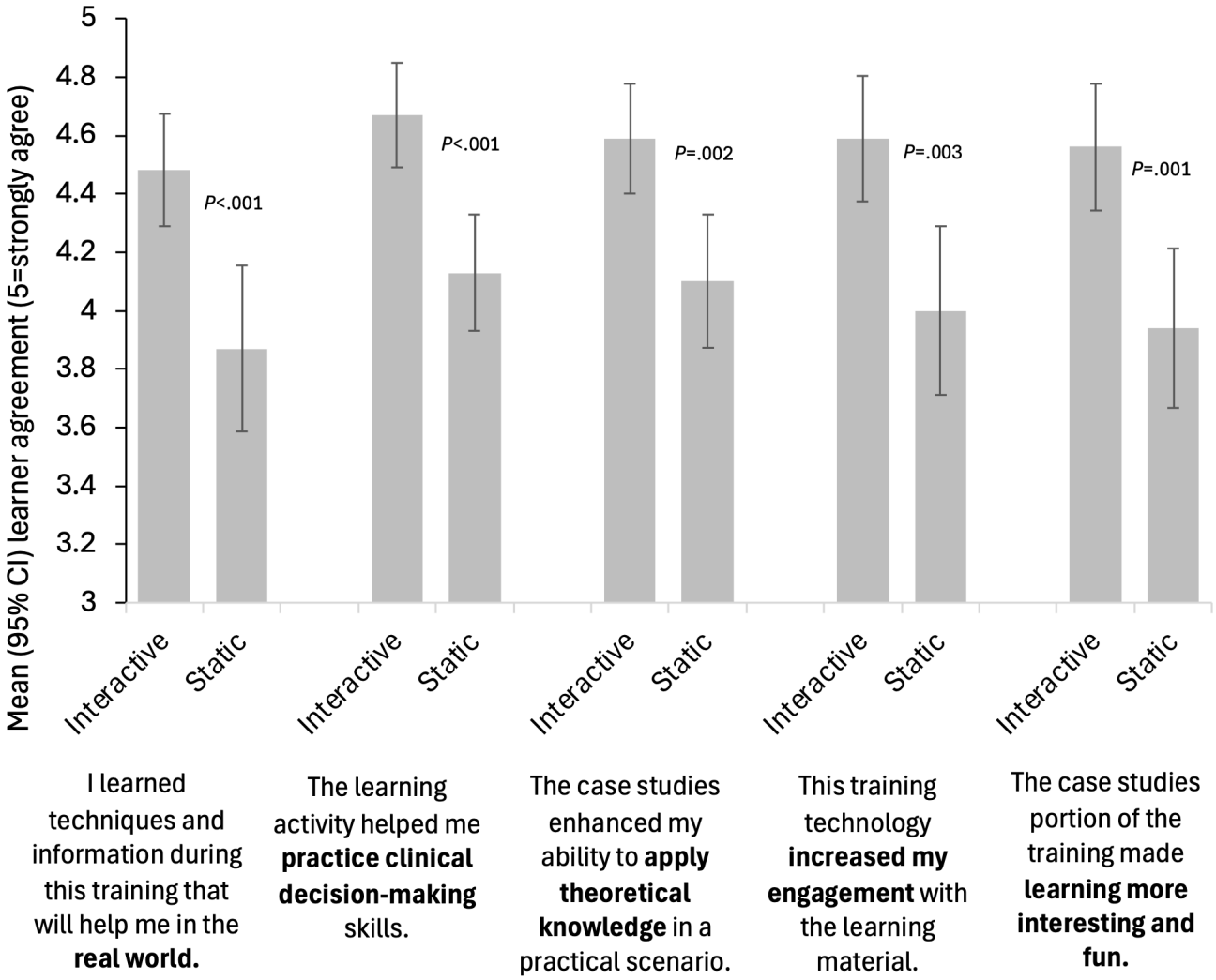
Internal consistency of the case discussion measure across the 9 dichotomously scored items (3 discussion questions after each of 3 cases) was low (Cronbach α=0.31). This is expected and does not necessarily indicate measurement error: although the same prompts followed each case, the correct responses differed across cases, which represented distinct clinical scenarios. The items were not intended to function as parallel indicators of a single latent construct; rather, they sampled performance across heterogeneous cases. Accordingly, the case discussion score is interpreted as an aggregate performance measure across cases rather than as a homogeneous psychometric scale. Participants in the interactive condition demonstrated significantly higher accuracy (mean 0.79, SD 0.15) than those in the written condition (mean 0.64, SD 0.11; t_{56}=4.48; P<.001), corresponding to an approximately 25% relative advantage for the interactive format.

Learner Engagement

Internal consistency of the 17-item engagement questionnaire was high (Cronbach α=0.95). Agreement scores were significantly higher in the interactive group than in the written group on 13 of the 17 items (P<.05). Figure 4 shows group-level means (95% CI) on five key indicator items: (1) I learned techniques and information during this training that will help me in the real world; (2) the learning activity helped me practice clinical decision-making skills; (3) the case studies enhanced my ability to apply theoretical knowledge in a practical scenario; (4) this training technology increased my engagement with the learning material; and (5) the case-studies portion made learning more interesting and fun. Full item-level results are shown in Table 2. Because these were 17 uncorrected, item-level comparisons, the pattern is interpreted as consistently supportive of greater engagement in the interactive condition rather than as a set of independent confirmatory findings.

XSL•FO
RenderX

Figure 4. Learner mean performance (95% CI) on key engagement survey items (interactive vs static).



XSL•FO
RenderX

Table . Mean Likert rating score[a] for learner engagement items among participants in the interactive virtual patient case group compared with those in the written case group[b].

Learner engagement survey item and group	Number, n	Mean (SD)	*t* test (*df*)	*P* value
The online training interface was easy to use.			1.75 (55.92)	.09
Interactive cases	27	4.56 (0.51)		
Written cases	31	4.29 (0.78)		
The instructions for completing the online training were clear and understandable.			0.77 (55.92)	.45
Interactive cases	27	4.41 (0.57)		
Written cases	31	4.29 (0.59)		
The online training platform is technologically sophisticated.			3.81 (47.69)	<.001
Interactive cases	27	4.48 (0.51)		
Written cases	31	3.74 (0.93)		
The online training platform is visually appealing.			3.46 (56)	.001
Interactive cases	27	4.52 (0.51)		
Written cases	31	3.74 (1.06)		
The case studies portion of the learning activity met the learning objectives.			2.33 (56)	.02
Interactive cases	27	4.52 (0.58)		
Written cases	31	4.16 (0.58)		
The case studies felt like “real-world” scenarios.			2.00 (55.36)	.052
Interactive cases	27	4.63 (0.56)		
Written cases	31	4.32 (0.60)		
The case studies portion of the training made learning more interesting and fun.			3.42 (56)	.001
Interactive cases	27	4.56 (0.58)		
Written cases	31	3.94 (0.77)		
I felt that I learned something from the case studies.			3.89 (56)	<.001
Interactive cases	27	4.59 (0.57)		
Written cases	31	4.00 (0.78)		
I learned techniques and information during this training that will help me in the real world.			3.39 (56)	.001
Interactive cases	27	4.48 (0.51)		
Written cases	31	3.87 (0.81)		
The learning activity helped me practice clinical decision-making skills.			3.88 (56)	<.001
Interactive cases	27	4.67 (0.48)		

XSL•FO
RenderX

Learner engagement survey item and group	Number, n	Mean (SD)	*t* test (*df*)	*P* value
Written cases	31	4.13 (0.56)		
The case studies enhanced my ability to apply theoretical knowledge in a practical scenario.			3.88 (55.90)	<.001
Interactive cases	27	4.59 (0.50)		
Written cases	31	4.10 (0.65)		
The case studies felt realistic.			1.74 (56)	.09
Interactive cases	27	4.41 (0.80)		
Written cases	31	4.03 (0.84)		
This training technology increased my engagement with the learning material.			3.15 (56)	.002
Interactive cases	27	4.59 (0.57)		
Written cases	31	4.00 (0.82)		
I am motivated to use activities like online case studies to enhance my learning experience.			3.30 (56)	.002
Interactive cases	27	4.56 (0.51)		
Written cases	31	3.97 (0.80)		
This training activity made learning more enjoyable.			3.47 (56)	.001
Interactive cases	27	4.52 (0.51)		
Written cases	31	3.90 (0.79)		
I am satisfied with this learning activity.			3.53 (56)	.001
Interactive cases	27	4.52 (0.51)		
Written cases	31	3.97 (0.66)		
I would recommend this learning activity to my peers.			4.56 (56)	<.001
Interactive cases	27	4.56 (0.51)		
Written cases	31	3.84 (0.69)		

[a]Items used a 5-point Likert scale (1="strongly disagree" to 5="strongly agree").

[b]All comparisons were conducted using 2-tailed independent-sample *t* tests (interactive cases, n=27; written cases, n=31). Where Levene test indicated unequal variances, the Welch correction was applied, and the adjusted *df* are reported.

Discussion

Principal Findings

In this randomized controlled trial, students assigned to the interactive case condition demonstrated higher short-term knowledge gains, higher case discussion accuracy, and higher engagement than students assigned to the written case condition with identical content. These findings support the value of an interactive virtual patient format for case-based learning in health professions education. They are, however, preliminary and primarily hypothesis-generating, and, as detailed below, the design does not permit attribution of the observed benefits to AI specifically.

Interpretation: An Interactive-Format Effect, Not an Isolated AI Effect

The strongest interpretation supported by this design is that an interactive, AI-enabled virtual patient format outperformed a static written format delivering the same content. The trial cannot separate the contribution of AI from the contributions of conversational interactivity, learner agency, adaptive questioning, or additional time-on-task because these features are bundled within the interactive condition. We therefore frame

XSL•FO
RenderX

the results as support for interactive virtual patient learning rather than as evidence of a distinct, AI-specific advantage. At a minimum, integrating an LLM did not detract from the benefits of interactive learning; quantifying the incremental value of AI over non-AI interactive formats will require trials that include a non-AI interactive comparison arm.

Comparison With Prior Work

Our findings align with and extend prior research on simulation-based and active learning, adding randomized evidence that an interactive virtual patient platform can yield both measurable learning advantages and operational scalability [5,6,10-12]. The use of AI to deliver clinical cases has been examined in a range of recent studies [20], and some work suggests that simple LLM-generated cases may be less engaging than complex real-world cases [21]; our results, using facilitator-authored cases delivered through an interactive interface, are consistent with the broader active-learning literature [13-15] showing the benefits of interactivity over passive formats. Although this study addressed pain management, the instructional mechanism is not necessarily content-specific, and the observed benefits may extend to other domains that rely on clinical reasoning, communication, and applied decision-making. The extent to which these effects generalize across content areas, learner levels, and curricular structures remains an empirical question.

Clinical Reasoning, Data Access, and Demographic Context

Interactive virtual patients can train skills beyond communication; in the platform used here, learners could access a structured medical record and order facilitator-defined assessments (eg, laboratory results, medication and allergy lists, and prior treatments), so the encounters were designed to *unfold* components of case information (as opposed to openly presenting them as part of the vignette, as in the text-based case format) in addition to facilitating conversation alone. Nonetheless, this study did not separately measure information-gathering behavior, and future work should evaluate how access to and use of clinical data within the simulation relate to reasoning outcomes. Second, the realism of any virtual patient depends on the demographic and socioeconomic context authored into the case. The 3 cases here specified patient demographics and social context, but 3 US-based cases cannot represent the full range of geographic, socioeconomic, and cultural circumstances that shape pain management; readers should not generalize these specific encounters to all settings, and future case libraries should deliberately span diverse contexts and be evaluated for bias.

AI Reliability and the Need for Verifiable Content

LLMs can generate fluent but incorrect content, exhibit stereotyped patterns, and vary across runs [21]. In this study, the substantive clinical content and the scoring of reasoning were fixed and facilitator-authored, and guardrail prompts constrained the model to in-case information, which limits, but does not eliminate, the risk that a conversational reply could be inaccurate. We did not perform automated, turn-by-turn verification of model responses against the source record. Educational deployments of LLM-based virtual patients should therefore incorporate safeguards such as authoritative source content, explicit guardrails, human review, and, where feasible, explainable or verification-oriented architectures [22,23], and they should also monitor for inaccurate or biased output.

Limitations

Several limitations should be acknowledged. First, this was a completer-based analysis: 16 of 74 enrolled participants (22%) discontinued before the case study portion, and because reasons for attrition are unknown, unmeasured differences between completers and noncompleters could have biased the observed group differences and may limit conclusions about feasibility, acceptability, and generalizability. Second, time-on-task was not measured; students in the interactive condition may have spent more time with the material, and this duration effect cannot be disentangled from the modality itself. Third, voluntary participation may introduce self-selection bias toward learners more favorably disposed toward technology-assisted learning. Fourth, the assessment instruments were brief, with modest (knowledge test; Cronbach α=0.56) to low (aggregate discussion score; Cronbach α=0.31) internal consistency, limiting the granularity of measured learning. Fifth, the study was conducted at a single center with a modest sample (n=58) and was powered only to detect medium-to-large effects; small effects could have been missed. Sixth, only immediate postintervention outcomes were assessed; data on knowledge and skill retention at 6 or 12 months were not collected and are an important gap. Finally, the effects of AI cannot be separated from the broader benefits of interactivity, and 2 authors (JJB and DH) are founders of the company that commercializes the platform under study; although this conflict is disclosed and the substantive content and scoring were author-fixed, independent replication by investigators without a commercial interest is essential.

Future Directions

Future research should include a non-AI interactive comparison arm to isolate any incremental AI effect; measure time-on-task and in-simulation data-gathering behavior; assess retention at 6 and 12 months and transfer to authentic clinical performance; use validated, multifaceted assessments; and recruit larger, multi-institutional, and demographically diverse samples through curricular integration. Work on explainable and verifiable AI [22,23] may help ensure that conversational responses remain clinically accurate and interpretable as these tools scale.

Conclusions

This preliminary randomized controlled trial suggests that an interactive virtual patient format may be a portable, scalable, and feasible way to realize the benefits of interactive, case-based learning. Compared with static written cases of identical content, students in the interactive condition showed greater short-term knowledge gains, higher case discussion accuracy, and stronger engagement. These results are hypothesis-generating and reflect an interactive-format effect that cannot be attributed to AI specifically; they should be interpreted in light of completer-based analysis with unexplained attrition, voluntary participation, brief assessments, the absence of time-on-task and retention data, and the authors' commercial interest. Larger,

XSL•FO
RenderX

multi-institutional, longitudinal, and independently conducted trials—ideally including non-AI interactive comparison conditions—are needed to confirm durability and generalizability and to determine the unique contribution, if any, of AI to learning outcomes.

Acknowledgments

Generative AI (Claude Opus 5) was used to proofread for grammar and spelling and to enhance the readability of this paper. The authors remain fully responsible for the accuracy, originality, and integrity of all content in the paper.

Funding

This work was supported by the National Institute on Drug Abuse (NIDA) of the National Institutes of Health under Small Business Technology Transfer (STTR) award number R41DA059281. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable written request.

Conflicts of Interest

JJB and DH are employed by the Medical University of South Carolina (MUSC) and are founders of Palmetto Innovative Education LLC, doing business as Figment Learning Labs—the company formed to commercialize the Figment Learning Labs software developed at MUSC with National Institutes of Health funding. The remaining authors declare no conflicts of interest.

Multimedia Appendix 1
Knowledge test items and case study information.
[DOCX File, 22 KB - nursing_v9i1e98700_app1.docx]

Checklist 1
CONSORT-EHEALTH (V 1.6.1) checklist.
[PDF File, 449 KB - nursing_v9i1e98700_app2.pdf]

References

1. Weimer M, Chou R, Sevarino K. Pain core curriculum. Providers Clinical Support System for Medications for Opioid Use Disorder (PCSS-MOUD). URL: https://pcssnow.org/education-training/pain-core-curriculum/ [accessed 2026-09-15]
2. Schmidt HG, Mamede S. How to improve the teaching of clinical reasoning: a narrative review and a proposal. Med Educ 2015 Oct;49(10):961-973. [doi: 10.1111/medu.12775] [Medline: 26383068]
3. Zendejas B, Wang AT, Brydges R, Hamstra SJ, Cook DA. Cost: the missing outcome in simulation-based medical education research: a systematic review. Surgery 2013 Feb;153(2):160-176. [doi: 10.1016/j.surg.2012.06.025] [Medline: 22884087]
4. Mafumo JL, Maputle MS. Limited resources in clinical facilities: experiences of nursing students during placements. Curationis 2025 May 26;48(1):e1-e8. [doi: 10.4102/curationis.v48i1.2732] [Medline: 40459092]
5. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011 Sep 7;306(9):978-988. [doi: 10.1001/jama.2011.1234] [Medline: 21900138]
6. McGaghie WC, Issenberg SB, Petrusa ER, Scalese RJ. A critical review of simulation-based medical education research: 2003-2009. Med Educ 2010 Jan;44(1):50-63. [doi: 10.1111/j.1365-2923.2009.03547.x] [Medline: 20078756]
7. Thistlethwaite JE, Davies D, Ekeocha S, et al. The effectiveness of case-based learning in health professional education. A BEME systematic review: BEME Guide No. 23. Med Teach 2012;34(6):e421-e444. [doi: 10.3109/0142159X.2012.680939] [Medline: 22578051]
8. Kim J, Park JH, Shin S. Effectiveness of simulation-based nursing education depending on fidelity: a meta-analysis. BMC Med Educ 2016 May 23;16(20):152. [doi: 10.1186/s12909-016-0672-7] [Medline: 27215280]
9. Oh PJ, Jeon KD, Koh MS. The effects of simulation-based learning using standardized patients in nursing students: a meta-analysis. Nurse Educ Today 2015 May;35(5):e6-e15. [doi: 10.1016/j.nedt.2015.01.019] [Medline: 25680831]
10. Post JH, Teslak KE, Tolsgaard MG, Rasmussen S, Friis ML. Cost of simulation-based mastery learning for abdominal ultrasound. BMC Med Educ 2023 Dec 5;23(1):921. [doi: 10.1186/s12909-023-04919-5] [Medline: 38053134]
11. Tan K, Seah B, Wong LF, Lee CCS, Goh HS, Liaw SY. Simulation-based mastery learning to facilitate transition to nursing practice. Nurse Educ 2022;47(6):336-341. [doi: 10.1097/NNE.0000000000001224] [Medline: 35667017]

XSL•FO
RenderX

12. Dawidziuk A, Miller G, Malawana J. Visualisation approaches in technology-enhanced medical simulation learning: current evidence and future directions. Adv Exp Med Biol 2023;1421:175-190. [doi: 10.1007/978-3-031-30379-1_8] [Medline: 37524988]
13. Boedeker P, Schlingmann T, Kailin J, et al. Active versus passive learning in large-group sessions in medical school: a randomized cross-over trial investigating effects on learning and the feeling of learning. Med Sci Educ 2025 Feb;35(1):459-467. [doi: 10.1007/s40670-024-02219-1] [Medline: 40144125]
14. Haidet P, Morgan RO, O'Malley K, Moran BJ, Richards BF. A controlled trial of active versus passive learning strategies in a large group setting. Adv Health Sci Educ Theory Pract 2004;9(1):15-27. [doi: 10.1023/B:AHSE.0000012213.62043.45] [Medline: 14739758]
15. Patnaik S, Behera MR, Patnaik L. Evaluating the effectiveness of the flipped classroom model in undergraduate medical education: a quasi-experimental study. Cureus 2025 Jun;17(6):e86312. [doi: 10.7759/cureus.86312] [Medline: 40688939]
16. van Gaalen AEJ, Brouwer J, Schönrock-Adema J, Bouwkamp-Timmer T, Jaarsma ADC, Georgiadis JR. Gamification of health professions education: a systematic review. Adv Health Sci Educ Theory Pract 2021 May;26(2):683-711. [doi: 10.1007/s10459-020-10000-3] [Medline: 33128662]
17. Yu H, Zhou J, Li L, et al. Simulated patient systems powered by large language model–based AI agents offer potential for transforming medical education. Commun Med (Lond) 2025 Dec 19;6(1):27. [doi: 10.1038/s43856-025-01283-x] [Medline: 41420084]
18. Borg A, Schiött J, Ivegren W, et al. AI-enhanced social robotic versus computer-based virtual patients for clinical reasoning training in medical education: observational crossover cohort study. J Med Internet Res 2025 Nov 27;27:e82541. [doi: 10.2196/82541] [Medline: 41309100]
19. Zeng J, Qi W, Shen S, et al. Embracing the future of medical education with large language model–based virtual patients: scoping review. J Med Internet Res 2025 Nov 13;27:e79091. [doi: 10.2196/79091] [Medline: 41232097]
20. Jiang J, Ye MZ, Kwok TTO, Wong JYH. GenAI-supported virtual patients in health care education: systematic review. J Med Internet Res 2026 May 7;28:e82756. [doi: 10.2196/82756] [Medline: 42098926]
21. Karampinis E, Bozi Tzetzi DA, Pappa G, et al. Use of a large language model as a dermatology case narrator: exploring the dynamics of a chatbot as an educational tool in dermatology. JMIR Dermatol 2025 Sep 16;8:e72058. [doi: 10.2196/72058] [Medline: 40957071]
22. Binsar F. Integrating chain of thought and explainable AI in BERT-based deep learning for interpretable medical diagnosis. Procedia Comput Sci 2025;269:389-399. [doi: 10.1016/j.procs.2025.08.291]
23. Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The Virtual Operative Assistant: an explainable artificial intelligence tool for simulation-based training in surgery and medicine. PLoS One 2020;15(2):e0229596. [doi: 10.1371/journal.pone.0229596] [Medline: 32106247]

Abbreviations

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
LLM: large language model
SBL: simulation-based learning

Edited by E Borycki; submitted 17.Apr.2026; peer-reviewed by E Karampinis, F Binsar, S Ghazi; revised version received 07.Aug.2026; accepted 07.Sep.2026; published 23.Sep.2026.

Please cite as:
Borckardt JJ, Henninger-Borckardt D, Barth K, Annan-Coultas D, Langdale L, Kascak K, Pelic C, de Arellano MA, Brady K Knowledge Acquisition, Case Discussion, and Engagement in Health Professional Students Using Interactive Virtual Patient Cases Versus Written Case Studies: Randomized Controlled Trial
JMIR Nursing 2026;9:e98700
URL: https://nursing.jmir.org/2026/1/e98700
doi:10.2196/98700

© Jeffrey J Borckardt, Debra Henninger-Borckardt, Kelly Barth, Dusti Annan-Coultas, Lisa Langdale, Kimberly Kascak, Christopher Pelic, Michael A. de Arellano, Kathleen Brady. Originally published in JMIR Nursing (https://nursing.jmir.org), 23.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Train-the-Trainer Evaluation of Technology-Enhanced Learning in Preservice Midwifery Education in Nigeria: Qualitative Descriptive Study

Isabella Garti[1], PhD; Nadyen Shikpup[2], MSc; Ophelia Tonto[3], BSc; Michelle Gray[4], PhD; Halima Musa Abdul[5], PhD; Awwal Mohammed Ladan[6], PhD

[1]Faculty of Health, Charles Darwin University, 7 Ellengowan Dr, Casuaurina, Northern Territory, Australia

[2]Department of Nursing Sciences, University of Jos, Jos, Plateau State, Nigeria

[3]Ghana Health Service, Accra, Greater Accra, Ghana

[4]School of Nursing and Midwifery, University of Newcastle, Newcastle, New South Wales, Australia

[5]College of Nursing Sciences, Ahmadu Bello University, Kaduna, Kaduna State, Nigeria

[6]Department of Nursing Sciences, Bayero University, Kano, Kano State, Nigeria

Corresponding Author:
Isabella Garti, PhD
Faculty of Health, Charles Darwin University, 7 Ellengowan Dr, Casuaurina, Northern Territory, Australia

Abstract

Background: Midwifery education in Nigeria is undergoing a transformation in pedagogical approach. Existing traditional teaching materials are being supplemented by digital technology-enhanced learning resources; however, educators' limited pedagogical capacity may hinder their adoption. Thus, train-the-trainer (TTT) workshops were implemented to build capacity among midwifery educators to prepare them for the implementation of a digital learning platform across 20 institutions in Nigeria.

Objective: This study aimed to describe the development and qualitative evaluation of a TTT program to support the implementation of a technology-enhanced learning platform within preservice midwifery education in Nigeria. We aimed to evaluate participants' perceptions of the training and their preparedness to implement their learning.

Methods: The workshop was developed using the ADDIE (Analysis, Design, Development, Implementation, and Evaluation) instructional design model. A 3-day TTT program was delivered in 2 cohorts in July 2025 to a purposive sample of 80 institution-nominated midwifery educators and information and communication technology officers from 20 pilot institutions. A qualitative descriptive study was conducted to explore participants' perceptions of the training and their readiness to implement and support the technology-enhanced learning platform within their institutions. Data were analyzed using qualitative content analysis.

Results: Pre-workshop responses demonstrated varying levels of prior exposure to and use of digital technologies, as well as differing expectations of the training. Four themes emerged from the post-workshop evaluation. Participants perceived improvements in their understanding of pedagogical and technology-enhanced learning approaches; reported greater readiness to support implementation and knowledge transfer within their institutions; highlighted infrastructure and technical support barriers to implementation; and emphasized the need for ongoing capacity building, monitoring, and support to promote long-term sustainability.

Conclusions: Rather than evaluating educational outcomes, this study provides early evidence of educator readiness, implementation considerations, and sustainability requirements for technology-enhanced learning in preservice midwifery education in low-resourced settings such as Nigeria. A multifaceted approach combining capacity building, ongoing support, technical requirements, and institutional commitment to continuous improvement is needed. Future phases of this implementation research will evaluate educational, implementation, and institutional outcomes associated with the digital learning platform.

(JMIR Nursing 2026;9:e93596) doi:10.2196/93596

KEYWORDS

midwifery; midwifery educator; experiences; train the trainer; digital learning; digital technology

XSL•FO
RenderX

Introduction

Nigeria has a maternal mortality rate of 993 per 100,000 live births, more than 3 times the global average of 211 per 100,000 live births [1]. The country continues to experience poor maternal and newborn health outcomes, partly attributable to deficiencies in the quality of midwifery education. Nigeria is unlikely to make substantial progress toward achieving the Sustainable Development Goals (SDGs) related to maternal and newborn health [2] without strengthening preservice midwifery education, particularly in northern states disproportionately affected by high maternal mortality, health workforce shortages, insecurity, and ongoing unrest associated with conflict and instability [3]. Digital learning approaches can support more flexible and scalable educational delivery, strengthen educators' capacity, and supplement existing preservice training with interactive, clinically relevant content [4-6]. Therefore, increased investment in midwifery education in Nigeria is necessary to ensure that international standards for quality education are met and to lower the extremely high maternal and newborn mortality rates [7,8].

Internationally, strengthening midwifery education is recognized as a critical strategy for reducing preventable maternal and newborn mortality and achieving the SDGs [9]. The International Confederation of Midwives (ICM) calls for midwifery education to meet international competency standards while remaining responsive to local population and health system needs [10]. Similarly, global initiatives such as the World Health Organization's Ending Preventable Maternal Mortality agenda and the Midwifery Accelerator Framework emphasize the importance of investing in a competent, well-supported midwifery workforce as a foundation for improving maternal and newborn outcomes [11,12]. High-quality midwifery education is therefore essential for equipping practitioners with the knowledge, clinical competence, and respectful care approaches required to deliver safe and effective maternity care across diverse settings [8].

Midwifery education in Nigeria is delivered primarily through schools of nursing and midwifery, colleges of health technology, and universities, and is regulated by the Nursing and Midwifery Council of Nigeria (NMCN). The Nigerian midwifery curriculum, recently revised, does not fully align with the ICM's Global Standards, which outline the minimum requirements for preparing competent midwives [13]. There are, however, significant challenges within preservice midwifery education in Nigeria. Midwifery training institutions face increasing pressure from workforce shortages, limited educational resources, and inconsistent access to clinical learning opportunities [8,14]. The expertise of midwifery educators in Nigeria is variable, with many tutors having strong clinical backgrounds but limited formal preparation in pedagogy, as well as limited access to contemporary teaching technologies, simulation resources, and ongoing educator development opportunities necessary to support competency-based and student-centered learning [7,14]. Thus, the delivery of education is predominantly teacher-centric, with heavy reliance on didactic instruction and rote learning [15]. Consequently, there is a growing need for innovative, student-centered educational approaches that support self-directed learning, critical thinking, clinical reasoning, and greater engagement with evidence-based midwifery practice.

Previous studies suggest that technology-enhanced learning can help address gaps in midwifery education by enriching teaching and learning and supporting more flexible, interactive, and student-centered educational approaches [6,16,17]. When integrated with traditional classroom teaching, digital technologies can accommodate diverse learning styles and promote greater learner engagement compared with predominantly didactic methods [16,18]. However, effective integration of digital technologies into midwifery curricula requires adequate educator preparation, technical support, and ongoing capacity strengthening to support implementation, adaptation, and sustainability [17,19]. As key drivers of digital transformation in education, midwifery educators require sufficient pedagogical and digital competence to effectively facilitate and support student learning in increasingly technology-enabled environments [20].

To support the effective integration and sustainability of technology-enhanced learning approaches in midwifery education, there is a need for structured capacity-building initiatives that strengthen the pedagogical and digital capabilities of educators and institutional support staff [19]. One approach increasingly used in health care education and workforce development is the train-the-trainer (TTT) model. The TTT method provides participants with the knowledge and expertise to become skilled trainers, enabling them to transfer their acquired knowledge and skills to others within their professional group. TTT models involve experts preparing a nucleus group of professionals to disseminate knowledge, skills, and evidence-based practices to their peers within their own organizations [21]. This approach supports scalable capacity strengthening and the development of local expertise, especially in limited-resource settings, promotes peer-supported learning, and may enhance the implementation of new practices within routine health care and educational settings [22]. Although TTT programs can facilitate the delivery of digital midwifery education at scale, robust and context-specific models are needed to ensure quality, sustainability, and local ownership.

Although technology-enhanced learning is increasingly being integrated into health professions education globally, the literature reports more on student learning outcomes than on the preparedness of educators and the institutional systems required to support sustainable implementation [15,23]. Limited evidence exists on how midwifery educators and institutional support staff in low- and middle-income countries experience and adapt to digital transformation initiatives, particularly within large-scale preservice education reforms [19].

Limited research has explored the design, implementation, and evaluation of training programs specifically developed to prepare midwifery educators for technology-enhanced learning approaches. Existing literature has largely focused on student experiences and digital learning outcomes, with comparatively less attention given to building the pedagogical and digital capabilities of educators responsible for integrating these approaches into preservice midwifery education. This gap is

XSL•FO
RenderX

particularly evident in low- and middle-income countries, where educators may have varying levels of exposure to digital technologies, instructional design, and online facilitation strategies.

This study provides novel evidence on a needs-informed TTT approach designed specifically to prepare midwifery educators to support technology-based learning in preservice midwifery education. Its contribution lies in addressing the limited evidence on how educators in resource-constrained settings can be equipped to integrate digital platforms and digital teaching approaches into existing midwifery programs. The findings advance knowledge by identifying practical strategies to strengthen midwifery educators' digital competencies and to support the effective and sustained implementation of technology-based learning within educational institutions. The study involved collecting baseline information on the perspectives of Nigerian midwifery educators and information and communication technology (ICT) officers regarding their participation in the initial TTT workshop.

Methods

Context of the Study

The NMCN digital learning platform was developed to deliver, manage, and monitor educational content for approximately 2000 preservice midwifery students across 20 pilot institutions in 11 states in Nigeria. In this pilot, 2 of the researchers, both international midwifery experts, codeveloped the digital modules. Each module features podcasts, simulations with first-person point-of-view videos, animations, and interactive technology-enhanced learning content focused on 3 maternal health conditions: postpartum hemorrhage and the E-MOTIVE bundle, multiple micronutrient supplementation for maternal anemia, and maternal sepsis. The 3 digital learning modules were developed to address major contributors to preventable maternal morbidity and mortality in Nigeria and other low-resource settings. They were designed to strengthen students' exposure to contemporary, evidence-informed maternal health content through interactive and technology-enhanced learning approaches that could supplement existing classroom and clinical education. Midwifery students across the pilot sites will access the platform over a 12-month implementation period, during which iterative refinement of the platform and educational content will be undertaken alongside ongoing evaluation of the technology-enhanced learning intervention's feasibility, acceptability, and usability.

TTT Program Development and Delivery

This study is part of a broader implementation research project introducing a digital learning platform into preservice midwifery education in Nigeria. The ADDIE (Analysis, Design, Development, Implementation, and Evaluation) instructional design model guided the development of the educational products and learning resources for the TTT program. This model is used as a structured framework for instructional design [24]. The model includes analysis of learner needs and educational objectives, design and development of instructional content and delivery approaches, implementation of the training, and evaluation to inform ongoing refinement and improvement. Although presented in phases, the ADDIE model is iterative and nonlinear in practice. As part of the evaluation phase, a qualitative descriptive study was conducted to explore the perspectives and initial experiences of midwifery educators and institutional ICT officers who participated in the TTT workshop.

The ADDIE Framework

Analysis

The analysis phase identified gaps in preservice midwifery education in Nigeria, including reliance on predominantly teacher-centered approaches that did not adequately support student learning outcomes or meet the objectives of the NMCN. A baseline assessment conducted between September and December 2024 identified shortages of midwifery educators, limitations in digital learning infrastructure and skills laboratories, and competency gaps among educators, preceptors, and final-year students in key maternal health areas, including postpartum hemorrhage management. These findings highlighted the need for strengthened educator capacity, improved learning resources, and innovative teaching approaches, including refresher training and technology-enhanced learning strategies.

Design

The design phase focused on planning the implementation of a technology-enhanced learning platform within existing preservice midwifery programs. Learning needs assessments identified limited digital literacy, insufficient experience with learner-centered teaching approaches, and varying levels of familiarity with digital technologies among educators. The TTT approach was selected to support scalable implementation and sustainability across pilot institutions. Educational content and module topics were predetermined within the broader project and developed collaboratively by Nigerian midwifery academics, project partners, and key stakeholders, including the NMCN, to ensure contextual relevance and alignment with curriculum standards. Two 3-day workshop sessions were designed for separate participant cohorts. Training objectives, session content, instructional activities, logistical requirements, and facilitation strategies were developed during this phase. The workshops incorporated practical activities, case studies, small-group discussions, and adult learning principles to promote active participation and peer learning. Nigerian midwifery educators led the delivery of the training, supported by the project team and technology partners, to enhance contextual appropriateness and communication.

Development

During the development phase, the original program developers created technology-enhanced learning resources and instructional materials for the TTT workshops, as they had the most in-depth understanding of the program. Due to contextual factors, including language and local communication styles, it was not feasible for the program developers to deliver the training directly. Therefore, it was planned that experienced Nigerian midwifery faculty with in-depth knowledge of the program would facilitate the training, with ongoing support and guidance from the original developers. A combination of theoretical and practical learning was used to equip trainers with the competencies needed to deliver high-quality cascade

XSL•FO
RenderX

training. Workshop materials were designed to model learner-centered and interactive teaching approaches while supporting participants' engagement with adult learning principles and instructional design concepts. A training handbook was developed to provide guidance on curriculum alignment, lesson planning, and integration of digital technologies into teaching and assessment. Additional workshop activities, including group-based exercises and interactive digital tools such as Mentimeter, were incorporated to support collaborative learning and participant engagement. Technology partners also developed platform-specific instructional resources covering platform architecture, navigation, login procedures, backend functionality, and user management. Effective adaptation strategies were introduced during the initial training and supported by planned supplementary training to reinforce learning and build confidence. Session schedules, durations, and training logistics were finalized during this phase (Multimedia Appendix 1).

Implementation

The TTT workshops were delivered in Abuja between July 14, 2025, and July 19, 2025, in 2 cohorts through intensive 3-day training sessions. The workshops aimed to prepare participants to support the implementation of the digital learning platform within their institutions and to cascade training to other midwifery educators. Training combined hands-on workshops facilitated by the technology partner with in-person and virtual sessions delivered by Nigerian educators and supported by the original program developers and the project team. Educational strategies included presentations, facilitated discussions, group work, demonstrations, and practical activities. Participants received an overview of the Moodle platform architecture and engaged in guided and independent navigation exercises exploring platform functionality, user roles, and educational content. Training explored challenging areas in greater depth, such as the integration of digital learning, and incorporated dedicated opportunities for questions and shared problem-solving. This initial training session focused on foundational digital literacy and platform engagement, with future training planned to support participants in independently creating and uploading digital learning resources and integrating the platform into existing teaching approaches.

Evaluation

This evaluation was not designed to formally assess feasibility, but rather to explore participants' perceptions and experiences of the TTT workshops to inform ongoing refinement of the training program and broader implementation processes. The evaluation used anonymous qualitative surveys administered before and after the training workshops. Pre-training questions explored participants' prior experiences, expectations, and perceptions of digital technologies, while post-training questions examined the perceived usefulness and acceptability of the training, as well as participants' perceived readiness to support implementation within their institutions. These insights were intended to inform educator readiness, engagement, and acceptability, while guiding future refinement, sustainability, and cascading of the intervention.

Research Design

A qualitative descriptive design [25] was used, as it enables researchers to remain close to the data while providing a comprehensive summary of participants' perspectives and experiences without aiming to generate theory or explore deeper, lived, or cultural meanings [26]. Qualitative description is particularly suited to examining who, what, and the context of events or experiences and is useful for exploring poorly understood phenomena through the accounts of participants [25]. This approach was selected because little is currently known about the perspectives and experiences of midwifery educators in Nigeria regarding the introduction of technology-enhanced learning resources and the expectations associated with adopting and disseminating new digital approaches within preservice midwifery education. Consequently, a qualitative descriptive methodology was considered appropriate for exploring participants' perceptions, experiences, and responses to this educational change introduced by the NMCN. Findings from the evaluation were used to identify areas requiring additional support, refinement, and future capacity-building activities.

Participants and Sampling

Participants were purposively selected by the NMCN from participating institutions based on their teaching responsibilities in courses linked to the digital learning modules, as course allocations are predetermined for the academic year. Institutional ICT officers were also included due to their roles in supporting digital infrastructure and platform implementation within the schools. Eligible participants included seasoned midwifery educators and ICT officers attending the TTT workshops who were not on leave, were involved in the relevant courses, and were willing to participate in the qualitative evaluation conducted before and after the training. A total of 80 participants attended 2 separate 3-day workshops. Although participants were nominated by their institutions to attend the training, participation in the qualitative evaluation component was voluntary, and participants were informed that their decision to participate or decline would not affect their involvement in the training program.

Data Collection

Qualitative data were collected through anonymous, written, open-ended responses before and after the workshops. This qualitative method can generate meaningful and rich insights, even when individual responses are brief, and minimizes participant discomfort [27]. This approach was selected based on participant preference, as some attendees were uncomfortable with audio-recorded interviews or group discussions, and it enabled participants to share their perspectives in their own words while minimizing burden within the intensive training schedule. Surveys were completed individually at the venue. Participants could remain at their allocated seats or choose another suitable location within the premises and were given approximately 30 minutes to record their responses. They wrote their responses by hand on the paper-based surveys and could request extra sheets as needed.

XSL•FO
RenderX

The pre-training survey comprised 5 questions that assessed participants' prior experience, expectations, and concerns about the training (Multimedia Appendix 2). The post-training survey included 9 open-ended questions that evaluated participant satisfaction, key learnings, intended actions, and areas requiring further training to inform future program planning (Multimedia Appendix 3).

Despite the anonymous, written format, social desirability bias may still have influenced responses, with participants potentially feeling inclined to provide favorable feedback due to their institutional nomination to attend the training and the broader implementation context.

Data Analysis

The research team followed the systematic steps of qualitative content analysis as described by Erlingsson and Brysiewicz [28], which include familiarization with the data, identifying meaning units, condensing those units, coding, categorizing, and formulating themes. Content analysis was chosen because it provides a systematic and transparent method for analyzing open-ended qualitative data across multiple participants. Data were analyzed manually without the use of qualitative software. Manual coding was selected to support collaborative analysis across the international research team and ensure accessibility for all researchers involved in the project. Two researchers independently reviewed and coded the responses before meeting to compare interpretations and discuss coding differences. Any discrepancies were resolved through discussion and consensus within the broader research team. Meaningful data segments relevant to the study objectives, including perceptions of the training, perceived learning gains, challenges, and anticipated application of skills, were identified and coded. The research team met regularly via Zoom to discuss coding decisions, resolve discrepancies through consensus, merge overlapping codes, and refine categories. Related codes were grouped into broader descriptive categories, which were subsequently developed into higher-level themes. Throughout the analytic process, categories and themes were repeatedly reviewed against the original written responses to ensure they accurately reflected participants' perspectives and experiences. Given the qualitative descriptive nature of the study and the use of brief written responses, the analysis aimed to achieve sufficient informational richness relevant to the study objectives rather than formal theoretical data saturation associated with interpretive methodologies.

Trustworthiness

To enhance the trustworthiness of the findings, several strategies were used throughout the analysis [29]. Credibility was enhanced through independent coding by members of the research team, followed by consensus discussions to resolve differences in interpretation and refine the coding framework. Dependability was supported through consistent data collection and analysis procedures, and by documenting coding decisions as the analysis progressed. To ensure confirmability, regular team meetings were held in which interpretations were critically examined and grounded in participants' responses rather than in individual assumptions. Transferability was facilitated by providing detailed information about the participants, training, and the study context. Researcher reflexivity was embedded throughout the analytic process and involved ongoing consideration of the researchers' roles, assumptions, and involvement in the program, as well as their potential influence on data collection, analysis, and interpretation.

Ethical Considerations

Ethical approval for this study was obtained from the Human Research Ethics Committee of Charles Darwin University, Australia (ethics number H25071), and the institutional Research Ethics Committee at the Liverpool School of Tropical Medicine (ethics number 24 - 074). The study was conducted in accordance with internationally accepted ethical principles for research involving human participants. All participants were provided with adequate information about the study objectives, procedures, potential risks, and benefits. Informed consent was obtained from all participants before data collection. Consent was sought both verbally and in writing, and participation was not compulsory. Confidentiality and anonymity were maintained throughout the study, and participation was entirely voluntary, with the right to withdraw at any stage without penalty. Prior to data collection, participants were informed verbally and in writing that completion of the evaluation activities was optional and would not affect their participation in the training or their professional roles. None of the participants objected to participating. Participants were also informed that they could decline to answer any question or withdraw at any time without consequence.

Results

In total, 80 participants took part in the study, comprising 60 female participants and 20 male participants. Most participants (n=58, 72.5%) were midwifery educators, alongside 21 ICT officers and 1 clinical instructor (Table 1).

XSL•FO
RenderX

Table . Participant characteristics.

Variable and category	Values, n (%)
Gender	
Male	20 (25)
Female	60 (75)
Participant role	
Midwifery educator	58 (72.5)
ICT[a] officer	21 (26.3)
Clinical instructor	1 (1.2)
Participating states	
Yobe	8 (10)
Katsina	8 (10)
Kaduna	8 (10)
Kano	8 (10)
Niger	8 (10)
Plateau	8 (10)
Borno	8 (10)
Ebonyi	8 (10)
Bauchi	8 (10)
Zamfara	4 (5)
Akwa-Ibom	4 (5)
Total, N	80 (100)

[a]ICT: information and communication technology.

Pre-Training Responses

Pre-training responses revealed considerable variation in participants' prior exposure to digital tools, current technology use, and expectations for the training. Although some participants reported prior digital training, these experiences were often described as basic, outdated, or insufficient, whereas others relied mainly on self-directed learning and trial-and-error. This finding suggests limited access to structured and ongoing digital capacity development across participating institutions.

Participants' current use of digital tools was generally limited to basic applications, such as PowerPoint, WhatsApp, and online videos, with only a few having prior experience with platforms like Moodle. Despite these limitations, participants expressed strong interest in developing practical digital teaching skills, particularly in navigating learning management systems, creating digital content, integrating multimedia resources, improving online student engagement, and enhancing digital assessment practices. The findings indicate a strong demand for hands-on, practice-oriented training to strengthen educators' readiness for technology-enhanced teaching.

Post-Training Responses

Post-Training Perceptions and Experiences

The post-training responses highlighted 4 themes that articulated participants' perceptions and experiences of the TTT session—theme 1: developing pedagogical and digital teaching competencies; theme 2: self-perceived confidence and readiness to support implementation; theme 3: perceived barriers and logistical challenges; and theme 4: sustaining technology-enhanced learning.

Theme 1: Developing Pedagogical and Digital Teaching Competencies

Following the training workshop, participants described a greater understanding of both educational pedagogy and digital concepts. They described acquiring foundational knowledge and practical skills to effectively use digital tools and navigate the teaching platform. This highlights that the training was instrumental in helping educators align digital resources with the existing curriculum, thereby strengthening the relevance and integration of technology within learning activities. The structured and progressive nature of the training appeared to help participants connect pedagogy, curriculum outcomes, and technology-enhanced learning in more practical ways.

> *The training was progressive, with each part building upon the previous one. It was a step-by-step process: you must understand the learning objectives in the curriculum because they go hand in hand.* [Cohort 2, Respondent 23]

The lesson-planning component was particularly valued because it strengthened participants' understanding of learning outcomes, learner characteristics, and Bloom's taxonomy, while improving

XSL•FO
RenderX

their ability to design effective lessons and organize courses sequentially.

> *I have also learned how to make a lesson plan, which will bring a breakthrough in my teaching activities and help me achieve the outcomes I want for my students. It [Lesson plan] also helps us, as educators, to create a sequential model that makes our work flexible and enhances student learning.* [Cohort 1, Respondent 1]

Participants further noted that the knowledge gained from lesson planning and course organization would improve student learning outcomes.

> *This will improve learning outcomes for my students..., understanding how a course is designed and organized online.* [Cohort 2, Respondent 22]

Participants described gaining a broader understanding of educational principles, including different types of learners, teaching approaches, learning activities, and Bloom's taxonomy. They perceived this knowledge as enhancing their understanding of how teaching and learning can be effectively supported through the learning management system.

> *I also gained knowledge on the different types of students, lecturers, and educators, the various kinds of learning activities, and the six types of Bloom's taxonomy. All these have helped me understand teaching and learning through the LMS.* [Cohort 1, Respondent 4]

Participants reported gaining practical skills in using the learning management system, including creating assessments, delivering lectures, uploading educational content, developing multimedia learning materials, and monitoring student engagement and participation.

> *I have gained a great deal from learning how to log in to the computer and navigate it. I also learned how to use the computer in teaching and incorporating videos. In addition, I gained knowledge of the learning management system, how to navigate the portal, and an overview of creating content.* [Cohort 1, Respondent 15]

Participants described the content and resources as appropriate, useful, and accessible, supporting both teaching and student engagement. Participants appreciated the use of Nigerian voiceovers and actors, which increased the realism and contextual relevance of the learning experience.

> *The parts of the training I found most useful and impactful were those that made teaching and learning more realistic, efficient, and effective by utilizing podcasts, videos, pictures, animations, drag-and-drop features, and scenarios. This is because these will make teaching and learning easier and more interesting for both the trainer and the learner.* [Cohort 2, Respondent 23]

These findings suggest that participants required support not only in the technical use of digital platforms but also in the pedagogical foundations necessary to effectively integrate technology-enhanced learning into existing teaching practices. The overlap between educational and digital competencies highlights that successful digital transformation requires more than technological infrastructure; it also requires educator preparedness and confidence in applying learner-centered teaching approaches.

Theme 2: Self-Perceived Confidence and Readiness to Support Implementation

Following the training, participants described greater perceived confidence and readiness to support the implementation of technology-enhanced learning within their institutions. Participants attributed this readiness to the practical, hands-on nature of the training, which enabled them to actively engage with the platform, explore its functionality, and apply newly acquired knowledge in a supportive learning environment. There were 2 subthemes: valuing practical, hands-on learning and commitment to training others.

Subtheme 1: Valuing Practical, Hands-On Learning

Participants consistently highlighted the workshop's practical, interactive design as a key facilitator of learning. Opportunities to navigate the platform, engage with activities, and apply concepts in real time were perceived as particularly valuable for building familiarity with technology-enhanced learning approaches.

> *I found the training to be very engaging and participatory, with everyone actively involved typing and clicking on their computers, logging in and out.* [Cohort 2, Respondent 3, post-training]

Another participant added:

> *The practical aspect of the training was the most useful because it made learning easier for me. It was better to do it myself, and I felt more confident by the second day. It has motivated me to use a laptop to teach. Henceforth, I will continue to try my hand at creating blended lectures.* [Cohort 2, Respondent 4, post-training]

The workbook and workshops enhanced participants' confidence and motivation, equipping educators with resources to support students' use of digital tools during implementation.

> *I was previously hesitant about using technology in training and teaching, but this program gave me the confidence to utilize digital tools effectively.* [Cohort 2, Respondent 5, post-training]

However, some participants felt that additional training sessions would help consolidate learning and build confidence over time. They noted that the breadth of content introduced during the workshop required additional time and practice to fully understand and apply it. They collectively emphasized that ongoing training and support would be essential to sustain their engagement.

> *Organise training 2 - 3 times per year...especially creating PowerPoints.* [Cohort 1, Respondent 20, post-training]

XSL•FO
RenderX

> *I suggest that more days should be allocated to this type of training next time, enabling us to fully grasp everything.* [Cohort 2, Respondent 20, post-training]

These findings indicate the importance of embedding continuous professional development, technical support, and implementation coaching within technology-enhanced learning initiatives to facilitate long-term adoption and integration into teaching practice.

Subtheme 2: Commitment to Training Others

Participants expressed a strong willingness to disseminate the knowledge they acquired through this process through step-down training, viewing this as essential to ensuring the platform's success and sustainability. Many described intentions to organize training sessions for colleagues, support the implementation of the digital learning platform, and begin developing or uploading their own educational content.

> *I plan to train other educators [on this training], involve students, and ensure that they can access the platform. I will be very excited to start uploading my own modules to improve the learning and teaching process.* [Cohort 1, Respondent 5]

Participants also outlined strategies for preparing to cascade the training, including reviewing workshop materials, revisiting the training workbook and notes, and developing action plans tailored to their institutional contexts.

> *This is how I plan to apply what I learned in my own teaching practice: first, to become very familiar with the process of logging in and using the platform so that I can effectively transfer the knowledge and skills to my colleagues and students. I will then ensure that my students are referred to the modules. I will also create content to further support their learning.* [Cohort 2, Respondent 3]

Another participant explained,

> *Before I train my colleagues, I will review my notes and develop an action plan.* [Cohort 2, Respondent 18]

These findings suggest that participants viewed themselves not only as recipients of training but also as potential change agents within their institutions. Their willingness to train colleagues, develop educational content, and plan implementation activities reflects their perceived readiness to support the broader adoption of technology-enhanced learning.

Theme 3: Perceived Barriers and Logistical Challenges

Participants identified infrastructure and ongoing support as critical determinants of successful implementation. Common concerns included unreliable internet connectivity, unstable power supply, limited sensitization, and access to timely technical support. The provision of free internet access was frequently proposed, underscoring the need to address systemic and infrastructural barriers to the sustainable integration of technology-enhanced learning.

> *All necessary facilities, such as internet access and power supply, should be ensured, as they are very important.* [Cohort 2, Respondent 17]

Participants identified technical support and access to adequate resources as essential for sustaining digital learning and the effective implementation of the program. Participants also emphasized that resource availability, including access to laptops and embedded training materials, could enhance the consistency and ease of training delivery.

> *Yes, my suggestion is that laptops should be provided to master trainers, with all the topics embedded into the laptop.* [Cohort 2, Respondent 20]

Participants highlighted the importance of collaboration between educators and ICT officers, emphasizing the need for shared training opportunities and equal access to the platform. This finding suggests that successful implementation of technology-enhanced learning requires a team-based approach, with clearly defined roles and shared responsibility to support sustainability and reduce reliance on individual staff members.

> *Both midwife educators and Information Technology officers should be trained at the same level so that the ICT officers will not be overwhelmed with loads of responsibilities.* [Cohort 1, Respondent 5]

Theme 4: Sustaining Technology-Enhanced Learning

Participants highlighted the importance of continuity in future training activities, expressing a preference for the same facilitators who introduced the platform to lead subsequent training sessions, thereby providing continuity in learning, reinforcing key concepts, and supporting a deeper understanding of both the platform and its educational applications.

> *There is a need for more training on content creation and uploads. The NMCN should organise such training two to three times per year.* [Cohort 2, Respondent 29]

> *Future training should include the same educators who introduced the platform to ensure continuity and a deeper understanding.* [Cohort 1, Respondent 8]

Participants identified several strategies to support successful implementation and ongoing engagement with the platform. These included regular system maintenance and updates to enhance usability, development of user-friendly navigation tutorials such as YouTube videos, routine check-ins to monitor progress and address challenges, and the establishment of clear expectations for appropriate behavior in both classroom and online learning environments.

> *a) Regular maintenance and updates that can enhance user experience, b) tutorials on navigating the platform, which can be in the form of YouTube videos, c) implementation of regular check-in strategies, and d) clear guidelines for acceptable behaviour within the classroom and online.* [Cohort 1, Respondent 22]

Participants generally valued the training but suggested that online sessions could be made more engaging through greater visual interaction with facilitators. This finding highlights the importance of facilitator presence in online learning

XSL•FO
RenderX

environments and reflects the practical challenges of delivering technology-enhanced learning in settings where limited internet bandwidth may restrict the use of video and other interactive features.

> *I suggest that the online lectures be made more engaging and less boring, perhaps by allowing us to see the presenters or facilitators.* [Cohort 2, Respondent 26]

Together, these findings suggest that sustained implementation requires ongoing capacity strengthening, resource investment, and responsive support systems, rather than one-off training interventions.

Discussion

Principal Findings

Participants in this study perceived that the TTT program enhanced their understanding of technology-enhanced learning and increased their confidence in supporting its implementation in preservice midwifery education. They also expressed a willingness to champion digital learning in their institutions. However, participants consistently identified inadequate digital infrastructure, limited technical support, and the need for ongoing training as important considerations for sustaining its implementation. The study's insights contribute to a growing evidence base by identifying contextual factors that educators perceive as important for integrating digital learning into preservice midwifery education in resource-constrained settings.

The findings suggest that the success of TTT programs for technology-enhanced learning extends beyond knowledge acquisition and requires attention to implementation readiness, contextual barriers, and strategies that support long-term sustainability, consistent with evidence synthesized in a recent scoping review [19]. Such evidence is scarce in midwifery education, particularly in the context of nationwide initiatives in countries with high maternal mortality, where strengthening the midwifery workforce is critical to improving maternal and newborn outcomes [9].

Prior studies have demonstrated that TTT initiatives can improve knowledge, confidence, and teaching competencies [21,30,31]. Although this study did not measure these outcomes quantitatively, participants reported increased confidence, readiness to train others, and motivation to apply newly acquired skills, all of which are required for successful implementation and sustainability [30,32]. In contexts where educator shortages and limited resources constrain professional development opportunities, such gains may be critical for supporting the successful implementation and scale-up of technology-enhanced learning. These qualitative findings extend the existing evidence by providing a deeper understanding of the processes through which TTT programs build implementation capacity and foster sustainable educational change.

Gaps in pedagogical and digital competencies can directly affect the quality, implementation, and sustainability of technology-enhanced learning, as well as training outcomes at different levels. Although participants were enthusiastic about the TTT program and eager to apply their new knowledge, many lacked a foundational understanding of teaching and learning principles, reflecting a reliance on traditional didactic approaches. These findings should be understood within the broader Nigerian educational environment, where educator shortages, limited professional development opportunities, and inadequate digital skills training are well-documented [7,15]. Within midwifery education, large student cohorts and high educator-to-student ratios further reinforce didactic teaching practices [15]. In this context, the program addressed an important capacity gap by strengthening educators' pedagogical and digital competencies and equipping them to support teaching and learning despite significant workforce and resource constraints.

In this study, resistance to educational innovation was not the issue. Participants demonstrated openness to change and a willingness to experiment with new teaching approaches. While participants were enthusiastic about technology-enhanced learning, they frequently highlighted unreliable internet connectivity and electricity supply as key barriers to implementation. These infrastructure challenges were viewed as important considerations for successful implementation, as reported in other settings [6,16,33,34], and are also consistent with broader reports describing digital infrastructure limitations across Nigeria [35,36]. While planned national investments in fiber-optic infrastructure may improve future connectivity, these findings suggest that digital education initiatives must be accompanied by investments in reliable internet connectivity, digital infrastructure, devices, and ongoing technical support that enable educators and students to fully engage with these resources to achieve their intended impact [34].

Sustaining digital learning initiatives requires supportive policy, leadership commitment, and ongoing training [34]. In low- and middle-income countries, where resources are constrained and priorities shift, understanding sustainability is critical [37]. Given Nigeria's size and security challenges, technology-enhanced learning strategies may offer a viable and scalable solution. Notably, there is a strong political commitment to advancing midwifery in Nigeria, as reflected in the Nigeria Strategic Directions for Nursing and Midwifery (NSDNM) 2025 to 2030, which acknowledges ongoing gaps in training, deployment, and retention of skilled midwives [38]. Situating participant-identified barriers within this national reform agenda highlights that, while targeted training to build educator capacity is essential, sustainable improvement will require continued investment in educational infrastructure, workforce development, and supportive policy frameworks.

The findings have several practical implications. First, digital capacity-building programs should incorporate both technological skills and foundational pedagogical training to support meaningful educational transformation. Second, implementation strategies should include ongoing mentorship, technical support, and opportunities for peer learning to reinforce skills gained during training [34]. This also includes diffusing expertise within the system through a strong community of practice, which was set up via WhatsApp group chats in this case. Third, policymakers and educational leaders should recognize that investments in educator capacity alone are unlikely to achieve sustainable change without concurrent

XSL•FO
RenderX

improvements in digital infrastructure, internet access, and institutional support systems [16].

Overall, the findings suggest that the TTT approach is a feasible and practical strategy for implementing technology-enhanced learning at scale. Future studies should examine the longer-term impact of the TTT model on behavioral outcomes, including educator practice, institutional adoption of technology-enhanced learning, and student learning outcomes. Evaluation of the planned training cascade will be particularly important in determining whether knowledge and skills are successfully transferred to additional educators and sustained over time. Further studies should also explore factors that influence implementation across different institutional and geographical contexts and identify strategies that support long-term sustainability of digital education initiatives in resource-constrained settings.

Limitations

This study has several limitations. First, a quantitative assessment of participants' pre-training and post-training knowledge and skills was not conducted, as the training was intentionally designed to foster a supportive, nonthreatening learning environment. Second, although a paper-based open-ended survey was deemed appropriate, it may have limited opportunities for deeper exploration of individual experiences and perspectives. Third, participants were nominated by their institutions to attend the training, which may have influenced perceptions of voluntariness, despite the evaluation being optional and anonymous. As the findings were based on self-reported written feedback collected within a funded implementation project, socially desirable responses cannot be excluded. Moreover, we did not objectively assess knowledge, skills, competence, or the extent of change attributable to the program. Therefore, the findings reflect perceived readiness, usefulness, and acceptability rather than measurable improvements in teaching practice or digital competence. In addition, the short duration of the training may have limited participants' confidence to cascade the training, and the absence of a structured buddying approach may have reduced opportunities for peer support and ongoing capacity building. As the broader implementation project is ongoing, future phases of the study will evaluate longer-term educational, institutional, and implementation outcomes associated with the technology-enhanced learning platform.

Conclusions

This study evaluated participants' perceptions of a needs-informed TTT program and their preparedness to implement a technology-enhanced learning platform in preservice midwifery education in Nigeria. It provides novel evidence in an underresearched area by showing how targeted educator training can strengthen digital competencies and support the implementation and sustained use of technology-based learning. Rather than evaluating educational outcomes or changes in practice, this study provides early evidence of educators' readiness to adopt student-centered innovative learning resources. Participants are motivated to share their learning and are enthusiastic about adopting the new resources on a digital learning platform. Nevertheless, our findings suggest that national implementation may require a structured strategy to equip educators with pedagogical principles to advance the widespread application across Nigeria. The findings also indicate that expanding technology-enhanced learning in preservice midwifery education in Nigeria is likely to require sustained institutional and government commitment to strengthening the infrastructure required to support digital learning platforms.

Acknowledgments

No generative AI tools were used in the preparation of this manuscript.

Funding

Funding for this study is provided by the Gates Foundation. The funder has no role in the study design, data collection, analysis, interpretation, or report writing.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to the ongoing nature of the implementation and the absence of participant consent for data sharing. All data relevant to the study's findings are included in the manuscript and the supplementary materials.

Authors' Contributions

IG conceptualized the study, mapped out the study design, supervised data collection and analysis, and spearheaded the drafting and review of the manuscript. NS contributed to the development of the instrument, recruitment of participants, and coordination of data collection and analysis. OT took part in data analysis, writing the original draft, the review and editing of the manuscript, and organized logistics needed during the study. MG contributed to the conceptualization and design of the study, drafting, review, and editing of the manuscript. HMA and AML contributed to the development of training resources, ensured contextual adaptation, and reviewed and edited the manuscript. Each author has read the completed work, given their approval for submission, and committed to taking personal responsibility for their contributions and for addressing any unanswered concerns.

XSL•FO
RenderX

Conflicts of Interest

Although this study formed part of a funded implementation project, the authors had limited involvement in the direct delivery of the train-the-trainer workshops. The workshops were primarily facilitated by Nigerian midwifery educators and technology partners. The authors did not hold supervisory, consultancy, or institutional authority roles over the participants.

Multimedia Appendix 1
Summary of the 3-day train-the-trainer workshop.
[DOCX File, 29 KB - nursing_v9i1e93596_app1.docx]

Multimedia Appendix 2
Pre-training open-ended questions.
[DOCX File, 21 KB - nursing_v9i1e93596_app2.docx]

Multimedia Appendix 3
Post-training open-ended questions.
[DOCX File, 22 KB - nursing_v9i1e93596_app3.docx]

References

1. Trends in maternal mortality 2000 to 2023: estimates by WHO, UNICEF, UNFPA, World Bank Group and UNDESA/Population Division. : World Health Organization; 2025 URL: https://iris.who.int/server/api/core/bitstreams/29f43a3d-2228-489c-b1e7-b2f28e9101ce/content [accessed 2026-01-18]
2. Alkema L, Chou D, Hogan D, et al. Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN Maternal Mortality Estimation Inter-Agency Group. Lancet 2016 Jan 30;387(10017):462-474. [doi: 10.1016/S0140-6736(15)00838-7] [Medline: 26584737]
3. Ige WB, Ngcobo WB. A model to strengthen the quality of midwifery education: a grounded theory approach. IME 2024 Nov 24;3(4):473-487. [doi: 10.3390/ime3040036]
4. Sadat A, Vasilica C. The impact of digital learning in the context of nursing pedagogy. Nurse Educ Pract 2026 Feb;91:104682. [doi: 10.1016/j.nepr.2025.104682] [Medline: 41448039]
5. Arbour MW, Nypaver CF, Wika JC. Innovative uses of technology in online midwifery education. J Midwifery Womens Health 2015;60(3):278-282. [doi: 10.1111/jmwh.12291] [Medline: 25952711]
6. Gray M, Vogel K, Bernloehr A, Downer T, Welfare M, Bauer NH. Digital teaching in midwifery education since 2021: a literature review of trends, technological advancements and challenges. Nurse Educ Pract 2026 Jul;94:104900. [doi: 10.1016/j.nepr.2026.104900] [Medline: 42335723]
7. Shikuku DN, Mohammed H, Mwanzia L, et al. Evaluation of the feasibility of a midwifery educator continuous professional development (CPD) programme in Kenya and Nigeria: a mixed methods study. BMC Med Educ 2024 May 14;24(1):534. [doi: 10.1186/s12909-024-05524-w] [Medline: 38745243]
8. Ige WB, Ngcobo WB. Exploring the conceptualization of strengthening the quality of midwifery education: a constructivist grounded theory study. Int J Afr Nurs Sci 2025;23:100881. [doi: 10.1016/j.ijans.2025.100881]
9. Strengthening quality midwifery education for universal health coverage 2030: framework for action. : World Health Organization; 2019 URL: https://iris.who.int/server/api/core/bitstreams/33d86f14-1958-429a-ae89-3555a416c108/content [accessed 2026-05-27]
10. ICM global standards for midwifery education (revised 2021). : International Confederation of Midwives (ICM); 2021 URL: https://www.internationalmidwives.org/a-moment-for-midwives-2021/pdf/global-standards-for-midwifery-education_2021_en-1.pdf [accessed 2025-04-15]
11. The midwifery accelerator: expanding health care for women and newborns. : United Nations Population Fund (UNFPA); 2025 URL: https://www.unfpa.org/sites/default/files/pub-pdf/UNFPA-Midwifery_15_Web.pdf [accessed 2026-05-27]
12. Ending preventable maternal mortality (EPMM). World Health Organization. 2021. URL: https://www.who.int/initiatives/ending-preventable-maternal-mortality [accessed 2026-05-27]
13. Adegoke AA, Mani S, Abubakar A, van den Broek N. Capacity building of skilled birth attendants: a review of pre-service education curricula. Midwifery 2013 Jul;29(7):e64-e72. [doi: 10.1016/j.midw.2012.08.009] [Medline: 23199532]
14. Iwu E, Elnakib S, Abdullahi H, et al. Rapid assessment of pre-service midwifery education in conflict settings: findings from a cross-sectional study in Nigeria and Somalia. Hum Resour Health 2025 Feb 3;23(1):6. [doi: 10.1186/s12960-025-00977-6] [Medline: 39901258]
15. Ujoh F, Dzunic-Wachilonga A, Noor R, et al. Digital vs. conventional instructor-led midwifery training in Benue State, Nigeria: a randomized non-inferiority trial. Front Educ 2024;9:1470075. [doi: 10.3389/feduc.2024.1470075]
16. Nukunu F, Odoi P, Oppong-Boateng V, Donkor W, Bennin L, Addy A. The journey to digitalization: the story of nursing and midwifery training colleges in Ghana. Ghana J Nurs Midwifery 2024;1(1):1-14. [doi: 10.69600/z3pyqm13]

XSL•FO
RenderX

17. Janes G, Ekpenyong MS, Mbeah-Bankas H, Serrant L. An international exploration of blended learning use in pre-registration nursing and midwifery education. Nurse Educ Pract 2023 Jan;66:103514. [doi: 10.1016/j.nepr.2022.103514] [Medline: 36610859]
18. Car J, Carlstedt-Duke J, Tudor Car L, et al. Digital education in health professions: the need for overarching evidence synthesis. J Med Internet Res 2019 Feb 14;21(2):e12913. [doi: 10.2196/12913] [Medline: 30762583]
19. Garti I, Tonto OS, Shikpup NJ, Gray M. Stakeholder perspectives of the adoption and effectiveness of digital learning in midwifery education in Africa: a scoping review. SAGE Open Nurs 2025;11:23779608251380333. [doi: 10.1177/23779608251380333] [Medline: 41132903]
20. Ndayisenga JP, Babenko-Mould Y, Kasine Y, et al. Blended teaching and learning methods in nursing and midwifery education: a scoping review of the literature. Res J Health Sci 2021;9(1):100-114. [doi: 10.4314/rejhs.v9i1.11]
21. Ben Amor A, Farhat H, Aounallah A, et al. Effectiveness of a train-the-trainer pedagogical pilot workshop in improving educators' knowledge and competency in objective structured clinical examination implementation. Cureus 2026 May;18(5):e109894. [doi: 10.7759/cureus.109894] [Medline: 42222620]
22. Nexø MA, Kingod NR, Eshøj SH, Kjærulff EM, Nørgaard O, Andersen TH. The impact of train-the-trainer programs on the continued professional development of nurses: a systematic review. BMC Med Educ 2024 Jan 4;24(1):30. [doi: 10.1186/s12909-023-04998-4] [Medline: 38178050]
23. Fontaine G, Cossette S, Maheu-Cadotte MA, et al. Efficacy of adaptive e-learning for health professionals and students: a systematic review and meta-analysis. BMJ Open 2019 Aug 28;9(8):e025252. [doi: 10.1136/bmjopen-2018-025252] [Medline: 31467045]
24. Cheung L. Using the ADDIE model of instructional design to teach chest radiograph interpretation. J Biomed Educ 2016 Jun 20;2016:1-6. [doi: 10.1155/2016/9502572]
25. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
26. Villamin P, Lopez V, Thapa DK, Cleary M. A worked example of qualitative descriptive design: a step-by-step guide for novice and early career researchers. J Adv Nurs 2025 Aug;81(8):5181-5195. [doi: 10.1111/jan.16481] [Medline: 39382252]
27. Braun V, Clarke V, Boulton E, Davey L, McEvoy C. The online survey as a qualitative research tool. Int J Soc Res Methodol 2021 Nov;24(6):641-654. [doi: 10.1080/13645579.2020.1805550]
28. Erlingsson C, Brysiewicz P. A hands-on guide to doing content analysis. Afr J Emerg Med 2017 Sep;7(3):93-99. [doi: 10.1016/j.afjem.2017.08.001] [Medline: 30456117]
29. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
30. Hoang VL, Bui HT, Le TT, et al. Effectiveness of a train-the-trainer workshop series on teaching methods among nurse educators: an exemplar from Vietnam. Front Educ 2024;9:1441481. [doi: 10.3389/feduc.2024.1441481]
31. Poitras ME, Bélanger E, Vaillancourt VT, et al. Interventions to improve trainers' learning and behaviors for educating health care professionals using train-the-trainer method: a systematic review and meta-analysis. J Contin Educ Health Prof 2021 Jul 1;41(3):202-209. [doi: 10.1097/CEH.0000000000000375] [Medline: 34292260]
32. van de Water BJ, Longacre AH, Sonnie M, et al. Effectiveness of a novel midwifery preceptor program on midwives' competence and confidence in Sierra Leone: a prospective pilot study. PLOS Glob Public Health 2026;6(5):e0006294. [doi: 10.1371/journal.pgph.0006294] [Medline: 42133636]
33. Barikzai S, Bharathi S V, Perdana A. Challenges and strategies in e-learning adoption in emerging economies: a scoping review. Cogent Edu 2024 Dec 31;11(1):2400415. [doi: 10.1080/2331186X.2024.2400415]
34. Melisa M, Susanti AI. Digitalizing midwifery education: trends, tools, and transformation. Curricula J Curr Dev 2025;4(1):633-648. [doi: 10.17509/curricula.v4i1.84476]
35. Nigeria speedtest connectivity report 2024. Ookla. 2024. URL: https://www.ookla.com/research/reports/nigeria-speedtest-connectivity-report-h2-2024 [accessed 2026-01-11]
36. Ogili A, Olewe B, Eneh M. Effect of tourism on socio-economic development in Nigeria: a study of Enugu state. NG J Soc Dev 2023;11(1):85-93 [FREE Full text]
37. Obiezu-Umeh C, Subramaniam DS, Nwaozuru U, et al. Advancing sustainability in low-resource settings: development and validation of a sustainability tool for evidence-based interventions and programs. Front Health Serv 2025;5:1618400. [doi: 10.3389/frhs.2025.1618400] [Medline: 40777589]
38. Nigeria advances nursing and midwifery reforms to strengthen health workforce. World Health Organization Regional Office for Africa. 2025. URL: https://www.afro.who.int/countries/nigeria/news/nigeria-advances-nursing-and-midwifery-reforms-strengthen-health-workforce [accessed 2026-01-01]

Abbreviations

ADDIE: Analysis, Design, Development, Implementation, and Evaluation
ICM: International Confederation of Midwives
ICT: information and communication technology

XSL•FO
RenderX

NMCN: Nursing and Midwifery Council of Nigeria
NSDNM: Nigeria Strategic Directions for Nursing and Midwifery
SDG: Sustainable Development Goal
TTT: train-the-trainer

Edited by J Sarvestan; submitted 15.Feb.2026; peer-reviewed by L Herbawani, S Manurung; revised version received 15.Jul.2026; accepted 16.Jul.2026; published 31.Aug.2026.

Please cite as:
Garti I, Shikpup N, Tonto O, Gray M, Abdul HM, Ladan AM
Train-the-Trainer Evaluation of Technology-Enhanced Learning in Preservice Midwifery Education in Nigeria: Qualitative Descriptive Study
JMIR Nursing 2026;9:e93596
URL: https://nursing.jmir.org/2026/1/e93596
doi:10.2196/93596

© Isabella Garti, Nadyen Shikpup, Ophelia Tonto, Michelle Gray, Halima Musa Abdul, Awwal Mohammed Ladan. Originally published in JMIR Nursing (https://nursing.jmir.org), 31.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Preliminary Psychometric Evaluation of the Arabic Version of the NURS-DIAL Questionnaire Among Nursing Students in Saudi Arabia: Cross-Sectional Study

Nader Alnomasy[1*], PhD; Habib Alrashedi[1*], PhD; Petelyne Pangket[2], PhD

[1]Medical Surgical Nursing Department, University of Ha'il, PO Box 2240, Ha'il, Saudi Arabia

[2]Community Nursing Department, Taif University, Taif, Mecca Region, Saudi Arabia

[*]these authors contributed equally

Corresponding Author:
Nader Alnomasy, PhD
Medical Surgical Nursing Department, University of Ha'il, PO Box 2240, Ha'il, Saudi Arabia

Abstract

Background: The rapid progress in digital transformation and the increasing reliance on AI in health care have created an urgent need for validated assessments of digital and AI literacy among nursing students. However, there is a substantial gap in culturally adapted assessment tools for Arabic-speaking populations.

Objective: This study aimed to translate, culturally adapt, and conduct a preliminary psychometric evaluation of the Arabic version of the Nursing Digital and AI Literacy Assessment (NURS-DIAL).

Methods: This cross-sectional study evaluated the Arabic-translated and culturally adapted version of the NURS-DIAL. The psychometric evaluation included content validity, confirmatory factor analysis (CFA) using diagonally weighted least squares estimation, internal consistency reliability, construct validity, and known-groups validity. Given the complexity of the 52-item, 5-factor NURS-DIAL model and the limited sample size, the CFA findings should be interpreted as preliminary.

Results: Content validity was strong, with item-level content validity indices ranging from 0.83 to 1.00 and a scale-level content validity index average of 0.94. CFA provided preliminary support for a 52-item, 5-factor model after removing 4 items. Model fit indices were as follows: comparative fit index was 0.930, Tucker-Lewis index was 0.926, root mean square error of approximation was 0.055, standardized root mean square residual was 0.103, and goodness-of-fit index was 0.947. The internal consistency estimates ranged from 0.856 to 0.919. Selected known-group comparisons also provided preliminary validity evidence.

Conclusions: The NURS-DIAL demonstrated preliminary evidence of content validity, structural validity, internal consistency reliability, and construct validity among nursing students at 1 university in Saudi Arabia. Because of the limited sample size and single-institution setting, the tool should currently be used for exploratory research and formative assessment in similar settings. Larger multi-institutional studies are needed to confirm its factor structure and broader applicability.

(JMIR Nursing 2026;9:e106044) doi:10.2196/106044

KEYWORDS

digital literacy; mHealth literacy; artificial intelligence literacy; nursing students; digital well-being

Introduction

The rise of digitalization in health care and nursing education has created an increasing demand to evaluate nursing students' competence in using digital and AI technologies. Contemporary digital competency models in higher education emphasize 5 key domains: information literacy, digital well-being, digital research, content creation, and online identity management as essential for safe and effective technology use in learning and practice [1]. Recent studies have identified gaps in nursing students' self-perceived competence in information literacy, digital research, and digital innovation, while also highlighting the importance of digital health literacy, digital skills, and positive attitudes toward AI [2,3]. However, studies in the Gulf and wider Arab region have documented gaps among nursing and health professions students in information handling, critical appraisal, and digital self-management [4-6]. Moreover, a recent study of e-learning adoption in Gulf Cooperation Council (GCC) higher education institutions showed that students' use of e-learning systems was influenced by their perceptions of system usability and accessibility. These perceptions may affect students' ability to engage effectively with online learning environments [7].

The role of AI in nursing practice has grown rapidly in recent years. Nurses now use AI in clinical decision-support systems, diagnostic imaging, predictive analytics, documentation, patient

XSL•FO
RenderX

monitoring, and robotic technology. As AI becomes more integrated into patient care, nurses need to understand how to use it safely. This includes foundational knowledge of basic AI principles, the ability to use AI-supported tools appropriately, and the use of clinical judgment when considering AI-generated recommendations [8,9]. Therefore, nursing students should receive exposure to and training in AI, including their roles and responsibilities, before entering clinical practice [10,11].

In Saudi Arabia, nursing students displayed a wide range of perceptions and readiness for AI [12]. Some students expressed a lack of knowledge of and uncertainty about the concept of AI, though a majority were positively inclined to the technology [12]. Students with higher AI literacy expressed a greater sense of empowerment and a reduction in their fears about the technology [13]. Systematic reviews of the nursing literature on AI indicated that the integration of AI education programs in nursing programs can reduce nursing students' fears and improve their understanding of the technology [14]. A review of the literature on the concept of AI in the Middle East and North Africa indicated that a majority of the studies conducted and published on the topic originated from that region, but the majority of the studies did not use instruments developed within those countries to measure students' attitudes and readiness to use the technology [9].

There is a significant amount of nursing literature that discusses how Middle Eastern and North African countries feel about AI in nursing education. Some of the areas that have been studied within this region include nurses' attitudes toward AI, their literacy regarding the technology, their readiness for the technology, and their intentions to adopt AI into nursing education [9]. Despite the wealth of studies conducted within this region, there is limited reporting in the literature of how these studies were performed. In many of the Arabic language nursing education texts, the scales and questions used in the studies were developed in non-Arabic speaking countries, and there has been limited reporting of how these scales have been formally translated, culturally adapted, and psychometrically evaluated for use within the Middle Eastern and North African countries. There is a need not only to describe nursing students' attitudes and readiness for AI but, more importantly, to establish culturally appropriate measures for assessing digital and AI-related skills in Arabic-speaking educational environments. This study addressed this need by conducting an Arabic adaptation and preliminary psychometric analysis of the Nursing Digital and AI Literacy Assessment (NURS-DIAL).

The NURS-DIAL was developed to assess 5 related domains among nursing students: information literacy, mobile health (mHealth) literacy, AI literacy, digital well-being, and perceived AI education topics. The present study aimed to translate, culturally adapt, and conduct a preliminary psychometric evaluation of the Arabic NURS-DIAL among nursing students in Saudi Arabia. The study examined content validity, internal consistency, structural validity using confirmatory factor analysis (CFA), intersubscale associations, and selected known-group comparisons. Translation and cultural adaptation were undertaken using the stages recommended by Beaton et al [15], including forward translation, synthesis, back-translation, expert-committee review, and pretesting with members of the target population.

Methods

Study Design

This methodological cross-sectional study aimed to translate and conduct a preliminary psychometric evaluation of the Arabic version of the NURS-DIAL questionnaire among nursing students.

Population/Setting/Sampling

Convenience sampling was used to collect data from enrolled nursing students. Eligible students were invited to complete an online survey via institutional announcements and online learning platforms. The sample included undergraduate and postgraduate nursing students of all years of study and specializations in the field. The inclusion criteria were age 18 years or older, ability to read Arabic written language, and electronic informed consent. Exclusion criteria were students enrolled in nonnursing programs and students who participated in the pilot phase of NURS-DIAL. The final sample included 196 nursing students.

Sample-Size Consideration

The final sample included 196 nursing students. Although this sample exceeded the absolute minimum of 100 participants, it was still small in relation to the complexity of the model. The final instrument contained 52 items across 5 dimensions. Therefore, the participant-to-item ratio was approximately 3.8:1. This ratio is less than the commonly recommended ratios of 5 to 10 participants per item and more conservative recommendations for complex CFA models. In other words, the sample size was adequate for the initial psychometric evaluation but not for the final confirmation of the factor structure. The CFA findings should be viewed as preliminary and verified in further studies using larger independent samples. Additionally, future studies should use a priori model-based sample size estimation or Monte Carlo power analysis that considers the number of factors and indicators, expected factor loadings, residual variances, correlations between factors, ordinal response categories, estimator type, and anticipated missing values.

Development and Governance of the NURS-DIAL

The NURS-DIAL was created through a multinational research collaboration (AI4Health consortium) to evaluate nursing students' knowledge of digital and AI literacy in 5 areas. Four previously used tools were integrated into the NURS-DIAL: the information literacy and digital well-being subscales of Martzoukou et al [2], Discipline-Based Digital Competence Survey Battery, mHealth Literacy Scale of Kim et al [16], Artificial Intelligence Literacy Scale of Wang et al [17], and Education Topics Needs Survey of Civaner et al [18] on AI.

The initiative involves nursing faculties and departments in several countries: UMIT TIROL in Austria, the University of Maribor Faculty of Health Sciences in Slovenia, and the University of Hail in Saudi Arabia, where the current sample was recruited. All participating institutions follow a common

XSL•FO
RenderX

research protocol, obtain approval from their local ethics committees, and translate and culturally adapt the questionnaire for their own settings.

The consortium coordinating team manages the governance of the instrument, including curation of the item pool and cross-site coordination, while the project coordinators manage requests to use the instrument through the NURS-DIAL project website. Researchers and collaborating institutions can use the questionnaire at no cost. In the present study, the questionnaire was administered locally using Qualtrics and Google Forms. Parallel validation work is underway in additional cohorts, and preliminary findings have been reported in an Austrian cohort. To our knowledge, this study is the first to present an Arabic-language adaptation and preliminary psychometric evaluation of the NURS-DIAL.

The NURS-DIAL research team added 4 items to the AI Literacy subscale to better reflect AI literacy in the context of nursing and health care, as the original instrument was developed for general users, expanding the original 12-item scale [17] that measured AI literacy. Also, the AI Education Topics subscale was expanded from 15 to 20 items to include additional AI-related curricular priorities determined to be relevant to nursing education.

The Information Literacy subscale consists of 8 items, using a 5-point rating scale from novice to expert. The mHealth Literacy subscale consists of 6 items, using a 5-point rating scale from strongly disagree to strongly agree. The AI Literacy subscale consists of 16 items, using a 7-point rating scale from strongly disagree to strongly agree. The Digital Well-being subscale consists of 6 items, using a 5-point rating scale from novice to expert. The AI Education Topics subscale includes 20 items on a 5-point scale from include to no need describing the perceived priorities for AI content in the nursing curricula by students. Four items (AIL_c, AIL_g, DW_e, and AIE_t) were deleted after CFA, resulting in a final Arabic NURS-DIAL of 52 items.

Translation, Cultural Adaptation, Content Validity, and Pretesting

The NURS-DIAL was translated and culturally adapted into Arabic using the 5-stage framework described by Beaton et al [15] for cross-cultural adaptation of self-report measures. Permission to translate, adapt, and publish the questionnaire was obtained from the AI4Health consortium, and copyright requirements were met.

In stage 1, 2 bilingual translators independently translated the original English NURS-DIAL into Arabic. Both translators were native Arabic speakers, fluent in English, and had backgrounds in nursing or health sciences. In stage 2, the interdisciplinary research team reviewed and reconciled the 2 translations. The review focused on terms related to digital technology, AI, and reverse-worded items to ensure clarity and cultural relevance in the Saudi Arabic version.

In stage 3, the reconciled Arabic version was independently back translated into English by 2 bilingual translators whose first language was English, who were fluent in Arabic, and who were unfamiliar with the original questionnaire. In stage 4, a multidisciplinary expert committee reviewed the forward translations, reconciled Arabic version, and back-translations. The committee assessed semantic, idiomatic, experiential, and conceptual equivalence and produced a prefinal Arabic version of the NURS-DIAL.

Six experts reviewed the prefinal 56-item Arabic NURS-DIAL for content validity. The panel included 3 nursing education academics with 8, 12, and 15 years of experience, a health informatics expert with 10 years of experience, a linguist with 7 years of experience, and a psychometrician with 11 years of experience. All held doctoral degrees and had experience in instrument development. Experts rated each item for relevance, clarity, and cultural appropriateness on a scale from 1 (not relevant) to 4 (highly relevant). The item-level content validity index (I-CVI) values for all items were at least 0.83, and the scale-level content validity index average was 0.94. Based on the expert review, minor wording changes were made to 4 items to improve clarity and cultural appropriateness for use in Saudi Arabia.

In stage 5, the prefinal Arabic NURS-DIAL was tested with 35 nursing students who met the same eligibility criteria as the main study sample. After completing the questionnaire, students commented on the clarity of the items, wording, response options, time needed to complete the questionnaire, and cultural relevance. It took approximately 15 to 20 minutes to complete. Students did not report any concerns about cultural appropriateness, and no items were added. The research team reviewed the feedback before the main psychometric evaluation. The Arabic items, English backtranslations, response options, scoring directions, and item-retention status are provided in Table S4 in Multimedia Appendix 1.

Data Collection Procedure

Participants completed a self-administered online survey using Qualtrics and Google Forms. The survey could be accessed at participants' convenience using a smartphone, tablet, or computer. The survey link was sent out to nursing students via email, learning management system (LMS), and faculty posts for all program levels and years of study.

All invitation communications to standardize survey administration provided information on the study purpose, eligibility criteria, voluntary participation, and anticipated completion time and directed students to 1 online interface. The first page of the survey contained an information sheet and an electronic consent form, followed by items presented in a fixed sequence. The majority of closed-ended items were required; skip patterns were used only when conceptually justified; and the survey was designed for mobile accessibility to minimize missing data.

Statistical Analysis

Statistical analysis of the psychometric properties of the Arabic version of the NURS-DIAL was carried out on the main validation dataset. After screening the data, special codes were treated as missing values, and some items were reverse scored. Missingness was patterned, increasing monotonically with item position consistent with survey dropout (Spearman ρ=0.995; Tables S1 and S2 in Multimedia Appendix 1); the primary CFA used weighted least squares mean and variance (WLSMV) with

XSL•FO
RenderX

pairwise-present information, and the findings were robust to complete-case and multiple-imputation sensitivity analyses (Table S3 in Multimedia Appendix 1). Structural validity was evaluated with CFA using an appropriate estimator for ordinal data and standard fit indices. Internal consistency was evaluated by Cronbach α (≥0.70 was considered acceptable). Construct validity was assessed by intersubscale correlations for convergent and discriminant validity, while known-group validity was assessed by 2-tailed *t* tests and 1-way ANOVA on subscale scores.

Sex Reporting

In this study, sex was self-reported by participants in the online survey (male/female) and is used in the known-group validity analyses. Biological sex as recorded in official records was not assessed.

Ethical Considerations

Prior to beginning data collection, ethical approval was obtained from University of Hail's Institutional Review Board (case number H-2025 - 784). As part of obtaining informed consent, each participant was required to complete an electronic form that included a brief description of the study's purpose and procedures, potential risks/benefits of participating, as well as a statement indicating that participation was optional and that the student could withdraw or refuse to participate at any point during the completion of the survey and would incur no negative consequences. The survey was completed anonymously and did not include any identifying information.

Results

Table 1 summarizes the demographic and educational characteristics of the participants. Age data were available for 193 participants, with a mean age of 22.5 (SD 5.9, range 18 - 53) years. Among the 196 participants, 118 (60.2%) were female. Most participants were undergraduate BSc nursing students (n=185, 94.4%) and were enrolled in the general nursing program (n=163, 83.2%). Third-year students represented the largest academic-year group (n=60, 30.6%).

Table . Demographic characteristics of the nursing student sample (N=196).

Characteristic and category	Values
Age (y; n=193), mean (SD; range)	22.5 (5.9; 18 - 53)
Sex, n (%)	
Male	77 (39.3)
Female	118 (60.2)
Missing	1 (0.5)
Level of study, n (%)	
Undergraduate (BSc)	185 (98.9)
Graduate nondoctoral (MSc)	11 (1.1)
Year of study, n (%)	
1st year	21 (10.7)
2nd year	46 (23.5)
3rd year	60 (30.6)
4th year	49 (25)
5th year	20 (10.2)
Specialization, n (%)	
General nursing	163 (83.2)
Midwifery	27 (13.8)
Mental health	2 (1.0)
Other	4 (2.0)

Missing responses across questionnaire items ranged from 20.4% to 43.4%, with an overall missingness rate of 31.5%. The proportion of missing responses increased consistently as participants progressed through the questionnaire (Spearman ρ=0.995; *P*<.001), suggesting that some participants may not have completed the survey. Overall, 106 (54.1%) participants completed all 52 questionnaire items. Sensitivity analyses using complete cases and multiple imputation produced similar results; therefore, missing data were not considered to have substantially affected the study findings (Tables S1-S3 in Multimedia Appendix 1).

The final version of the Arabic NURS-DIAL is presented in Table 2. It includes the 5 retained subscales of the NURS-DIAL instrument: mHealth literacy, information literacy, AI literacy, digital well-being, and AI education topics. Each subscale was derived from an existing instrument, with some items modified or eliminated. In total, the final Arabic NURS-DIAL comprised 52 items across the 5 subscales.

XSL•FO
RenderX

All 8 information literacy items were retained for the information literacy scale. Participants responded on a 5-point scale ranging from novice to expert. All 6 mHealth literacy items were retained for the mHealth literacy scale. The responses to this scale were reversed, with higher scores indicating more mHealth literacy. The AI literacy scale was reduced from 16 to 14 items and participants responded on a 7-point scale from strongly disagree to strongly agree. The digital well-being scale was reduced from 6 to 5 items and retained its 5-point novice to expert response scale. Finally, the AI education topics scale was reduced from 20 to 19 items and used a reversed 5-point scale, with higher scores indicating greater importance of these topics.

Table . Final Arabic Nursing Digital and AI Literacy Assessment (NURS-DIAL) structure after confirmatory factor analysis (CFA): subscales, item counts, response scales, and items removed.

Subscale	Source	Original items	Final items	Items removed	Response scale
Information Literacy	Martzoukou et al [2] 2023	8	8	None	1 - 5 (Novice → Expert)
mHealth Literacy	Kim et al [16] 2024	6	6	None	1 - 5 (reversed; higher=greater mHealth literacy)
AI Literacy	Wang et al [17] 2023 (expanded)	16	14	AIL_c, AIL_g	1 - 7 (Strongly Disagree → Strongly Agree)
Digital Wellbeing	Martzoukou et al [2] 2023	6	5	DW_e	1 - 5 (Novice → Expert)
AI Education Topics	Civaner et al [18] 2022 (expanded)	20	19	AIE_t	1 - 5 (reversed; higher=greater perceived importance)

The CFA fit indices for the 5-factor Arabic NURS-DIAL model are presented in Table 3. The CFA results provided preliminary support for the proposed 5-factor structure. The model demonstrated acceptable fit according to several indices; however, the standardized root mean square residual (SRMR) exceeded conventional thresholds, and the relatively small sample size in relation to the 52-item model limits confidence in the stability and generalizability of the estimated parameters. The chi-square statistic was significant (χ^2_{1264}=1641.69; P<.001), with a χ^2/df ratio of 1.30.

Table . Confirmatory factor analysis (CFA) fit indices for the 5-factor Arabic Nursing Digital and AI Literacy Assessment (NURS-DIAL) model.

Index	Value
χ^2_{1264}	1641.69 (P<.001)
χ^2/df	1.30
CFI[a]	0.930
TLI[b]	0.926
RMSEA[c]	0.055 (90% CI 0.047 - 0.063)
SRMR[d]	0.103
GFI[e]	0.947

[a]CFI: comparative fit index.

[b]TLI: Tucker-Lewis index.

[c]RMSEA: root mean square error of approximation.

[d]SRMR: standardized root mean square residual.

[e]GFI: goodness-of-fit index.

The comparative fit index (CFI=0.930) and Tucker-Lewis index (TLI=0.926) both surpassed the benchmark of 0.90. The root mean square error of approximation (RMSEA) was 0.055 (90% CI 0.047 - 0.063), indicating acceptable fit; however, the SRMR was 0.103, exceeding the conventional ≤0.08 cutoff and indicating that the model did not fully reproduce the observed item covariances. We did not apply post hoc modifications to avoid capitalizing on chance in a modest sample (N=196); the weak-loading items identified in Table 4 (AILkR, λ=0.22; AILeR, λ=0.27; AIEpR, λ=0.36) are likely contributors to the elevated SRMR and are flagged for revision in future validation work. The goodness-of-fit index (GFI=0.947) further supported the adequacy of the specified 5-factor structure for the Arabic NURS-DIAL.

XSL•FO
RenderX

Table . Standardized factor loadings and explained variance (R^2) for Arabic Nursing Digital and AI Literacy Assessment (NURS-DIAL) items by subscale[a].

Item code	Item label (short)	λ	R^2
Information Literacy (8 items)			
IL_a	Web search engines	0.63	0.397
IL_b	Academic databases	0.73	0.530
IL_c	Online collection tools	0.72	0.523
IL_d	Evaluate trustworthiness	0.90	0.819
IL_e	Organize information	0.79	0.618
IL_f	Academic integrity	0.76	0.574
IL_g	Share/acknowledge work	0.81	0.650
IL_h	Referencing	0.77	0.596
mHealth Literacy (6 items, reverse-coded)			
mHL_a_R	Compare health information	0.76	0.577
mHL_b_R	Awareness of health apps	0.90	0.807
mHL_c_R	Familiarity with health apps	0.89	0.789
mHL_d_R	Assess quality	0.85	0.723
mHL_e_R	Use for decisions	0.76	0.573
mHL_f_R	Access to support	0.46	0.212
AI Literacy (14 items)			
AIL_a	Distinguish smart devices	0.56	0.309
AIL_b_R	Don’t know if AI helps	0.42	0.179
AIL_d	Skillfully use AI	0.80	0.647
AIL_e_R	Hard to learn AI	0.27	0.074
AIL_f	Use AI for efficiency	0.75	0.558
AIL_h	Choose AI solutions	0.73	0.530
AIL_i	Choose best AI for task	0.86	0.748
AIL_j	Comply with ethics	0.90	0.816
AIL_k_R	Not alert to privacy	0.22	0.050
AIL_l	Alert to AI abuse	0.88	0.768
AIL_m	Ethical implications	0.87	0.750
AIL_n	Aware of AI biases	0.82	0.680
AIL_o	Protect personal data	0.66	0.438
AIL_p	Evaluate AI accuracy	0.71	0.499
Digital Wellbeing (5 items)			
DW_a	Comfortable with digital tech	0.84	0.698
DW_b	Recognize digital stress	0.71	0.509
DW_c	Online behavior consequences	0.83	0.687
DW_d	Against cyberbullying	0.84	0.705
DW_f	Access well-being services	0.73	0.531
AI Education Topics (19 items)			
AIE_a_R	Knowledge of AI apps	0.60	0.359
AIE_b_R	AI reduces errors	0.70	0.491
AIE_c_R	Training in AI ethics	0.64	0.411

XSL•FO
RenderX

Item code	Item label (short)	λ	R^2
AIE_d_R	AI disease risk	0.77	0.599
AIE_e_R	AI patient compliance	0.83	0.690
AIE_f_R	AI clinical decisions	0.63	0.402
AIE_g_R	AI in epidemics	0.75	0.569
AIE_h_R	AI predicts health	0.66	0.437
AIE_i_R	AI genetic risk	0.77	0.594
AIE_j_R	AI in emergency care	0.63	0.400
AIE_k_R	AI in research	0.69	0.475
AIE_l_R	AI preventive apps	0.74	0.542
AIE_m_R	Robotic surgery	0.63	0.399
AIE_n_R	AI and patient control	0.85	0.724
AIE_o_R	Chatbots	0.52	0.275
AIE_p_R	AI psychiatric diagnosis	0.36	0.127
AIE_q_R	Legal/regulatory issues	0.59	0.348
AIE_r_R	Data privacy	0.66	0.439
AIE_s_R	Critical evaluation	0.62	0.381

[a]Abbreviated item labels are presented to facilitate interpretation of standardized factor loadings and explained variance. Full item wording in Arabic, original English items, English backtranslations, response options, scoring directions, reverse-coded status, and item retention status are provided in Table S4 in Multimedia Appendix 1.

Table 4 presents the standardized factor loadings and explained variance (R^2) for items across the five Arabic NURS-DIAL subscales. Most items showed moderate-to-high standardized factor loadings; however, several items had weaker loadings and should be evaluated in future validation studies.

The internal consistency reliability of the Arabic NURS-DIAL subscales is presented in Table 5. All domains showed good-to-excellent internal consistency. The Information Literacy subscale (8 items) had a Cronbach α of 0.890, the mHealth Literacy (6 items) and AI Literacy (14 items) subscales each showed good reliability (α=0.856), the Digital Wellbeing subscale (5 items) also had good reliability (α=0.868), and the AI Education Topics subscale (19 items) had excellent internal consistency (α=0.919). Coefficients were computed from all available responses; item-level valid sample sizes ranged from 149 to 156 for Information Literacy, 146 to 149 for mHealth Literacy, 131 to 144 for AI Literacy, 130 to 131 for Digital Wellbeing, and 111 to 129 for AI Education Topics (Tables S1 and S2 in Multimedia Appendix 1).

Table . Internal consistency reliability of the Arabic Nursing Digital and AI Literacy Assessment (NURS-DIAL) subscales[a].

Subscale	Items	Valid n per item[b]	Cronbach α[c]	Interpretation
Information Literacy	8	149 - 156	0.890	Good
mHealth Literacy	6	146 - 149	0.856	Good
AI Literacy	14	131 - 144	0.856	Good
Digital Wellbeing	5	130 - 131	0.868	Good
AI Education Topics	19	111 - 129	0.919	Excellent

[a]All α ≥0.70 indicate acceptable internal consistency.

[b]Valid n varied across items within each subscale due to missing data; the range of valid responses per item is reported.

[c]Cronbach α was computed using listwise deletion within each subscale.

Table 6 shows the intersubscale correlations of the Arabic NURS-DIAL and provides evidence of construct validity across the 5 domains. The information literacy subscale was positively correlated with mHealth literacy (r=0.263; P=.001), AI literacy (r=0.218; P=.009), and digital well-being (r=0.507; P<.001), but not correlated with AI education (r=−0.004). There were positive correlations between mHealth literacy and AI literacy (r=0.338; P<.001), digital well-being (r=0.244, P=.004), and AI education (r=0.290, P<.001). This indicates that higher mHealth literacy is somewhat associated with higher AI-related skills and education. AI literacy correlated significantly with digital well-being (r=0.402; P<.001) and AI education (r=0.327; P<.001). Digital well-being was positively related to AI Education (r=0.234; P<.01; small-to-moderate effect size).

XSL•FO
RenderX

Table . Inter-subscale correlations for the Arabic Nursing Digital and AI Literacy Assessment (NURS-DIAL; construct validity).

Subscale	Information Literacy	mHealth Literacy	AI Literacy	Digital Wellbeing	AI Education
Information Literacy	1.000	0.263[a]	0.218[a]	0.507[b]	−0.004
mHealth Literacy	0.263[a]	1.000	0.338[b]	0.244[a]	0.290[b]
AI Literacy	0.218[a]	0.338[b]	1.000	0.402[b]	0.327[b]
Digital Wellbeing	0.507[b]	0.244[a]	0.402[b]	1.000	0.234[a]
AI Education	−0.004	0.290[b]	0.327[b]	0.234[a]	1.000

[a]*P*<.01.

[b]*P*<.001 (2-tailed).

Table 7 presents the results of exploratory comparisons of Arabic NURS-DIAL subscale scores by sex and year of study. In Table 7, no statistically significant sex differences were found for information literacy, AI literacy, digital well-being, or AI education topics (all *P*>.05), with small-to-moderate effect sizes (absolute Cohen *d*=0.19-0.28). However, female students scored significantly higher than male students on mHealth literacy (female: mean 4.03, SD 0.64; male: mean 3.81, SD 0.64; t_{142}=–2.02, Cohen *d*=–0.34; *P*=.045). Table 8, no significant differences were found across years of study for information literacy, mHealth literacy, AI literacy, or digital well-being (all *P*>.05); however, AI education topics differed significantly by year of study ($F_{4\ 1\ 1\ 4}$=5.91, η^2=0.172; *P*<.001). These findings provide preliminary evidence of known-group validity and should be interpreted as exploratory.

Table . Exploratory comparisons of Arabic Nursing Digital and AI Literacy Assessment (NURS-DIAL) subscale scores in the current Saudi validation sample by sex of study[a].

Subscale	Male, mean (SD)	Female, mean (SD)	*t* test (df)	*P* value	Cohen d
Information Literacy	3.00 (0.92)	2.83 (0.87)	1.20 (159)	.23	0.19
mHealth Literacy	3.81 (0.64)	4.03 (0.64)	−2.02 (142)	.045	−0.34
AI Literacy	5.44 (0.95)	5.67 (1.00)	−1.43 (141)	.15	−0.24
Digital Wellbeing	3.68 (0.69)	3.45 (1.02)	1.55 (134)	.13	0.27
AI Education Topics	3.62 (0.58)	3.80 (0.73)	−1.59 (128)	.11	−0.28

[a]Subscale scores are prorated means requiring ≥80% of items.

Table . Exploratory comparisons of Arabic Nursing Digital and AI Literacy Assessment (NURS-DIAL) subscale scores in the current Saudi validation sample by year of study[a].

Subscale	*F* test (df)	*P* value	η^2
Information Literacy	1.34 (4, 141)	.26	0.037
mHealth Literacy	0.81 (4, 127)	.52	0.025
AI Literacy	1.04 (4, 126)	.39	0.032
Digital Wellbeing	1.10 (4, 120)	.36	0.035
AI Education Topics	5.91 (4, 114)	<.001	0.172

[a]Subscale scores are prorated means requiring ≥80% of items.

Table 9 summarizes the internal consistency reliability, standardized factor-loading ranges, and global CFA fit indices for the Arabic NURS-DIAL. All 5 subscales demonstrated good-to-excellent internal consistency. Most items showed moderate-to-high factor loadings; however, several items had weaker loadings. The overall 5-factor model showed acceptable fit on most indices, although the SRMR exceeded the conventional cutoff.

XSL•FO
RenderX

Table . Summary of the internal consistency reliability, standardized factor-loading ranges, and global confirmatory factor analysis (CFA) fit indices for the Arabic NURS-DIAL[a].

Subscale	Final items	Cronbach α	Standardized loading range
Information Literacy	8	0.890	0.63 - 0.90
mHealth Literacy	6	0.856	0.46 - 0.90
AI Literacy	14	0.856	0.22 - 0.90
Digital Wellbeing	5	0.868	0.71 - 0.84
AI Education Topics	19	0.919	0.36 - 0.85

[a]The global fit indices for the 52-item, 5-factor model were as follows: comparative fit index was 0.930, Tucker-Lewis index was 0.926, root mean square error of approximation was 0.055 (90% CI 0.047-0.063), standardized root mean square residual was 0.103, and goodness-of-fit index was 0.947. Detailed comparative fit index results, item-level loadings, and reliability results are shown in Tables 3-5.

Discussion

Principal Findings

This study provides preliminary evidence supporting the Arabic NURS-DIAL as a culturally adapted instrument for assessing digital and AI literacy among nursing students in Saudi Arabia. The findings support content validity; internal consistency reliability; and aspects of structural, construct, and known-groups validity. The implications of the findings should be interpreted considering the study's methodological limitations.

Cross-Cultural Adaptation of the Arabic NURS-DIAL

The Arabic NURS-DIAL was developed through a structured translation and cultural adaptation process, including forward translation, reconciliation, back-translation, expert review, content validation, and student pretesting, in accordance with Beaton et al [15]. The content-validity findings and student feedback support the clarity and cultural acceptability of the Arabic version in this sample. The Arabic and English versions of the questionnaire, English backtranslations, response options, scoring directions, and item-retention status are provided in Table S4 in Multimedia Appendix 1.

The translation and adaptation process is important when assessing digital and AI literacy in Arabic-speaking nursing students. Previous studies in the Middle East and North Africa have explored attitudes toward AI, readiness to use AI, AI literacy, and willingness to use AI. Many of these studies used tools translated from English into Arabic but did not clearly explain how the Arabic versions were culturally adapted or psychometrically tested [9,19]. This study offers preliminary support for an Arabic tool that covers information literacy, mHealth literacy, AI literacy, digital well-being, and perceived AI education needs. Further studies are needed with larger and more diverse groups of Arabic-speaking nursing students.

Structural Validity

The CFA showed initial support for the 5-factor model. Most fit indices were acceptable, but the SRMR was above the recommended cutoff. The 3 retained items AILeR, AILkR, and AIEpR had low loadings. This was related to the higher SRMR. These items need further review in a larger independent sample. No post hoc changes were made because the sample size was limited.

Reliability and Construct Validity

The subscales showed good reliability. Information literacy, mHealth literacy, AI literacy, and digital well-being had good internal consistency, while AI education topics had excellent internal consistency. This means that the items within each section were generally consistent in measuring the same area. These results support the use of the Arabic NURS-DIAL with nursing students for research and low-stakes educational assessment [20,21].

Construct validity was also supported by the relationships between the subscales. Information literacy, mHealth literacy, AI literacy, and digital well-being were positively related, suggesting that students who were more confident in 1 area often reported stronger skills in other digital areas. However, AI education topics showed weaker relationships with the other areas, especially information literacy. This suggests that students' views about what AI topics should be included in nursing education are related to, but different from, their digital and AI skills. Overall, the findings suggest that the NURS-DIAL measures are related but separate areas of digital and AI literacy [22,23].

Known-Group Validity

The known-groups analysis provided early, partial support for the validity of the Arabic NURS-DIAL. Female students scored higher than male students in mHealth literacy, which may suggest that the scale can identify sex-related differences in the use of digital health tools. This finding is consistent with Lupton and Maslen [24], who reported that women often use digital technologies to seek health information and support. However, because this difference was small, it should be interpreted cautiously. No clear sex differences were found in information literacy, AI literacy, digital well-being, or AI education topics. This is in line with Alshowkan and Shdaifat [13], who suggested that sex differences in digital and AI literacy may be decreasing as students have more similar exposure to technology in education.

Students' views about AI education topics differed by year of study. This may indicate that students at different stages of the nursing program consider different AI topics important to their education. Bakarman et al [12] suggested that students may develop a greater understanding of and openness to AI as they gain clinical experience and progress through nursing education.

XSL•FO
RenderX

However, the cross-sectional design did not allow the assessment of changes in students' views over time.

The findings of El Arab et al [9] similarly showed that nursing students and nurses in the Middle East and North Africa generally had positive attitudes toward AI but reported gaps in knowledge, readiness, and experience using AI in clinical practice. In the present study, the Arabic NURS-DIAL identified lower AI literacy in areas related to privacy and the perceived difficulty of learning AI.

Apart from differences in AI education topics by year of study, no other year-group differences were observed. These exploratory findings require replication in larger and more diverse samples.

Implications for Nursing Curricula and AI-Enabled Education

The findings suggest that the Arabic NURS-DIAL may help educators identify areas in which nursing students require further support in digital and AI literacy.

First, the 5 subscales may provide information for educators and curriculum committees about students' strengths and learning needs. In this study, students had lower scores in information literacy, suggesting a need for learning activities focused on searching for information, evaluating information sources, and referencing academic work. Students also showed lower AI literacy in areas related to AI privacy and the perceived difficulty of learning AI. Educational activities addressing the safe use of AI, privacy, and the critical evaluation of AI-generated information may therefore be beneficial.

Second, students in different years of the nursing program had different views about which AI topics were important. AI content can be introduced gradually as students move through the program. Topics such as clinical decision support, ethics, and checking AI-generated information are more suitable for students with more clinical experience. Assessing students at different stages can also help faculty see whether curriculum changes improve digital and AI literacy over time.

There are currently no normative data or validated cutoff scores for the NURS-DIAL. Therefore, scores should be used to identify learning needs and support teaching, rather than for diagnosis, institutional comparison, or high-stakes decisions.

Strengths and Limitations

This study has several strengths. The Arabic NURS-DIAL was translated and culturally adapted using an established framework, reviewed by a multidisciplinary expert panel, and pretested with nursing students. The study also evaluated several psychometric properties, including content validity, structural validity, internal consistency reliability, construct validity, and known-groups validity. In addition, CFA was conducted using an estimator appropriate for ordinal data.

The findings should be interpreted in light of several limitations. The sample was recruited by convenience sampling from 1 university and consisted of undergraduate students in general nursing, which limits generalizability to other institutions, educational levels, specialties, and Arabic-speaking populations. The sample size was also limited in relation to the 52-item, 5-factor CFA model. Therefore, the factor structure should be considered preliminary. Test-retest reliability, criterion-related validity against objective digital or AI competence measures, and measurement invariance across groups were not assessed. Future studies should use multi-institutional and more diverse samples to confirm the factor structure, examine measurement performance across subgroups, and evaluate stability over time.

Conclusions

The findings provide preliminary support for the Arabic NURS-DIAL as a measure of digital and AI literacy among nursing students in the Saudi setting studied. The evidence for content validity, structural validity, internal consistency reliability, and construct validity was obtained. The limited sample size and single-institution design require cautious interpretation. At present, the instrument is most appropriate for exploratory research and formative assessment in comparable settings. Multi-institutional studies are needed to confirm the factor structure and establish its suitability for broader use.

Generative AI Use

The authors used Claude Sonnet 5.5 to enhance writing and language and grammar.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

The

Nursing Digital and AI Literacy Assessment (NURS-DIAL) questionnaire, including the Arabic version, is available for research and educational use from the coordinating team via the project page [25] and from the corresponding author on reasonable request. No subscription, license fee, or institutional membership is required. The full Arabic item set, with literal English back-translations, is provided in S4 in Multimedia Appendix 1.

Authors' Contributions

Conceptualization: NA
Data curation: NA, HA

XSL•FO
RenderX

Formal analysis: NA
Funding acquisition: NA
Investigation: NA, HA, PP
Methodology: NA
Project administration: NA
Supervision: NA
Writing – original draft: NA
Writing – review and editing: NA, HA, PP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables: Table S1 (item-level missing-data frequencies and percentages by subscale), Table S2 (missingness pattern by questionnaire position), Table S3 (sensitivity analyses comparing primary confirmatory factor analysis results with complete-case and multiple-imputation results), and Table S4 (Arabic Nursing Digital and AI Literacy Assessment [NURS-DIAL] questionnaire with full item wording, English backtranslations, and removal rationale for dropped items).
[DOCX File, 49 KB - nursing_v9i1e106044_app1.docx]

References

1. Martzoukou K, Fulton C, Kostagiolas P, Lavranos C. A study of higher education students' self-perceived digital competences for learning and everyday life online participation. J Doc 2020 Jun 29;76(6):1413-1458. [doi: 10.1108/JD-03-2020-0041]
2. Martzoukou K, Luders ES, Mair J, et al. A cross-sectional study of discipline-based self-perceived digital literacy competencies of nursing students. J Adv Nurs 2024 Feb;80(2):656-672. [doi: 10.1111/jan.15801] [Medline: 37489586]
3. Abou Hashish EA, Alnajjar H. Digital proficiency: assessing knowledge, attitudes, and skills in digital transformation, health literacy, and artificial intelligence among university nursing students. BMC Med Educ 2024 May 7;24(1):508. [doi: 10.1186/s12909-024-05482-3] [Medline: 38715005]
4. Zhao Y, Pinto Llorente AM, Sánchez Gómez MC. Digital competence in higher education research: a systematic literature review. Comput Educ 2021 Jul;168:104212. [doi: 10.1016/j.compedu.2021.104212] [Medline: 36568577]
5. Alhejaili AA, Aljohani M, Muslihi A, et al. Nursing students' information seeking behavior in Saudi Arabia: a mixed-method study. SAGE Open Nurs 2026;12:23779608261477509. [doi: 10.1177/23779608261477509] [Medline: 42602293]
6. Alnomasy N, Alrashedi H, Alsayed S, Pangket P, Alsayed R. Informatics competency and technology self-efficacy levels among undergraduate nursing students: cross-sectional study. JMIR Nurs 2026 Jun 8;9:e85283. [doi: 10.2196/85283] [Medline: 42258796]
7. Akbar M, Ogalo H, Sankar JP, Ramayah T. Exploring adoption of e-learning in the Gulf Cooperation Council (GCC) education sector. Discov Sustain 2025;6(1):207. [doi: 10.1007/s43621-025-00950-2]
8. El Arab RA, Al Moosa OA, Sagbakken M, et al. Integrative review of artificial intelligence applications in nursing: education, clinical practice, workload management, and professional perceptions. Front Public Health 2025;13:1619378. [doi: 10.3389/fpubh.2025.1619378] [Medline: 40823249]
9. El Arab RA, Alshakihs AH, Alabdulwahab SH, et al. Artificial intelligence in nursing: a systematic review of attitudes, literacy, readiness, and adoption intentions among nursing students and practicing nurses. Front Digit Health 2025;7:1666005. [doi: 10.3389/fdgth.2025.1666005] [Medline: 41079691]
10. Bozkurt SA, Aydoğan S, Dursun Ergezen F, Türkoğlu A. A systematic review and sequential explanatory synthesis: artificial intelligence in healthcare education, a case of nursing. Int Nurs Rev 2025 Jun;72(2):e70018. [doi: 10.1111/inr.70018] [Medline: 40243390]
11. Ma J, Wen J, Qiu Y, et al. The role of artificial intelligence in shaping nursing education: a comprehensive systematic review. Nurse Educ Pract 2025 Mar;84:104345. [doi: 10.1016/j.nepr.2025.104345] [Medline: 40168750]
12. Bakarman SS, Al-Shammari A, Aboshaiqah A. Nursing students' perception of and readiness for artificial intelligence in Saudi Arabia. Nurs Open 2025 Dec;12(12):e70386. [doi: 10.1002/nop2.70386] [Medline: 41329788]
13. Alshowkan A, Shdaifat E. Impact of AI literacy on well-being among nursing students-mediating roles of empowerment and anxiety: cross-sectional study. JMIR Nurs 2025 Dec 22;8:e79789. [doi: 10.2196/79789] [Medline: 41428941]
14. El-Banna MM, Sajid MR, Rizvi MR, Sami W, McNelis AM. AI literacy and competency in nursing education: preparing students and faculty members for an AI-enabled future—a systematic review and meta-analysis. Front Med (Lausanne) 2025;12:1681784. [doi: 10.3389/fmed.2025.1681784] [Medline: 41384110]
15. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]

XSL•FO
RenderX

16. Kim H, Schnall R, Yoon N, Koh SJ, Lee J, Cheon JH. Development and Validation of a Mobile-Centered Digital Health Readiness Scale (mDiHERS): health literacy and equity scale. J Med Internet Res 2024 Aug 13;26:e58497. [doi: 10.2196/58497] [Medline: 39137409]
17. Wang B, Rau PLP, Yuan T. Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. Behav Inf Technol 2023;42(9):1324-1337. [doi: 10.1080/0144929X.2022.2072768]
18. Civaner MM, Uncu Y, Bulut F, Chalil EG, Tatli A. Artificial intelligence in medical education: a cross-sectional needs assessment. BMC Med Educ 2022 Nov 9;22(1):772. [doi: 10.1186/s12909-022-03852-3] [Medline: 36352431]
19. Ta'an WF, Rababah JA, Al-Hammouri MM, Yousef J, Mukattash TL, Williams B. Validation and cross-cultural adaptation of the six-dimension scale of nursing performance- Arabic version. BMC Nurs 2024 Jan 19;23(1):55. [doi: 10.1186/s12912-024-01740-3] [Medline: 38238753]
20. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55. [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
21. Nunnally JC, Bernstein IH. Psychometric Theory, 3rd edition: McGraw-Hill; 1994.
22. Markus A, Carolus A, Wienrich C. Objective measurement of AI literacy: development and validation of the AI competency objective scale (AICOS). Comput Educ Artif Intell 2025 Dec;9:100485. [doi: 10.1016/j.caeai.2025.100485]
23. Navas-Echazarreta N, Juárez-Vela R, Martínez-Sabater A, et al. Measuring instruments for media health literacy: a systematic review of psychometric properties. Nurs Rep 2024 Oct 6;14(4):2795-2818. [doi: 10.3390/nursrep14040206] [Medline: 39449443]
24. Lupton D, Maslen S. How women use digital technologies for health: qualitative interview and focus group study. J Med Internet Res 2019 Jan 25;21(1):e11481. [doi: 10.2196/11481] [Medline: 30681963]
25. NURS-DIAL: Nursing Digital and Ai Literacy Assessment. AI4Health. URL: https://ai4health.si/projects/nurs-dial/ [accessed 2026-10-17]

Abbreviations

CFA: confirmatory factor analysis
CFI: comparative fit index
GCC: Gulf Cooperation Council
GFI: goodness-of-fit index
I-CVI: item-level content validity index
LMS: learning management system
mHealth: mobile health
NURS-DIAL: Nursing Digital and AI Literacy Assessment
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index
WLSMV: weighted least squares mean and variance

Edited by L Donelle; submitted 01.Jul.2026; peer-reviewed by C Urquhart, T Hidayat; revised version received 24.Sep.2026; accepted 05.Oct.2026; published 30.Oct.2026.

Please cite as:
Alnomasy N, Alrashedi H, Pangket P
Preliminary Psychometric Evaluation of the Arabic Version of the NURS-DIAL Questionnaire Among Nursing Students in Saudi Arabia: Cross-Sectional Study
JMIR Nursing 2026;9:e106044
URL: https://nursing.jmir.org/2026/1/e106044
doi:10.2196/106044

© Nader Alnomasy, Habib Alrashedi, Petelyne Pangket. Originally published in JMIR Nursing (https://nursing.jmir.org), 30.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Manifestations and Potential Consequences of Information Distortion in Electronic Nursing Records: Qualitative Study

Jianan Wang*, BS; Yihong Xu*, MS; Wen Li, MS; Zhichao Yang, MS; Yushu Sun, MS; Xiaoxiao Zhang, MS; Hongying Pan, MS

Nursing Department, Sir Run Run Shaw Hospital, 3 Qingchun East Road, Hangzhou, Zhejiang, China

*these authors contributed equally

Corresponding Author:
Hongying Pan, MS
Nursing Department, Sir Run Run Shaw Hospital, 3 Qingchun East Road, Hangzhou, Zhejiang, China

Abstract

Background: Nursing records are a critical reflection of nursing practices and play a key role in ensuring the quality of patient care. However, these records sometimes fail to accurately represent the real-world health truth, resulting in information distortion. The manifestations and potential consequences of such distortion remain understudied, which may hinder efforts to improve nursing practices.

Objective: This study aims to explore the primary manifestations and potential consequences of information distortion in electronic nursing records (ENRs).

Methods: We conducted semistructured interviews at a Class-A tertiary hospital in China. A total of 12 nurses from various care units were recruited through purposive and snowball sampling. Inductive content analysis was used to analyze the qualitative data.

Results: Qualitative analysis revealed that manifestations of information distortion in ENRs fall into 4 main categories: untimeliness, inaccuracy, incompleteness, and inconsistency. These categories are further subdivided into 7 subcategories: premature recording, delayed recording, incorrect data entry, fabrication, account misuse, missing data, and inconsistent data entry. The potential consequences of information distortion encompass 3 main categories: patient-related, nurse-related, and data-related. These categories are further subdivided into 5 subcategories: threats to patient safety, increases in negative emotions, damage to professional image, erosion of nursing professionalism, and impairment of secondary data use.

Conclusions: The identified manifestations encompass a wide range of nursing activities, with potential consequences affecting patients, research, and caregivers. Strengthening ethics education, implementing monitoring mechanisms, establishing standardized guidelines, and leveraging technological solutions are recommended to reduce information distortion in ENRs. Further research on contributing factors and effective countermeasures is needed to mitigate information distortion in ENRs.

(JMIR Nursing 2026;9:e96679) doi:10.2196/96679

KEYWORDS

electronic nursing records; information distortion; information ethics; data quality; nursing

Introduction

Health data reveal personal information about an individual's past, present, or future physical or mental health status [1], playing a vital role in patient treatment, diagnosis, and decision-making [2]. Most health data are stored in health information systems and are reflected in medical and nursing documentation. Nursing documentation represents the care provided and the severity of a patient's condition [3]. Proper documentation serves as a communication tool, quality assurance measure, educational resource, and research material [4,5]. In addition, it is a critical reference for doctors to understand the patient's condition and take appropriate medical actions [6]. Therefore, nurses have a professional responsibility to ensure their records are complete and accurate [7,8].

However, discrepancies between nursing records and actual patient care can be problematic. Some research has found that certain nursing activities are performed but not documented [4,5,9]. De Marinis et al [10] observed that only 40% of nursing activities are recorded in nursing documentation, while Inan and Dinç [11] found that the consistency between actual nursing care and nursing records was only 77.6%. A study measuring health care data quality revealed that every 1% reduction in data completeness resulted in a 1.21% increase in missing events [12], highlighting the need to address data quality issues. In fact, the discrepancies in nursing records constitute some forms of information distortion, which occur when information

XSL•FO
RenderX

documented in electronic nursing records (ENRs) fails to accurately reflect the underlying clinical reality [13]. Current studies focus primarily on constructing various data quality frameworks for health records [14,15]. Although information distortion is closely related to poor data quality and certain data quality dimensions (eg, incomplete documentation and documentation errors) overlap with manifestations of information distortion, it represents a broader ethical and behavioral phenomenon involving multiple forms of deviation and a wide range of contributing factors [16]. However, little research has explored information distortion in practical contexts. Existing studies have focused on isolated manifestations, such as insufficient documentation during nurse rounding or skin assessments [4,5]. A comprehensive understanding of the full spectrum of information distortion in nursing records and its potential consequences is lacking, limiting the development of effective strategies to prevent and mitigate information distortion.

Nursing records can be in paper or electronic formats. With technological advancements, ENRs have become prevalent in health care settings, often replacing paper records either partially or entirely [17]. Therefore, it is crucial to adapt to these changes and explore the manifestations of information distortion within the context of ENRs.

Although China's efforts in informatization started later than those of high-income countries, the Chinese government has actively promoted informatization through policies such as the 14th Five-Year Plan for National Informatization and the High-Quality Development Promotion Action of Public Hospitals. In recent years, the integration of advanced information technologies such as 5G, big data, and the Internet of Things into Chinese hospitals has expanded significantly in both breadth and depth [18]. According to the National Health Commission, the health information technology development index in China reached 74.14 in 2024 [19]. Data from the China Hospital Information Management Association also show that nearly 100 hospitals in China have achieved level 5 or higher in electronic medical record (EMR) adoption. The implementation of information systems such as clinical decision support systems, hospital information systems, EMRs, and picture archiving and communication systems has optimized clinical data management and improved the quality of health care services [20,21]. With the rapid growth of informatization, it is crucial to address information ethics issues to support the sustainable development of these advancements. Examples of such issues include information disclosure, distortion, alienation, and injustice. Notably, information distortion compromises data accuracy, which can directly jeopardize patient health and safety, making it a critical area that requires focused attention [13].

Therefore, this study aims to explore information distortion from a Chinese perspective, specifically examining its manifestations and potential consequences in ENRs. The goal is to enhance the understanding of information distortion and its implications for nursing practice and patient care within the Chinese context.

Methods

Study Design

When exploring social phenomena, qualitative research is suitable [22]. Therefore, a qualitative descriptive design was employed. Semistructured interviews were chosen as they allow for flexible yet guided questions, facilitating an in-depth exploration of participants' thoughts and experiences [23]. All interviews were conducted face-to-face. This study also adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines [24] (Checklist 1).

Setting

The study was conducted in a Class-A tertiary hospital in Zhejiang Province, which has been using electronic documentation for nearly a decade. The hospital has over 20 years of experience with electronic documentation and integrated a nursing decision support system into its nursing information system in 2016. Its advanced level of digitalization has provided nurses with a comprehensive technological environment, equipping them with extensive experience in using the system. This experience is invaluable for gaining insights into issues related to information distortion.

Recruitment

The interviewed nurses were recruited using purposive sampling and snowball sampling. To gain deeper insights into the complexities of information distortion, we chose to invite members of the Information Committee to participate in the interviews. Each department has a representative in the committee. These committee members are nurses who are actively engaged in clinical practice while also assuming additional roles, such as addressing system usage concerns and sharing technical expertise with their colleagues. They interact with information frequently, making them more likely to have a deeper understanding of the phenomenon of information distortion. We explained the basic concept of information distortion and the objectives of the research to committee members, allowing them to decide freely whether to participate. Since information systems differ across specialties, we used maximum variation sampling based on departmental categories to ensure that participants represented a broad range of clinical nursing practices. Additionally, committee members were invited to recommend other nurses who had been involved in clinical informatics-related work or projects as potential interviewees. The inclusion criteria for nurses were as follows: (1) clinical registered nurses employed at the hospital for at least 1 year and (2) voluntary participation with signed informed consent. Nurses from other hospitals who were present for training, as well as those absent during the study period (eg, on leave, on vacation, or attending external training), were excluded.

Data Collection

A researcher (JW), a female PhD student trained in qualitative methods, conducted a total of 12 face-to-face semistructured interviews between January 5, 2024, and April 8, 2024. The interviews were conducted in Chinese, either in a coffee shop or in a demonstration classroom within the hospital. Data

XSL•FO
RenderX

saturation was reached when no new insights emerged during the interviews. No other individuals were present besides the participants and the interviewer. The interviewer had no prior relationship with the participants, which helped reduce biases and potential ethical concerns. As part of the interview process, participants completed a short questionnaire about their personal characteristics, including gender, age, education, position, professional title, work experience, and work department. A topic guide with open-ended questions (Multimedia Appendix 1), developed through a review of existing literature and expert opinions and tested by 2 nurses beforehand, was used to cover all relevant topics during the interviews. Participants were asked to describe instances of information distortion they had experienced or observed in their clinical practice, as well as the potential consequences, including those they had witnessed among their colleagues. The interviewer took field notes during the interviews to enhance the data and mark key moments [25]. Interviews lasted between 35 and 91 minutes and were recorded with the written consent of the participants. The audio recordings were transcribed into Chinese by the researcher (JW) and then verified by the participants. No repeat interviews were conducted.

Data Analysis

Inductive content analysis was conducted to analyze the qualitative data, following 3 phases: preparation, organizing, and reporting [26]. Two researchers (JW and YX) independently read the transcripts line by line, noting as many headings as necessary to capture all aspects of the content [27], using NVivo 12 (QSR International Pty Ltd). The headings were compiled into coding sheets, and categories were generated and supported by quotations derived from the interviews. The 2 researchers then compared and discussed each code and category, refining them and adding relevant codes and potential categories as needed. Participants also received a summary of the identified categories along with their own quotations for review, offering them an opportunity to provide feedback on the accuracy of the findings in reflecting their experiences. Finally, the entire research team reviewed and discussed the identified categories and subcategories to ensure coherence.

Since this study is presented in English, analyzing data in Chinese might raise concerns regarding trustworthiness. To address this, a qualified translator first translated the initial transcripts into English. During the process of identifying categories, subcategories, and representative quotations, the bilingual authors (JW and YX) carefully reviewed and adjusted the translations to ensure they accurately captured the intended meanings within the Chinese context in nursing. The other authors then independently reviewed the content for accuracy and provided feedback.

Rigor

The rigor of this study was ensured by the trustworthiness of the data (credibility, transferability, dependability, and confirmability) [28]. Credibility was ensured by having 2 researchers independently code the data, followed by discussions within the research group to resolve any ambiguities or discrepancies. Furthermore, member checks were conducted by inviting participants to review and validate the identified categories and subcategories. Transferability was ensured by providing sufficient contextual information about the information environment in China, enabling readers to relate the findings to their own positions. Dependability was ensured through a detailed description of the participants' inclusion and exclusion criteria, data collection and analysis procedures, and participant characteristics. Finally, confirmability was ensured through regular reflection on the researchers' biases and a commitment to transparency and openness. The researchers also received comprehensive training in qualitative research methods, equipping them with the essential skills and knowledge needed to conduct the study reliably.

Ethical Considerations

This study is part of a larger research project, and ethical approval was obtained from the Ethics Committee of Sir Run Run Shaw Hospital, Zhejiang University School of Medicine (approval number 2024; research number 0270). Prior to the interviews, each participant signed a written informed consent form after full disclosure of the study's aim. All participants were informed that they could withdraw from the study at any time without any obligation. Data analysis and reporting were conducted anonymously to protect participants' privacy.

Results

Participant Characteristics

All participating nurses were women, with an average age of 32 (SD 3.5, range 26 - 37) years and work experience ranging from 4 to 14 years. They worked in various nursing units, including internal medicine, surgery, and the intensive care unit. Participant characteristics are further described in Table 1.

XSL•FO
RenderX

Table . Characteristics of participants (N=12).

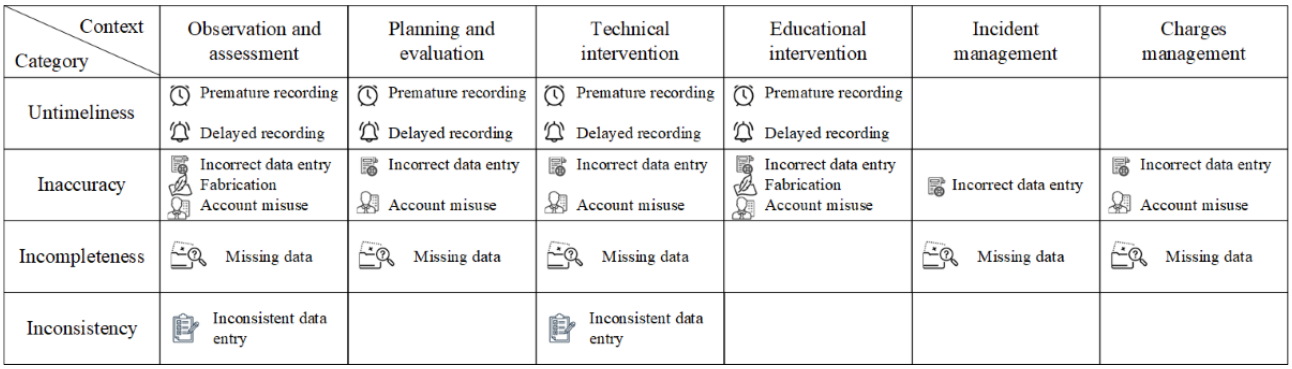
Characteristics	Values, n (%)
Sex	
Female	12 (100)
Age (y)	
25 - 29	3 (25)
30 - 34	5 (41.7)
35 - 39	4 (33.3)
Education	
Bachelor's degree	10 (83.3)
Master's degree	2 (16.7)
Post	
Clinical nurse	8 (66.7)
Advanced practice nurse	2 (16.7)
Nurse educator	2 (16.7)
Professional title	
Primary nurse	3 (25)
Charge nurse	9 (75)
Work experience (y)	
1-5	2 (16.7)
6-10	6 (50)
11-15	4 (33.3)
Work unit	
Internal medicine	7 (58.3)
Surgery	4 (33.3)
Intensive care unit	1 (8.3)

Manifestations of Information Distortion in ENRs

Overview of Manifestations

The content analysis revealed that manifestations of information distortion in ENRs fall into 4 main categories: untimeliness, inaccuracy, incompleteness, and inconsistency. Seven subcategories were derived: premature recording, delayed recording, incorrect data entry, fabrication, account misuse, missing data, and inconsistent data entry. In addition, we identified specific nursing contexts in which these manifestations are more likely to occur, including observation and assessment, planning and evaluation, technical intervention (preventive and curative), educational intervention, incident management, and charges management. The findings are synthesized in Figure 1.

Figure 1. Manifestations of nurses' information distortion in electronic nursing records.



XSL•FO
RenderX

Untimeliness

This type of manifestation occurs when nurses write ENRs at inappropriate times. This category is subcategorized into 2: premature recording and delayed recording.

Premature Recording

Nurses indicated that they would record times before the events actually occurred. Nurses stated that they might complete the assessment forms and nursing care plans before actually checking patients.

> *Some people have a habit of copying the assessment forms and care plans for their patients in advance, then doing rounds to check the patients. Then, they go back and make changes. With so many patients, it could take an hour to finish, but they don't change the recorded time on the computer.* [N3]

Another nurse stated that she recorded the time earlier than the actual time when performing interventions, such as infusing blood:

> *We have drip rate requirements for blood transfusions, for red blood cells, it's 60 drops per minute, and there's a formula to figure out when it should be done. But sometimes, with all the factors affecting the tubing, it's impossible to meet that exact time, so I'll close the loop in the PDA (Personal Digital Assistant), you know, I have to mark it as finished in the record, even though it's still infusing.* [N7]

Nurses also stated that the education time was sometimes earlier than the actual time:

> *We sometimes check off the education part when the medical order is issued, then give the education, like providing education when we're actually giving out the medication. So the timing isn't quite right.* [N5]

Delayed Recording

Nurses mentioned that they might also record event times later than they actually occurred. A nurse said that they might forget to assess or give care plans and make up for it later:

> *There are so many things, so sometimes I forget to update the assessments or the care plans. Then, right before leaving, I'll check and go, 'Oh no, I missed that', and quickly fill it in.* [N1]

Nurses also explained that this happens during curative interventions. When they administer medication while awaiting a delayed doctor's order, they give the medication first and document the time afterward to comply with regulations. One nurse said the following:

> *We (with doctors) agreed to administer 8 units of insulin, but the doctor's order only specifies 6 units. We might ask the doctor to update the order to add the extra 2 units later. However, you might end up administering the insulin before the order is updated and yet the time on your records looks perfect.* [N3]

In educational terms, in addition to premature recording, nurses mentioned that they might sometimes delay recording:

> *When the meds are prescribed, they pop up one by one, and I can't just sit there and watch each one to check it off, right? So I'll catch up on it later. Like, I might go through all the education items from the morning at noon. I don't care what time they were ordered—7, 8, 9, or 10 o'clock—I just check them all off at once.* [N2]

Inaccuracy

This type of manifestation indicates that the contents of ENRs cannot reveal the truth. This category is further divided into 3 subcategories: incorrect data entry, fabrication, and account misuse.

Incorrect Data Entry

There are times when events described in ENRs do not accurately reflect the truth. When assessing the patient's condition, some issues may arise, such as errors in location. Some nurses noted:

> *Once during a tube assessment, when I took over from another nurse, I found out that for nearly half a month, the record said the tube was in the left nostril when it was actually in the right one.* [N1]

Nurses also mentioned that they would intentionally assess the patient's condition as more severe to alert doctors:

> *To use restraints, we need a doctor's order. Before the doctor gives the order, the patient's risk score has to be pretty high. So we might rate them higher than actual to make sure they get the restraints.* [N12]

Nurses also found that the nursing care plans they create are sometimes not suitable for patients, and their evaluation faces similar issues. One nurse noted:

> *In fact, some of patients even know more than we do, but we still put down "knowledge deficit" (when making the nursing care plan)…While the patient is in the hospital, it's rare for us to mark nursing evaluations as "completed." We usually just keep marking 'partially completed' until they're discharged, even if they no longer have the issue.* [N3]

When performing preventive interventions, nurses mentioned that they choose to continue providing nursing care for patient safety even after the order had been discontinued:

> *For example, at 9 o'clock, the doctor's restrictive order expired. We documented in our nursing records that we stopped the restriction, but in reality, we kept enforcing it to keep the patient safe, without asking the doctor to renew the order.* [N12]

Educational templates offer convenience by significantly reducing nurses' workloads. However, they can lead to errors if nurses fail to apply their own judgment. One nurse said the following:

> *We have a template for VTE (venous thromboembolism) discharge education, it should be adjusted based on the patient's needs. However, we often use it without any change…Like the template*

XSL•FO
RenderX

suggests quitting smoking but the patient doesn't even smoke. [N6]

Additionally, nurses mentioned that when they have reported an incident, such as a pressure ulcer, they may not be able to confirm that it has healed. One nurse said:

I think pressure ulcers are also a problem. We may think they have healed, but the assessment forms are still filling up, no one records "healed." [N6]

Regarding charges, nurses admitted that they sometimes intentionally overestimate a patients' condition to increase their incomes:

We sometimes score the Braden scale at 18 or below on purpose so we can charge for it, as only scores of 18 or below qualify for high-risk pressure ulcer fees. For some patients who have undergone ERCP (Endoscopic Retrograde Cholangiopancreatography), their actual scores might be above 18, but we score them lower to fit the charging criteria. [N1]

Fabrication

There have been instances in which nurses documented tasks in ENRs that were not actually performed. Under the pressure of excessive workloads, some nurses may resort to fabricating their observation records for convenience:

For critical care charts, we have to record data every hour. Sometimes if we're really busy and the patient seems fine, we'll just fill in that hour based on other times. [N9]

If I'm in charge of rooms 9 to 16, I'll record the same position for everyone in those rooms at a certain time. Like at 10 o'clock, they all 'left side'. And at 12 o'clock, they all 'right side.' [N4]

Moreover, nurses highlighted that fabrications are more common in their education records:

Education problems are pretty big. When handing out oral medications, there are too many medications, like a dozen or more. We might miss the education step, we don't explain what the new medications are for, we just focus on checking the patient's identity, even though the records show that we provided the education…We just don't have the time or energy to explain what the basic prevention measures for VTE include and what needs to be done. But the records do show that we provided education on VTE prevention. [N2]

Account Misuse

Nurses have their own accounts on ENRs and personal digital assistants, and misuse of these accounts can sometimes occur. One nurse said:

Some people don't log out of the PDA after their shift, and sometimes I don't check carefully. So I might use the PDA all day with the previous shift's account. All the charges and records for the day end up under their name. [N11]

The hospital has specific rules for correcting records, allowing only nurse leaders to make changes after 24 hours have passed. However, nurses often use their leaders' accounts to make corrections without informing them.

We all know our leader's code. Sometimes, if there's a mistake in the record, we use the leader's code to make corrections. [N4]

Incompleteness (Missing Data)

This type of manifestation indicates that the content of ENRs is less comprehensive than the actual nursing activities performed. This category consisted of a single subcategory: missing data.

Nurses stated that they tend to omit records either intentionally or unintentionally. One nurse mentioned that she did not record her thoughts during assessment because her experienced colleagues did not write them down:

They (experienced nurses) didn't write it, so maybe it's not necessary. [N3]

Nurses also stated that new nurses are especially more likely to miss data. One nurse described this issue when discussing the nursing care plan:

Some less experienced nurses, for example, might focus all their efforts on bringing down a patient's fever after the patient has had it for a day. But they might forget to include the issue of the high temperature in their documentation. [N8]

When performing technical interventions, nurses also acknowledged that they sometimes omit certain details from documentation.

When an infusion gets interrupted, you usually don't really pay attention to it. If a patient needs to go for a test and there's half a bag of saline still hanging, you won't remember to mark it as interrupted. [N2]

Incident management helps prevent error recurrence and improve care quality [29]. However, nurses tend not to report medication errors:

In our oncology unit, patients receiving chemotherapy also get an anti-nausea medication like ondansetron every 8 hours. For instance, if it's given around 11 or 12 in the morning, we prioritize using the gastroprotective medication before chemotherapy. But with so many other medications to give throughout the day, sometimes as late as 10 or 11 at night, the timing for that small dose of ondansetron gets pushed back. It's actually a near miss, but we don't report it. [N6]

When it comes to charges, nurses stated that their established work habits sometimes lead to omissions in documentation.

Our work habit is to charge for services in the morning. So, if a nebulizer treatment starts around noon, we won't charge for that day, just once. In this case, a mask is given for that day, but no fee is collected. [N1]

XSL•FO
RenderX

Inconsistency (Inconsistent Data Entry)

This type of manifestation refers to contradictions in data within ENRs. There may be disagreement between data in ENRs or between ENRs and another data source such as EMRs. This category consists of one subcategory: inconsistent data entry.

The lack of interoperability could hinder health professionals from exchanging information. Nurses stated that insufficient integration of information systems across different departments or systems could lead to inconsistent data in assessment. The lack of standard medical codes could lead to data recorded about the same patient having different descriptions:

> *This patient was transferred from the ER (Emergency Room) to our ward, but I can only find his medical records by checking the discharge notes. If I look in our system, there's nothing there, it's like I'm dealing with a new patient.* [N1]

> *There are situations where the doctor and we nurses both assess the patient, but we record different chief complaints. The doctor might write "diagnosed with pancreatic cancer over seven months ago," while we write "a few days post-chemotherapy."* [N7]

In addition, the introduction of data capture technologies, which make nursing practices more labor-saving and effective, can sometimes also lead to inconsistencies with manually entered records. Nurses expressed this issue during curative interventions:

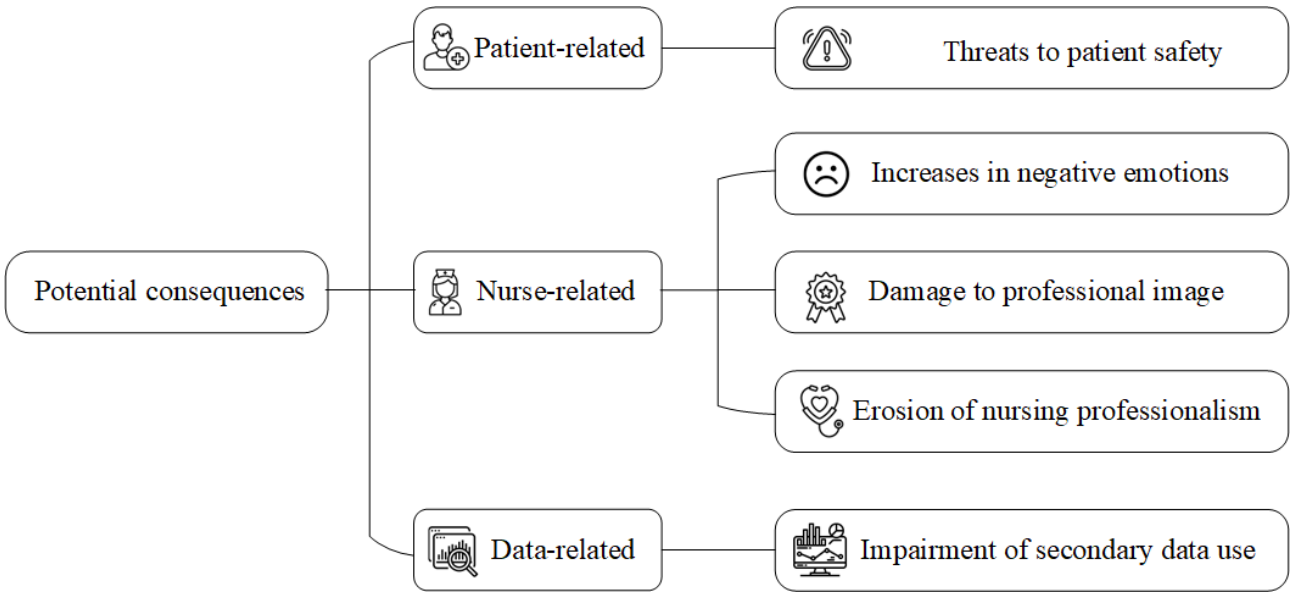
> *If a patient's oxygen saturation drops, we handle the situation first and document it later. But vital signs are updated periodically, so while you might note the low oxygen levels, the recorded vital signs at that time could show normal, creating an inconsistency between the two.* [N7]

Potential Consequences of Information Distortion

Overview of Potential Consequences

The data indicated that the possible consequences of information distortion could be classified into 3 categories: patient-related, nurse-related, and data-related. These can be further divided into 5 subcategories: threats to patient safety, increases in negative emotions, damage to professional image, erosion of nursing professionalism, and impairment of secondary data use. The findings are synthesized in Figure 2.

Figure 2. Potential consequences of information distortion in electronic nursing records.



Patient-Related: Threats to Patient Safety

Nurses indicated that failing to document changes in a patient's condition could impact both nurses' and doctors' judgments about the patient. Inaccurate or omitted documentation may obscure the patient's true clinical status, delaying the recognition of deterioration and the timely escalation of care, which may ultimately result in delayed treatment. Furthermore, distorted documentation may contribute to medication errors by leading clinicians to misjudge the patient's condition and make inappropriate clinical decisions, thereby posing significant risks to patient safety.

> *We often rely on the critical care sheet or assessment sheet to track significant changes in a patient's condition. If our previous records aren't accurate, it will definitely affect how we judge the patient's condition later, right? We wouldn't know when the changes actually started, and this could throw off our focus.* [N10]

> *If a patient's blood pressure reads 180 but he's shivering, and you only record the high blood pressure and inform the doctor, the doctor will trust that data and might prescribe other medications. If they prescribe a blood pressure med, like oral nitroglycerin, and you give it to the patient who's actually shivering, which could mean they have an infection, right? If they're in septic shock, their blood pressure would already be dropping. Giving them a blood pressure medication could be really dangerous.* [N5]

XSL•FO
RenderX

Nurse-Related

Increases in Negative Emotions

Nurses expressed that the appearance of information distortion could make them feel regret or upset, intensifying their negative emotions.

> *(If I don't have the information distortion) I won't be regretting at home that I didn't write accurate records, thinking about what I should have done at that time.* [N6]

> *If they find my record is distorted, I'll definitely feel a bit down about it. I mean, who doesn't want their records to be perfect?* [N12]

Damage to Professional Image

When information in the records does not match the actual situation, the likelihood of medical disputes may increase, and patients' trust in nurses may decline, affecting nurses' professional image.

> *If there's a problem with the drug's expiration and patients notice it, especially if it says something like "12 hours", it can cause a medical dispute cause issues. And patients might feel they can't fully trust we nurses, thinking, 'How could they make mistakes on something like this?'* [N8]

Erosion of Nursing Professionalism

Constantly distorted information can desensitize nurses, causing them to overlook important possibilities. This can erode nursing awareness and expertise, ultimately undermining the thoughtful and conscientious spirit essential to the profession.

> *Many patients, including those who've had surgery, often have a care plan for a long time before they're discharged. We rarely make any changes, so there's hardly any real awareness of their care needs.* [N4]

Data-Related: Impairment of Secondary Data Use

Secondary data use refers to the use of data for purposes other than those for which it was originally collected, such as quality improvement or research analysis [30]. As ENRs become increasingly prevalent in clinical practice, secondary analysis of these data is also commonplace. In terms of incident management, nurses expressed that distorted information about the cause of an incident can introduce bias in quality improvement, leading to ineffective improvement measures. As a result, incidents may continue to occur, rendering significant efforts futile.

> *Because this (incident reporting) involves the subsequent analysis, why did this unexpected event happen? Were your measures effective? I think it is necessary to restore the actual situation. For example, if I say I put a restraint mitt on the patient but they still pulled out the tube, during analysis, wouldn't they think that even with the mitt, the patient could pull out the tube? But if the mitt was not actually used, then maybe if it had been, the patient wouldn't have pulled out the tube. Without an accurate record of the real situation, you can't analyze the reason properly.* [N12]

From a research perspective, nurses noted that distorted records could undermine the validity of findings generated from such data, ultimately compromising the reliability of scientific research.

> *If I'm working on a pressure ulcer risk prediction model, things like how long a patient stays in bed or how often their position is changed are key factors. If those records aren't accurate, they will affect the model's accuracy. In that case, what's the point of developing such a model?* [N7]

Discussion

Principal Findings

This research is the first to clearly identify manifestations of information distortion in ENRs. Although previous studies do not explicitly use the term "information distortion," some specific manifestations of distortion have been explored, including failure to record in a timely manner [31,32], inaccurate data [30,33], nursing practices performed but not recorded [4,5], and discrepancies in the same information [34]. These findings are related to our results, which include manifestations of delayed recording, incorrect data entry, missing data, and inconsistent data entry. Our study further contributes by adding manifestations of premature recording, fabrication, and account misuse within the context of Chinese nursing. Some of the identified manifestations raise ethical and professional concerns, as they may undermine the principles of accountability, responsibility, and expertise outlined in the International Council of Nurses Code of Ethics for Nurses [35]. Structured ethics education and training are needed to strengthen nurses' awareness of accurate and responsible documentation practices. In addition, real-time documentation monitoring systems could be developed to provide oversight and support early identification of documentation-related issues. However, it is important to recognize that not all instances of information distortion represent intentional misconduct. For example, some behaviors identified in this study, such as overassessment and continued restrictions, were driven by nurses' intentions to enhance patient safety, even when this required them to assume additional responsibilities. Therefore, considering the complexities of the clinical context in which these behaviors occur, it is inappropriate to see the issues in black and white.

Another important finding is that manifestations of information distortion permeate almost every facet of nursing practice, which is concerning. Previous studies have explored manifestations of information distortion in various nursing contexts, including assessments, evaluations, technical and educational interventions, and incident reporting [5,10,36]. Our study expands on this by identifying information distortion in nursing care plans and charge management, which may be related to the hospital's comprehensive electronic nursing systems, which provide nurses with more intuitive user experiences. In our interviews, participants from every department reported experiencing or witnessing information distortion in ENRs, suggesting that this issue is not limited to specific departments,

XSL•FO
RenderX

such as those that are extremely busy or poorly managed. Considering the widespread occurrence of information distortion, standardized documentation guidelines tailored to different nursing contexts are needed. In addition, review and feedback mechanisms should be established not only through hierarchical supervision but also through peer-to-peer feedback to improve documentation quality.

We also found that the potential consequences of information distortion in ENRs affect patients, caregivers, and the secondary use of data. Some studies on data quality have pointed out that incomplete or inaccurate data can put patients' health care at risk [37] and affect the reuse of data for analysis [38,39], which is consistent with some of our findings, including threats to patient safety and impairment of secondary data use. Our research further explores the nurse-related potential consequences, offering an example of the impact of information distortion on health care professionals—an area seldom investigated in previous studies. Future studies could further develop and refine frameworks for understanding potential harms associated with health data based on these findings [40].

Although the issue of information distortion has existed since the era of paper-based records, manifesting in behaviors such as reluctance to report incidents or falsification of records, some of the manifestations identified in this study suggest that advancements in information technology may be influencing this phenomenon. Examples include the "copy then adjust" habit, the inappropriate use of templates, and the misuse of accounts, which may align with Burnum's perspective that the development of health information technology, such as electronic health records, has not improved the quality of recorded data; instead, it has resulted in the recording of a larger amount of poor-quality data [41]. However, the trend toward greater digitization is irreversible. Given the pervasive manifestations of information distortion and the severe potential consequences, further research is needed to explore the influencing factors and develop strategies to address information distortion in ENRs, with the ultimate aim of enhancing the quality of records. Our research has highlighted factors such as heavy workload, years of experience, lack of interoperability, and the absence of standardized medical codes, which align with existing literature [34,42]. Potential countermeasures, such as the adoption of AI scribes and improvements in interoperability among electronic health record systems, could be considered to reduce information distortion [43]. Further research is needed to explore more related factors and countermeasures based on our findings.

Limitations

This study has several limitations. Although the sample size may appear small, data saturation was achieved, which is consistent with the typical range of 12 to 20 interviews required for qualitative research [44]. Although nurses were recruited from diverse departments, the study was conducted in a single hospital in China, and the findings may have been influenced by its specific organizational and cultural context. Nevertheless, this study provides valuable exploratory insights into information distortion, highlighting potential trends that can be further explored in multicenter studies to strengthen generalizability. Furthermore, the findings were based on participants' self-reported experiences. Because the interviews explored ethically sensitive documentation practices, participants may have underreported or minimized problematic behaviors due to social desirability bias, despite assurances of confidentiality. Their willingness to disclose such practices may also have been influenced by organizational culture and workplace norms. In addition, this study relied solely on interview data and did not incorporate methodological triangulation, such as direct observation or document analysis. Future research using multiple data sources could provide a more comprehensive understanding of information distortion in ENRs. Finally, all participants were female, and some were purposively recruited for their informatics-related experience. While these participants provided rich insights into electronic nursing documentation, their perspectives may differ from those of nurses without such experience, potentially limiting the applicability of the findings to the broader nursing population. Nevertheless, given the predominance of female nurses in the Chinese nursing workforce, the lack of male participants was difficult to avoid. Moreover, we included nurses with limited informatics-related experience, and purposive sampling was adopted to maximize variation in informatics experience and departmental backgrounds, thereby enriching the breadth and depth of the data.

Conclusions

This research identified 4 main categories of information distortion manifestations in ENRs and 3 categories of their potential consequences. These findings underscore the importance of raising awareness of information distortion in health care data. Ethics training, monitoring mechanisms, standardized documentation guidelines, and technology-enabled solutions are recommended to address this issue. Further research is needed to analyze the contributing factors and develop effective countermeasures to improve the quality and reliability of nursing records.

Acknowledgments

The authors sincerely thank all the participants for their time and valuable contributions. No generative AI tools were used in the preparation of this manuscript.

Funding

This work was supported by the Health Commission of Zhejiang Province (grant number 2024KY1142) and by the China Scholarship Council program.

XSL•FO
RenderX

Authors' Contributions

JW designed the study protocol and collected the data. JW and YX coded and analyzed the data under HP's guidance and supervision. JW drafted the initial manuscript, and all other authors contributed to critical reviews and revisions of the manuscript. All authors have approved the final manuscript and agreed to its publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Topic guide.
[DOCX File, 19 KB - nursing_v9i1e96679_app1.docx]

Checklist 1
COREQ checklist.
[DOCX File, 20 KB - nursing_v9i1e96679_app2.docx]

References

1. Recital 35 Health data. GDPR.EU. 2018. URL: https://gdpr.eu/recital-35-health-data [accessed 2024-10-19]
2. Li Y, Yang J, Zhang Z, Wen J, Kumar P. Healthcare data quality assessment for cybersecurity intelligence. IEEE Trans Ind Inf 2023;19(1):841-848. [doi: 10.1109/TII.2022.3190405]
3. Brazelton NC, Knuckles MC, Lyons AM. Clinical documentation improvement and nursing informatics. Comput Inform Nurs 2017 Jun;35(6):271-277. [doi: 10.1097/CIN.0000000000000367] [Medline: 28591064]
4. Lee S, Yu S, Park H, Kim TG. Assessing how well electronic nursing records reflect changes in the nursing profession's scope of practice in Korea. Comput Inform Nurs 2020 Jun;38(6):312-319. [doi: 10.1097/CIN.0000000000000585] [Medline: 31913862]
5. Park H, Yu S, Lee SH. Evaluating the correlation between nursing practice and electronic nursing records using importance-performance analysis. Comput Inform Nurs 2021 Apr 13;39(9):492-498. [doi: 10.1097/CIN.0000000000000737] [Medline: 33871385]
6. Kebede M, Endris Y, Zegeye DT. Nursing care documentation practice: the unfinished task of nursing care in the University of Gondar Hospital. Inform Health Soc Care 2017 Sep;42(3):290-302. [doi: 10.1080/17538157.2016.1252766] [Medline: 27918228]
7. Goldsmith J. The NMC code: conduct, performance and ethics. Nurs Times 2011;107(37):12-14. [Medline: 22010552]
8. ANA's principles for nursing documentation: guidance for registered nurses. : American Nurses Association; 2010 URL: https://anaprodsite1.nursingworld.org/globalassets/docs/ana/ethics/principles-of-nursing-documentation.pdf [accessed 2024-09-26]
9. Brooks JT. An analysis of nursing documentation as a reflection of actual nurse work. Medsurg Nurs 1998 Aug;7(4):189-196. [Medline: 9782888]
10. De Marinis MG, Piredda M, Pascarella MC, et al. 'If it is not recorded, it has not been done!'? Consistency between nursing records and observed nursing care in an Italian hospital. J Clin Nurs 2010 Jun;19(11-12):1544-1552. [doi: 10.1111/j.1365-2702.2009.03012.x] [Medline: 20438599]
11. Inan NK, Dinç L. Evaluation of nursing documentation on patient hygienic care. Int J Nurs Pract 2013 Feb;19(1):81-87. [doi: 10.1111/ijn.12030] [Medline: 23432893]
12. Johnson SG, Speedie S, Simon G, Kumar V, Westra BL. Quantifying the effect of data quality on the validity of an eMeasure. Appl Clin Inform 2017 Oct;8(4):1012-1021. [doi: 10.4338/ACI-2017-03-RA-0042] [Medline: 29241241]
13. Wang J, Xu Y, Zhang R, Yang Z, Zhang X, Pan H. Information distortion in electronic health records: a concept analysis. J Adv Nurs 2026 Jun;82(6):5881-5891. [doi: 10.1111/jan.70233] [Medline: 40984662]
14. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. J Am Med Inform Assoc 2013 Jan 1;20(1):144-151. [doi: 10.1136/amiajnl-2011-000681] [Medline: 22733976]
15. Syed R, Eden R, Makasi T, et al. Digital health data quality issues: systematic review. J Med Internet Res 2023 Mar 31;25:e42615. [doi: 10.2196/42615] [Medline: 37000497]
16. Wang J, Xu Y, Yang Z, et al. Factors influencing information distortion in electronic nursing records: qualitative study. J Med Internet Res 2025;27:e66959. [doi: 10.2196/66959] [Medline: 40202777]
17. Häyrinen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. Int J Med Inform 2008 May;77(5):291-304. [doi: 10.1016/j.ijmedinf.2007.09.001] [Medline: 17951106]
18. Feng Y, Nie C. The effect of government informatization construction on corporate digital technology innovation: new evidence from China. Business Ethics Env & Resp 2025 Jul;34(3):657-676. [doi: 10.1111/beer.12671]

XSL•FO
RenderX

19. Release of the national health informationization development index (2024) [Article in Chinese]. National Health Commission's Statistical Information Center. 2024. URL: http://www.nhc.gov.cn/mohwsbwstjxxzx/s8553/202407/6142249688c14e13968041b42d3c35c1.shtml [accessed 2026-08-06]
20. 2021-2022 survey report on the informatization status of Chinese hospitals [Report in Chinese]. : China Hospital Information Management Association (CHIMA); 2023 URL: https://www.chima.org.cn/Sites/Uploaded/File/2024/08/20240821001.pdf [accessed 2026-08-06]
21. Medical informatization news weekly review issue 441 | Public announcement of the list of medical institutions passing level 5 and above with electronic medical records in 2023 [Article in Chinese]. China Hospital Information Management Association (CHIMA). 2024. URL: https://www.chima.org.cn/Html/News/Articles/16879.html [accessed 2026-08-06]
22. Malterud K. Qualitative research: standards, challenges, and guidelines. Lancet 2001 Aug 11;358(9280):483-488. [doi: 10.1016/S0140-6736(01)05627-6] [Medline: 11513933]
23. PLoS Medicine Editors. Qualitative research: understanding patients' needs and experiences. PLoS Med 2007 Aug;4(8):e258. [doi: 10.1371/journal.pmed.0040258] [Medline: 17760496]
24. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
25. Phillippi J, Lauderdale J. A guide to field notes for qualitative research: context and conversation. Qual Health Res 2018 Feb;28(3):381-388. [doi: 10.1177/1049732317697102] [Medline: 29298584]
26. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
27. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
28. Shenton AK. Strategies for ensuring trustworthiness in qualitative research projects. Educ Inf 2004;22(2):63-75. [doi: 10.3233/EFI-2004-22201]
29. Hamed MMM, Konstantinidis S. Barriers to incident reporting among nurses: a qualitative systematic review. West J Nurs Res 2022 May;44(5):506-523. [doi: 10.1177/0193945921999449] [Medline: 33729051]
30. Kahn MG, Callahan TJ, Barnard J, et al. A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. EGEMS (Wash DC) 2016;4(1):1244. [doi: 10.13063/2327-9214.1244] [Medline: 27713905]
31. Skyttberg N, Vicente J, Chen R, Blomqvist H, Koch S. How to improve vital sign data quality for use in clinical decision support systems? A qualitative study in nine Swedish emergency departments. BMC Med Inform Decis Mak 2016 Jun 4;16(61):61. [doi: 10.1186/s12911-016-0305-4] [Medline: 27260476]
32. Wilk M, Marsh DWR, De Freitas S, Prowle J. Predicting length of stay in hospital using electronic records available at the time of admission. Stud Health Technol Inform 2020 Jun 16;270:377-381. [doi: 10.3233/SHTI200186] [Medline: 32570410]
33. Lyons AM, Dimas J, Richardson SJ, Sward K. Assessing EHR data for use in clinical improvement and research. Am J Nurs 2022 Jun 1;122(6):32-41. [doi: 10.1097/01.NAJ.0000832728.09164.3f] [Medline: 35551125]
34. Jetley G, Zhang H. Electronic health records in IS research: quality issues, essential thresholds and remedial actions. Decis Support Syst 2019 Nov;126:113137. [doi: 10.1016/j.dss.2019.113137]
35. The ICN code of ethics for nurses. : International Council of Nurses (ICN); 2021 URL: https://www.icn.ch/sites/default/files/2024-12/ICN_Code-of-Ethics_EN_Web.pdf [accessed 2026-07-18]
36. Farag A, Blegen M, Gedney-Lose A, Lose D, Perkhounkova Y. Voluntary medication error reporting by ED nurses: examining the association with work environment and social capital. J Emerg Nurs 2017 May;43(3):246-254. [doi: 10.1016/j.jen.2016.10.015] [Medline: 28359712]
37. Makeleni N, Cilliers L. Critical success factors to improve data quality of electronic medical records in public healthcare institutions. S Afr J Inf Manag 2021;23(1). [doi: 10.4102/sajim.v23i1.1230]
38. Toftdahl AKS, Pape-Haugaard LB, Palsson TS, Villumsen M. Collect once—use many times: the research potential of low back pain patients' municipal electronic healthcare records. Stud Health Technol Inform 2018;247:211-215. [Medline: 29677953]
39. Lewis AE, Weiskopf N, Abrams ZB, et al. Electronic health record data quality assessment and tools: a systematic review. J Am Med Inform Assoc 2023 Sep 25;30(10):1730-1740. [doi: 10.1093/jamia/ocad120] [Medline: 37390812]
40. Wawrzyniak B. The interplay between data-related harm and the secondary use of health data. Medium. 2024. URL: https://medium.com/odi-research/the-interplay-between-data-related-harm-and-the-secondary-use-of-health-data-09ffd59f7cf0 [accessed 2026-07-18]
41. Burnum JF. The misinformation era: the fall of the medical record. Ann Intern Med 1989 Mar 15;110(6):482-484. [doi: 10.7326/0003-4819-110-6-482] [Medline: 2919852]
42. Pourasghar F, Malekafzali H, Koch S, Fors U. Factors influencing the quality of medical documentation when a paper-based medical records system is replaced with an electronic medical records system: an Iranian case study. Int J Technol Assess Health Care 2008;24(4):445-451. [doi: 10.1017/S0266462308080586] [Medline: 18828939]

XSL•FO
RenderX

43. Olson KD, Meeker D, Troup M, et al. Use of ambient AI scribes to reduce administrative burden and professional burnout. JAMA Netw Open 2025 Oct 1;8(10):e2534976. [doi: 10.1001/jamanetworkopen.2025.34976] [Medline: 41037268]
44. Guest G, Bunce A, Johnson L. How many interviews are enough? Field methods 2006 Feb;18(1):59-82. [doi: 10.1177/1525822X05279903]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
EMR: electronic medical record
ENR: electronic nursing record

Edited by E Borycki; submitted 30.Mar.2026; peer-reviewed by FA Dhabbari, M Tarabeih; revised version received 19.Jul.2026; accepted 28.Jul.2026; published 26.Aug.2026.

Please cite as:
Wang J, Xu Y, Li W, Yang Z, Sun Y, Zhang X, Pan H
Manifestations and Potential Consequences of Information Distortion in Electronic Nursing Records: Qualitative Study
JMIR Nursing 2026;9:e96679
URL: https://nursing.jmir.org/2026/1/e96679
doi:10.2196/96679

© Jianan Wang, Yihong Xu, Wen Li, Zhichao Yang, Yushu Sun, Xiaoxiao Zhang, Hongying Pan. Originally published in JMIR Nursing (https://nursing.jmir.org), 26.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Nurses' Expectations of a Knowledge Management System in Nursing Practice: Qualitative Study

Magdalena Vogt[1], MScPH; Sebastian Müller[2], MSc; Glorianna Wagner-Jagfeld[3], PhD; Renate Ranegger[3], PhD; Sabin Zürcher[4], MSCN; Janine Vetsch[1], Prof Dr

[1]Institute of Health Sciences, Eastern Switzerland University of Applied Sciences, Rosenbergerstrasse 59, St. Gallen, Switzerland

[2]Institute of Information and Process Management, Eastern Switzerland University of Applied Sciences, St. Gallen, Switzerland

[3]Department of Research and Development, LEP AG, St. Gallen, Switzerland

[4]Department of Nursing Development, Lindenhofgruppe Bern, Bern, Switzerland

Corresponding Author:
Magdalena Vogt, MScPH
Institute of Health Sciences, Eastern Switzerland University of Applied Sciences, Rosenbergerstrasse 59, St. Gallen, Switzerland

Abstract

Background: Evidence-based practice is essential for delivering safe, high-quality nursing care; however, its implementation remains challenging due to barriers such as limited knowledge, a lack of supportive organizational culture, and insufficient access to relevant knowledge at the point of care. Knowledge management systems (KMSs) have the potential to bridge this gap by integrating evidence into the nursing process through technological support. Despite growing interest, the integration of KMS into daily nursing practice is still underexplored, especially from the perspective of frontline nurses.

Objective: The aim of this study was to explore nurses' perspectives on the requirements for a KMS that supports evidence-based practice at the point of care, with a focus on usability, process integration into the electronic nursing care plan and patient chart, and implementation challenges and benefits.

Methods: A qualitative study was conducted in a Swiss hospital using observations, focus groups, and individual interviews with 6 registered nurses, 9 advanced practice nurses, 2 nursing managers, and 1 head physician. Data were analyzed using thematic analysis.

Results: The analysis revealed four main categories and ten subcategories: (1) content of the KMS, (2) personal and structural factors of knowledge management, (3) technical conditions of the KMS, and (4) implementation of a KMS. Participants emphasized the need for an intuitively structured, process-integrated system that links evidence-based information directly to nursing interventions in the electronic nursing care plan and patient chart. Organizational support, interprofessional collaboration, and clear responsibilities were identified as critical for successful implementation.

Conclusions: There is a clear need for a KMS that is user-friendly, seamlessly integrated into clinical workflows, and supports quick, reliable access to evidence-based knowledge. A KMS could enhance nurses' access to reliable knowledge, promote evidence-based decision-making, and strengthen professional confidence at the point of care. By embedding evidence directly into the electronic nursing care plan and patient chart, such systems can streamline workflows, reduce time spent searching for information, and support more consistent application of best practices. These capabilities may improve information retrieval and contribute to a safer, more consistent nursing practice.

(JMIR Nursing 2026;9:e78395) doi:10.2196/78395

KEYWORDS

evidence-based practice; knowledge management; knowledge management system; qualitative research; point of care; nursing

Introduction

Background

Delivering safe, high-quality patient care is a central goal of health care institutions [1] and evidence-based practice (EBP) plays a key role in achieving this [2]. Despite strong advocacy, the integration of scientific evidence into everyday nursing practice remains inconsistent [3]. Studies report that barriers such as insufficient EBP knowledge and skills, lack of mentors and facilitators, perceptions that EBP takes too much time, unsupportive organizational cultures, and environments hinder nurses from using evidence at the point of care [2,3]. At the same time, there is an exponential growth in the body of evidence-based knowledge, which needs to be accessed and integrated into daily nursing practices in a timely and contextually relevant manner [4].

To address these challenges, the concept of knowledge management, widely used in other industries, is gaining traction

XSL•FO
RenderX

in health care settings [5]. Knowledge management refers to programs or systems to create, capture, store, organize, and share knowledge and information effectively within organizations [6]. In health care settings, knowledge management has the potential to strengthen nursing performance [7] by facilitating access to both scientific knowledge and the expertise or practice knowledge of team members [5]. However, effective knowledge management in nursing practice requires more than just access; it requires integration into clinical workflows, supportive leadership, and a culture of continuous learning [6,8].

Knowledge management systems (KMSs), as a technological solution, offer a way to embed both evidence-based and practice-based knowledge directly into the nursing process [4]. KMSs are designed to support and enhance organizational processes for creating, storing, retrieving, transmitting, and applying knowledge [9]. When effectively designed and implemented, KMSs can support nurses in making informed decisions, promote EBP, and improve the quality of nursing care [9,10]. Despite this potential, research shows that such systems are rarely used in health care, especially in nursing contexts. To date, there are few descriptions of the development, implementation, and evaluation of KMSs in nursing practice [4]. There is a need to investigate factors on the adoption of a KMSs that are integrated into the nursing process in hospitals from different perspectives [11].

Prior Work and Research Gap

In a prior study, Ranegger et al [12] demonstrated the theoretical feasibility of linking evidence-based knowledge to standardized nursing interventions through a mapping project. While this work provided an essential foundation for embedding evidence in structured nursing documentation, it did not explore how such a system could meet the practical and contextual needs of nurses in clinical settings. Consequently, little is known about what nurses expect from a KMS, how they envision it supporting their workflow, and which organizational factors are required for successful implementation [12].

Aim of This Study

Building on this gap, our study focuses on advancing current research on KMSs in the health care sector by adding a user-centered perspective to support nurses at the point of care.

Therefore, the aim of this study was to qualitatively explore nurses' perspectives on the requirements for a KMS that supports EBP at the point of care, with a focus on usability, process integration into the electronic nursing care plan and patient chart, and implementation challenges and benefits. By identifying these requirements, this study contributes to the development of a KMS that is not only theoretically feasible but also contextually relevant, usable, and sustainable in clinical practice.

Methods

Study Design

An exploratory qualitative study design based on inductive thematic analysis was conducted to gain an in-depth understanding of nurses' perspectives, expectations, and experiences related to the development and implementation of a KMS to support EBP at the point of care. The study followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines to ensure methodological rigor and transparency [13]. The study was underpinned by a pragmatic theoretical orientation, which assumes that knowledge is constructed through experience and that research should focus on understanding real-world problems and generating practical solutions. This framework guided the exploration of nurses' expectations of a KMS, emphasizing the practical relevance of the findings for system design and implementation.

Researchers' Characteristics

Two researchers collected the data. The first researcher was a female research associate with expertise and training in nursing and health sciences. She holds a master's degree in public health, is specialized in EBP, and has worked as a nurse previously. The second researcher was a male research associate with a master's degree in information systems with research experience in digital health. The researchers were not known to the participants before the study. Participants were informed about the researchers' professional backgrounds, institutional affiliations, and the aim of the study. They also knew about the researchers' roles within the project and that participation was voluntary and anonymous. The researchers were aware that their professional backgrounds could influence how they collected and analyzed data. They therefore reflected these potential biases throughout the analysis to support a balanced understanding of the data.

Participants and Setting

The study was conducted in a hospital in Switzerland that is part of a private hospital group comprising 3 hospitals. The hospital group employs approximately 2500 staff and treats over 140,000 patients annually, including around 27,000 inpatients. At the time of the study, a new intranet was planned to centralize knowledge resources and improve search capabilities.

The study focused on nurses with diverse work experience and role profiles because the KMS was intended primarily for nursing practice. Additionally, 1 physician was included to provide an interprofessional perspective, as physicians are involved in the current system. Only 1 physician was included because the study primarily focused on nursing workflows and physician involvement in the planned KMS was limited during the recruitment period. A purposive sampling strategy was applied via the head of the nursing development department to recruit participants for observations, individual interviews, and focus group interviews. All participants were directly or indirectly involved in nursing-related knowledge management, either as users of or contributors to knowledge sources. Eligibility criteria included active involvement in nursing care or related managerial or educational functions within the hospital. The initial plan was to invite 15 to 20 participants, which was achieved, with 18 individuals confirming attendance.

XSL•FO
RenderX

Data Collection

Prior to data collection, an observation guide and a semistructured interview guide were developed based on a literature review. The researchers first conducted independently 4 hours of open, participatory observation on a ward in November 2023, focusing on the activities of 3 registered nurses during their shifts on a surgical and internal medicine ward and took field notes according to the observation guide. These 3 nurses were not further part of the interviews.

Subsequently, all interviews were conducted using the semistructured interview guide, which was adapted after the observations. The 2 authors held 2 face-to-face focus group interviews in the hospital, with participants grouped by professional hierarchy to encourage open discussion. The first focus group involved 9 advanced nurse practitioners (ANPs) with master's degrees from different wards in the participating hospital (92 min). The second focus group comprised 3 registered nurses with a diploma degree from the surgical ward to capture another perspective (50 min). Three additional online interviews using Microsoft Teams (30 min each) were held with a head physician, a division manager in nursing care, and a co-nursing manager to include different viewpoints. Sociodemographic data from all participants were collected verbally (Table 1). All interviews were conducted between November 2023 and January 2024 and were audio-recorded. Field notes were taken during the interviews. No repeated interviews were carried out.

Table . Sociodemographic characteristics (N=18).

Sociodemographic characteristics	Value
Sex, female, n (%)	18 (100)
Role, n (%)	
Advanced practice nurse	9 (50)
Registered nurse	6 (33.4)
Nursing management	2 (11.1)
Physician	1 (5.5)
Field of work, n (%)	
Surgical	8 (44.4)
Internal medicine	5 (27.8)
Other (eg, orthopedics and oncology)	3 (16.7)
Expert in a field (eg, delirium and breast care)	2 (11.1)
Years of working experience, n (%)	
<5	1 (5.5)
5-10	7 (38.9)
10-15	5 (27.8)
>15	5 (27.8)

Observation notes were translated, summarized, and thematically clustered. The single and focus group interviews were audio-recorded and transcribed by hand. Data analysis took place in parallel with data collection. Data saturation was considered achieved after the third online interview. Consistency between the 2 data collectors was ensured through continuous discussion during data collection and analysis to align interpretation and maintain reflexivity.

Data Management and Analysis

Thematic analysis was conducted following the 6-phase approach described by Braun and Clarke [14]. In addition to the categories already formed a priori through a literature review and by developing the semistructured interview guide, 1 author performed the initial coding of all transcripts using MAXQDA 2022 [15]. Coding decisions and theme development were subsequently discussed with 2 additional authors to ensure analytic consistency and to confirm the relevance of identified categories. Disagreements were resolved through discussion until consensus was achieved. The analysis resulted in main categories and subcategories, which were then translated from German to English.

Ethical Considerations

All participants were informed verbally about the purpose and procedures of the study, data confidentiality, and voluntary participation. Informed consent was obtained before participation, and withdrawal of consent was permitted at any stage, including after data collection. Audio recordings were transcribed verbatim, anonymized to remove any potentially identifiable information, and assigned participant codes before recordings were subsequently deleted. All data were stored securely on password-protected institutional servers in accordance with data protection regulations. No participants withdrew consent for the use of their data in this study. According to Swiss legislation, this study did not require approval by a cantonal ethics committee. In accordance with the Swiss Human Research Act (Humanforschungsgesetz, HFG),

XSL•FO
RenderX

ethical approval is mandatory only for research involving human participants where health-related personal data are collected or where interventions are performed [16]. The present study focused exclusively on healthcare professionals' perspectives on KMSs. No patients were involved, no health-related personal data were collected, and no interventions were performed. Therefore, the study does not fall within the scope of the Swiss Human Research Act and did not require formal ethical approval by a Swiss ethics committee.

However, the study was followed in accordance with the World Medical Association's Declaration of Helsinki.

Results

Sociodemographic Characteristics

All 18 participants were nurses with different degrees and roles, except 1 was a head physician. The participants from the observations and interviews had at least 1 year of professional experience and worked in different roles and fields in the hospital (Table 1).

Categories

The thematic analysis resulted in 4 main categories with 10 subcategories, each of which will be discussed in the following sections (Table 2).

Table . Main categories and subcategories of the thematic analysis.

Main categories	Subcategories
Content of KMS[a]	• Information sources • Format of information
Personal and structural factors of knowledge management	• Information retrieval skills • Time pressure and efficiency
Technical conditions of KMS	• Integration into workflow • Knowledge access and architecture
Implementation of a KMS	• Barriers • Facilitators • Expected benefits • Potential quality indicators

[a]KMS: knowledge management system.

Content of KMS

Information Sources

Participants described a clear distinction in information sources used by different roles. At the point of care, registered nurses primarily relied on in-house nursing instructions and team members, which was also observed.

In contrast, ANPs accessed a wider range of formal evidence sources, including databases, guidelines, professional networks, and conferences, which they used to update or develop new nursing instructions. Although digital advancements were mentioned, none of the participants reported using artificial intelligence (AI) tools in their knowledge work. Instead, maintaining clear, up-to-date, and evidence-based nursing instructions was viewed as a central way to ensure consistent practice. Most ANPs and nurses from focus groups would support the inclusion of brief synopses of studies explaining changes and evidence updates in the in-house nursing instructions. These would offer nurses an optional, deeper insight into the rationale behind changes. However, some ANPs and the co-nursing manager were critical of this and questioned whether nurses at the point of care would be using this due to the high workload and limited skills in scientific working.

Format of Information

Participants acknowledged that the current nursing instructions were logically structured and helpful, often featuring tables of contents and uniform formatting. Nurses were instructed to use nursing standards as the main source of information in nursing practice. At the time point of the interviews and observations, it was therefore important that the nursing standards were written in simple language and regularly updated according to the latest evidence.

Registered nurses and ANPs from both focus groups and observations expressed a need for varied formats as a source of information, such as checklists, videos, and schematics, as long as the content remained concise and practice-oriented. The information in the KMS should not be overloaded and it should summarize the most important information as briefly as possible, as an ANP said:

> *I think you have to be careful not to overload nurses with information, to be honest. You have to focus on what you really need in practice. The more it is broken down to the practical situation, the more the knowledge is used.* [P1]

Personal and Structural Factors of Knowledge Management

Information Retrieval Skills

Participants reported that while they were able to locate nursing instructions within their own specialty, accessing materials outside of their immediate practice area was often time-consuming and frustrating. Nurses, particularly those who

XSL•FO
RenderX

were new, part-time, or less experienced, struggled to find information when documents were not intuitively filed or when search paths were long and complex. A nurse confirmed this during the observation. Many participants noted that there was no systematic onboarding to teach information-seeking or navigation strategies. Although some suggested additional training, they emphasized that intuitive structure and powerful search functions were more impactful than teaching workarounds. An ANP summarized it as follows:

> *If the search function is poor, it doesn't matter how well you know the system. You still can't find what you need.* [P2]

Time Pressure and Efficiency

Time constraints were a significant concern in information use and acquisition across all participants. Nurses commonly relied on team members and ANPs to obtain information quickly, particularly during high workload periods. In the observations, the nurses asked more experienced nurses or a physician in some cases before searching available documents. All participants would find it helpful to have faster access to information sources at the point of care. These sources should be process integrated, which means embedded in the electronic nursing care plan and patient chart. An ANP said:

> *I often hear that nurses know that a certain nursing instruction exists. They still ask me as an ANP if I can't just tell them the answer quickly so that they do not have to search for the document.* [P1]

Technical Conditions of KMS

Integration Into Workflow

Participants envisioned a KMS integrated into every phase of the nursing process, from patient admission and assessment to diagnosis, intervention, and evaluation. They found it important that the information would be available and could be retrieved exactly when they needed it. The nurses from the focus groups saw the greatest benefit in linking information to nursing interventions, for example, to check how a central venous catheter needs to be connected. The head physician also recognized potential in areas like diagnosis support and medication information:

> *For example, if I select permanent catheters in the nursing care plan, the relevant nursing instruction should be stored there. If access to the information is clearly visible in the nursing care plan, my attention would be drawn to it and I can just click on it. And then the information just comes up. Because if it is not obvious and I don't see it, I won't click on it and won't get to the information. It has to be obvious to me.* [P5]

Knowledge Access and Architecture

All participants criticized the current dual document storage system, which resulted from an ongoing transition to a new intranet. Most participants found the folder structure confusing and the search function ineffective due to a lack of semantic features. Old or irrelevant documents still appeared in search results, adding to the inefficiency. An ANP mentioned:

> *With the folder system, for example, there are folders from the pharmacy, where I think there is a great need for training. Because sometimes you go to an instruction but do not realize that there is also something about [eg,] potassium substitution. And there would be very helpful practical [nursing] instructions. But [most nurses] do not know that they exist.* [P2]

Suggestions from the participants were to install links in the electronic nursing care plan and patient chart with direct access to information. Two ANPs had the idea to create question mark buttons or to provide the information when clicking on nursing interventions or diagnoses of the electronic nursing care plan and patient chart (eg, dressing a wound, assessing the risk of malnutrition, and administering a medication), which was supported by the other ANPs. Links to documents should always point to the latest version, avoiding discrepancies between sources. An additional idea from the interviewed head physician was to link medication prescriptions directly to the electronic nursing care plan and patient chart with instructions for administration. Additionally, powerful search functions and filter options to quickly find relevant information would be helpful for nurses. The goal from the interviewed division manager in nursing care would be a single-source approach where updated instructions were universally accessible. The division maker in nursing care, therefore, said:

> *It must be ensured that the latest version of the nursing instruction is available via the KMS. For example, if you open a link to the nursing instruction from the electronic nursing care plan, the revisions made should also be changed in this document […]. And if something is changed there, I always have the latest version, no matter where I access the document from.* [P7]

Implementation of a KMS

Barriers

Time, money, and personnel constraints were mentioned as the main barriers to the development and implementation of a KMS. The co-nursing manager stressed that the decision-maker of the hospital needs to be convinced of the KMS, as it requires financial investment. The head nurse emphasized that time and financial resources of the hospital must be used sparingly and that the benefits need to outweigh the costs. Additionally, the lack of clarity around responsibilities for integrating KMS content into hospital IT systems was problematic from the head nurses' perspective.

Facilitators

The ANPs saw themselves as responsible for content conceptualization within the KMS. They proposed that IT staff and KMS providers manage the structural and technical implementation. Strong interprofessional collaboration, clear role descriptions, and leadership support were emphasized as important, as an ANP said:

XSL•FO
RenderX

> *The conceptual aspect is for sure with us ANPs. Anything else would be inefficient. But we would not be unhappy if someone else takes care of linking the documents between KMS and the hospital information system.* [P8]

Expected Benefits

Participants believed the KMS would facilitate faster information retrieval, better alignment with current standards, and improved interdisciplinary collaboration. From the head nurse's point of view, this meant that knowledge in nursing could be better preserved and shared. The nurses were convinced that documents were more likely to be used if they were integrated into the nursing process and could be accessed quickly. This could also increase the nurses' sense of safety, as they would always use the correct and updated documents. Moreover, the responsibility for finding the right document would no longer lie with the nurses themselves, as a registered nurse said:

> *And I think it would be of particular benefit to patients, and that is an interprofessional interest. If the nursing staff can stand up afterwards and say, these are our instructions, we have to implement them. The better you know the content of the nursing instructions and the faster you find them, the better you can argue.* [P9]

Potential Quality Indicators

Participants proposed a range of indicators on how to measure the effectiveness of the KMS. The ANPs mentioned direct KMS-related indicators such as time to retrieve information (eg, reduced time to find nursing instructions), task-completion rate (eg, conducting a nursing intervention), need for help in terms of knowledge retrieval (eg, contacting ANP), and user satisfaction with the system. Indirect quality indicators could be downstream outcomes such as quality of care and patient safety. The nurses from the second focus group mentioned the nurses' subjective sense of security when performing nursing interventions as an additional indicator. The head physician and co-nursing manager particularly mentioned the quality of the intra- and interdisciplinary communication, including the perceived ease and frequency of collaboration as further quality indicators. The co-nursing manager said:

> *For me, relevant indicators are the satisfaction and nurses' sense of security in their daily work. The nurses need the information to provide the patient with adequate care.* [P6]

The participants emphasized that an effective KMS should directly support clinical decision-making and increase confidence during care delivery. Nurses frequently linked quick access to correct information with improved performance, lower stress levels, and better patient outcomes. The ANPs and registered nurses believed that evaluating the system's impact should go beyond technical metrics and include experiential factors, such as how secure, informed, and supported they felt while using the system. Furthermore, participants stressed that if a KMS was truly helpful, it would minimize the need for ad hoc knowledge-seeking from team members, reduce errors, and encourage standardized practice across wards. An ANP mentioned:

> *If the information is easy to access whenever they need it, the more they use this information. This, I guess, brings satisfaction because nurses do not have to search a long time for the information and this also indicates a higher sense of security because they know, where they find the information and are well informed.* [P3]

Discussion

Main Results

This study explored nurses' expectations and needs for a KMS integrated into the electronic nursing care plan and patient chart. Participants found the existing hospital information system fragmented and time-consuming. In-house nursing instructions were well-structured but difficult to access due to a confusing filing system and poor search functionality. Nurses often relied on colleagues or ANPs for quick answers, especially under time pressure. Nurses expressed a strong need for a KMS that was integrated into the electronic nursing care plan and patient chart. They envisioned context-sensitive information access, such as clickable links or icons, at each step of the nursing process, from assessment through intervention to evaluation. The system should offer a simplified structure, powerful search functions, and information presented in practical, user-friendly formats like checklists, videos, or brief summaries. To support safe and efficient care, nurses emphasized that information must be both easily retrievable and always up to date. They saw clarity about responsibilities for maintaining the system as essential. Ultimately, they imagined that a well-designed KMS would enhance care quality, streamline workflows, and strengthen nurses' professional confidence at the point of care.

Integrating Knowledge Into Clinical Workflow

Our results show that the current system does not adequately support quick and reliable access to nursing-relevant information at the point of care. Nurses reported relying on team members or navigating complex document systems, often under time pressure. This aligns with findings that emphasize the importance of integrating knowledge tools directly into clinical workflows to reduce search time and cognitive load [4]. Existing help buttons and intranet instructions were appreciated; however, they were not sufficient for efficient knowledge access during daily work. This underscores the importance of embedding knowledge directly into digital workflows. Chorney et al [9] recommended this because they found that integrating KMS into clinical systems significantly improved access and usage. Knowledge embedded in systems not only reduces variation of information and nursing interventions but also supports EBP, given that the content is reliable and up to date [17]. This resonates with the Technology Acceptance Model, which emphasizes perceived usefulness and ease of use as key predictors of usage [18]. The desire for an intuitively designed, workflow-integrated KMS illustrates that these dimensions are central to successful use and implementation.

XSL•FO
RenderX

Information Literacy and the Role of Training

Nurses described variability in their ability to retrieve and apply information, especially among new staff, part-time workers, or those returning from leave. This reflects a broader challenge of information and digital literacy in nursing practice. Training was seen as critical to ensuring consistent access to and use of available knowledge resources. These findings are consistent with earlier studies that show age and experience influence confidence with electronic clinical systems [10] and that tailored onboarding and continued training support more effective system use [17]. Moreover, nurses' literacy influences their attitudes towards and intentions to use KMS [19]. While technical solutions are necessary, they must be accompanied by accessible training formats and support structures to ensure equitable use across roles and experience levels [17].

Evidence Flow and the Role of ANPs

Our study revealed a distinct division of tasks around knowledge sources: while nurses primarily relied on in-house nursing instructions and team members, ANPs engaged with external evidence sources. This distinction reflects the layered process of knowledge use, translation, and transfer outlined by Shahmoradi et al [4]. ANPs acted as translators, adapting external evidence to the hospital's context, while nurses at the point of care used this adapted knowledge. In addition to the application of nursing instructions, information was also transferred via other communication channels such as direct exchange, emails, or newsletters. This confirms findings from Al-Busaidi [20], who emphasized that knowledge transfer in health care often depends on informal systems that are neither systematic nor easily evaluated.

KMS Quality, Functionality, and Usability

A consistent theme in the interviews was the desire for a system that was intuitive, accessible, and available throughout the nursing process. This aligns with previous findings that ease of access and integration into clinical routines are critical success factors for KMS adoption [9]. Participants suggested that its functionalities should include a logical filing system, powerful search capabilities, and support for multiple content formats. This reflects a need for information to be both concise and adaptable to diverse learning preferences [9].

The absence of AI use among participants in the period before and during data collection in 2023 and 2024 also reflects broader hesitations in clinical environments. While AI integration was not expected by participants, its future role in enhancing clinical KMS remains a promising area for development [4]. Regardless of the technology used, the success of the KMS depends on its ability to fit seamlessly into the existing workflow and meet users' needs for quick and trustworthy information [20].

Evaluation and Trust in the System

Participants proposed a range of indicators to evaluate a future KMS, including efficiency gains, time savings, and perceived improvements in quality of care. These are consistent with indicators described by Al-Busaidi [20], who emphasized both organizational and individual-level outcomes such as improved learning, collaboration, and job satisfaction. Nurses in the interviews also framed evaluation in terms of emotional and ethical relief, particularly the idea that linked and validated instructions could reduce their burden of manually searching the "right" document. This does not imply a reduction in professional responsibility but highlights how a well-maintained KMS can support nurses in fulfilling their responsibilities more safely and confidently. This emotional dimension adds a new perspective in understanding trust in digital systems. Trust is shaped not only by technical reliability but also by how systems redistribute responsibility and reduce the risk of error [11]. When knowledge is institutionalized within a centrally maintained KMS, nurses can rely on the organization rather than the individual for ensuring accuracy. This shift reflects a rebalancing of cognitive and ethical responsibility, which can enhance professional confidence and perceived safety in clinical decision-making [21,22].

Organizational Conditions for Success

From the point of view of the head physician, division manager in nursing care, and co-nursing manager interviewed, organizational support emerged as an important factor for KMS success. They highlighted the need for leadership support, funding, and clear roles. These themes are confirmed across multiple studies, which identify infrastructure, staffing, policy support, and leadership engagement as critical to implementation success [1,20]. The findings also align with the Normalization Process Theory, which highlights the processes through which new interventions become embedded in everyday practice [23]. The constructs of shared understanding, cognitive participation, and practical integration are evident in participants' emphasis on collaboration and institutional backing. Organizational culture also plays a key role, as collaborative and open cultures have been found to facilitate KMS adoption more effectively than hierarchical, profit-driven environments [11]. Interviewed ANPs acknowledged the potential value of a KMS, particularly in terms of reducing redundant work and saving time. These findings support the statement from Chorney et al [9], that the success of a KMS is not only a technical or clinical matter, but also a strategic one. For sustainable implementation, the system must align with institutional priorities, demonstrate clear value, and receive long-term support from decision-makers in the setting [20].

Limitations

The main strength of the study was the inclusion of nurses with different levels of work experience and role profiles. This approach allowed for the consideration of multiple perspectives in the implementation of a KMS that was grounded in practical nursing requirements. The interdisciplinary research team balanced clinical and technical expertise but acknowledged that professional backgrounds might have influenced interpretation. Reflexivity was maintained through ongoing discussion to ensure balanced representation of participants' views. Conducting the study in a single hospital allowed for detailed observation of local workflows and knowledge management practices but limits the transferability of findings to other settings with different structures or digital maturity. Another limitation concerns the conceptual nature of the topic, as the study explored expectations for a KMS that has not yet been developed. Finally, the data were collected in German and then

XSL•FO
RenderX

translated into English. These translations were rigorously checked by authors fluent in both languages.

Implications for Nursing Practice and Research

Our findings underscore the importance of designing a KMS that supports nurses' real-time information needs at the point of care. Seamless integration into the electronic nursing care plan and patient chart, intuitive navigation, and access to up-to-date, evidence-based instructions in various formats were seen as essential. Nurse managers should prioritize training, onboarding processes, and continuous support, especially for new, part-time, or returning staff.

There is a need for further research on the design and usability of KMS tools, especially those that leverage emerging technologies such as AI for knowledge synthesis and decision support. Future studies should also explore the implementation and effects of KMS at the point of care. Further investigation into the quality indicators identified by nurses for measuring KMS impact could support the development of validated evaluation frameworks. Future projects from the authors focus on developing and piloting an AI-supported KMS. It aims to provide personalized, evidence-based recommendations tailored to nurses' skill levels and workflows, thereby enhancing safety, quality, and efficiency at the point of care.

Conclusions

Participants expressed a clear need for a KMS that is user-friendly, seamlessly integrated into clinical workflows, and supports quick, reliable access to evidence-based knowledge. A well-designed KMS may have the potential to not only improve care quality and efficiency but also to enhance nurses' confidence and sense of safety in their daily work.

Acknowledgments

The authors thank all the participants who took part in the interviews, focus groups, and observations.

Funding

This study was funded by an Innovationsscheck supported by Innosuisse (INNO-SBM Nr. 66971.1) and by LEP AG, for which we would like to express our thanks. The funder played no role in the study design, data collection, analysis, interpretation, or writing of the report.

Authors' Contributions

Conceptualization: RR, JV (equal), SZ (supporting)
Data curation: MV, JV (equal)
Formal analysis: MV
Funding acquisition: RR, JV (equal)
Methodology: MV, SM, GW-J, RR, JV (equal)
Project administration: JV
Writing – original draft: MV
Writing – review & editing: SM, GW-J, RR, SZ, JV (equal)

Conflicts of Interest

The authors declare that they have no competing interests. LEP provides health care intervention classifications but no knowledge management system and therefore has no competing interests in this study.

References

1. Karamitri I, Talias MA, Bellali T. Knowledge management practices in healthcare settings: a systematic review. Int J Health Plann Manage 2017 Jan;32(1):4-18. [doi: 10.1002/hpm.2303] [Medline: 26153388]
2. Connor L, Dean J, McNett M, et al. Evidence-based practice improves patient outcomes and healthcare system return on investment: findings from a scoping review. Worldviews Evid Based Nurs 2023 Feb;20(1):6-15. [doi: 10.1111/wvn.12621] [Medline: 36751881]
3. Melnyk BM, Fineout-Overholt E, Gallagher-Ford L, Kaplan L. The state of evidence-based practice in US nurses: critical implications for nurse leaders and educators. J Nurs Adm 2012 Sep;42(9):410-417. [doi: 10.1097/NNA.0b013e3182664e0a] [Medline: 22922750]
4. Shahmoradi L, Safadari R, Jimma W. Knowledge management implementation and the tools utilized in healthcare for evidence-based decision making: a systematic review. Ethiop J Health Sci 2017 Sep;27(5):541-558. [doi: 10.4314/ejhs.v27i5.13] [Medline: 29217960]
5. Park M, Jang I, Lim Kim S, et al. Evaluating the performance of an integrated evidence-based nursing knowledge management (I-EBNKM) platform in real-world clinical environments. Int J Med Inform 2023 Nov;179:105239. [doi: 10.1016/j.ijmedinf.2023.105239] [Medline: 37783190]

XSL•FO
RenderX

6. Knowledge management is key to public health planning, new study shows 2023. World Health Organization. URL: https://www.who.int/azerbaijan/news/item/13-11-2023-knowledge-management-is-key-to-public-health-planning--new-study-shows [accessed 2024-11-19]
7. Lee EJ, Kim HS, Kim HY. Relationships between core factors of knowledge management in hospital nursing organisations and outcomes of nursing performance. J Clin Nurs 2014 Dec;23:3513-3524. [doi: 10.1111/jocn.12603] [Medline: 24698408]
8. Lunden A, Teräs M, Kvist T, Häggman-Laitila A. A systematic review of factors influencing knowledge management and the nurse leaders' role. J Nurs Manag 2017 Sep;25(6):407-420. [doi: 10.1111/jonm.12478] [Medline: 28580645]
9. Chorney J, Johnson Emberly D, Jeffrey J, et al. Implementation of a knowledge management system in mental health and addictions: mixed methods case study. JMIR Form Res 2023 Feb 6;7:e39334. [doi: 10.2196/39334] [Medline: 36745489]
10. Hamari L, Parisod H, Siltanen H, et al. Clinical decision support in promoting evidence-based nursing in primary healthcare: a cross-sectional study in Finland. JBI Evid Implement 2023 Sep 1;21(3):294-300. [doi: 10.1097/XEB.0000000000000375] [Medline: 37102429]
11. Yun EK. Predictors of attitude and intention to use knowledge management system among Korean nurses. Nurse Educ Today 2013 Dec;33(12):1477-1481. [doi: 10.1016/j.nedt.2013.05.018] [Medline: 23806194]
12. Ranegger R, Haug S, Vetsch J, Baumberger D, Bürgin R. Providing evidence-based knowledge on nursing interventions at the point of care: findings from a mapping project. BMC Med Inform Decis Mak 2022 Nov 28;22(1):308. [doi: 10.1186/s12911-022-02053-8] [Medline: 36437450]
13. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
14. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
15. MAXQDA 2022 online manual 2021. maxqda.com. URL: https://www.maxqda.com [accessed 2025-02-15]
16. Humanforschungsgesetz (HFG) – federal act on research involving human beings (SR 810.30). Swiss Federal Law Compilation. URL: https://www.fedlex.admin.ch/eli/cc/2013/617/de [accessed 2026-01-20]
17. El-Jardali F, Bou-Karroum L, Hilal N, et al. Knowledge management tools and mechanisms for evidence-informed decision-making in the WHO European Region: a scoping review. Health Res Policy Syst 2023 Oct 31;21(1):113. [doi: 10.1186/s12961-023-01058-7] [Medline: 37907919]
18. Lee AT, Ramasamy RK, Subbarao A. Understanding psychosocial barriers to healthcare technology adoption: a review of TAM technology acceptance model and unified theory of acceptance and use of technology and UTAUT frameworks. Healthcare (Basel) 2025 Jan 27;13(3):250. [doi: 10.3390/healthcare13030250] [Medline: 39942440]
19. Diya RT, Ghezzi A. Digital documentation in healthcare: empowering nurses and patients for optimal care. Stud Health Technol Inform 2024 Jul 24;315:47-51. [doi: 10.3233/SHTI240104] [Medline: 39049224]
20. Al-Busaidi KA. Knowledge workers' perceptions of potential benefits and challenges of inter-organizational knowledge sharing systems: a Delphi study in the health sector. Knowl Manag Res Pract 2014 Nov;12(4):398-408. [doi: 10.1057/kmrp.2013.4]
21. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
22. Jones SS, Rudin RS, Perry T, Shekelle PG. Health information technology: an updated systematic review with a focus on meaningful use. Ann Intern Med 2014 Jan 7;160(1):48-54. [doi: 10.7326/M13-1531] [Medline: 24573664]
23. Murray E, Treweek S, Pope C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. BMC Med 2010 Oct 20;8:63. [doi: 10.1186/1741-7015-8-63] [Medline: 20961442]

Abbreviations

AI: artificial intelligence
ANP: advanced nurse practitioner
COREQ: Consolidated Criteria for Reporting Qualitative Research
EBP: evidence-based practice
KMS: knowledge management system

XSL•FO
RenderX

Edited by E Borycki; submitted 02.Jun.2025; peer-reviewed by A Nickum, CA Rodriguez-Suarez; revised version received 25.Nov.2025; accepted 06.Dec.2025; published 21.Jan.2026.

Please cite as:
Vogt M, Müller S, Wagner-Jagfeld G, Ranegger R, Zürcher S, Vetsch J
Nurses' Expectations of a Knowledge Management System in Nursing Practice: Qualitative Study
JMIR Nursing 2026;9:e78395
URL: https://nursing.jmir.org/2026/1/e78395
doi:10.2196/78395

© Magdalena Vogt, Sebastian Müller, Glorianna Wagner-Jagfeld, Renate Ranegger, Sabin Zürcher, Janine Vetsch. Originally published in JMIR Nursing (https://nursing.jmir.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Voice-Based Structured Nursing Documentation Using Automatic Speech Recognition and Large Language Models: Development and Evaluation Study

Meng-Han Su[1], MS; Wei-Chun Wang[1], MD; Yi-Min Hsu[2], PhD; Shih-Yen Hou[1], MS; Su-Jung Chuang[2], MHA; Shih-Sheng Chang[1], MD, PhD

[1]Artificial Intelligence and Robotics Innovation Center, China Medical University Hospital, China Medical University, No. 2, Yude Rd, North Dist, Taichung, Taiwan

[2]China Medical University Hospital, Taichung, Taiwan

Corresponding Author:
Wei-Chun Wang, MD
Artificial Intelligence and Robotics Innovation Center, China Medical University Hospital, China Medical University, No. 2, Yude Rd, North Dist, Taichung, Taiwan

Abstract

Background: For clinical nurses, manually entering information into hospital information systems (HISs) remains time-consuming and prone to omissions. Although speech recognition can reduce the need for manual entry, its use in clinical settings has historically been limited by code-switching, medical terminology, and noisy ward environments. Recent advances in customized automatic speech recognition (ASR) and large language models (LLMs) now make speech-based, structured documentation aligned with nursing frameworks such as DART (data, action, response, and teaching) increasingly feasible.

Objective: This study developed and evaluated an integrated ASR and LLM system that transforms spoken nursing input into structured DART notes and evaluated its accuracy, usability, and clinical feasibility within HIS workflows.

Methods: A code-switching nursing speech corpus from emergency and ward settings was used to fine-tune the Whisper large-v2 model with parameter-efficient adaptation. The LLM generated schema-constrained DART records from ASR transcripts, which were verified by nurses before being uploaded to the corresponding HIS fields. Evaluation included mixed error rate for ASR accuracy, F_1-scores, and agreement statistics for DART classification, hallucination assessments based on factual correctness, and analysis of nurse feedback on system use.

Results: The fine-tuned ASR model reduced the mixed error rate from 44.79% to 6.67%. DART generation achieved a macroaveraged F_1-score of 0.82 (95% CI 0.80 - 0.84) and met the noninferiority margin relative to human transcripts (δ=−0.04). The hallucination rate was 2.51%. During deployment, the monthly volume of valid nursing notes generated through voluntary use of the ASR system increased from 32,724 to 65,417, where each note represented a single documentation entry generated per patient care episode. Among 120 participating nurses, 91 (75.8%) reported reduced workload and improved completeness.

Conclusions: The integrated ASR and LLM system was feasible and showed strong performance, with good acceptance among clinical nurses. It reduced the manual documentation burden, improved record completeness, and demonstrated the value of an ASR- and LLM-supported workflow for nursing documentation.

(JMIR Nursing 2026;9:e88567) doi:10.2196/88567

KEYWORDS

automatic speech recognition; code-switching; large language model; nursing; documentation; nursing records

Introduction

Taiwan has faced a sustained shortage of clinical nurses, driven largely by low retention among licensed professionals rather than limited training capacity. The national nursing practice rate—representing the percentage of licensed nurses actively working in the profession—remains at only 59.1% [1]. This low retention of qualified staff exacerbates the clinical shortage, and excessive workloads and long shifts are strongly associated with burnout and turnover intention [2-4]. Documentation occupies a substantial portion of clinical workflow time, and existing digital systems have not fully reduced the need for manual input. As hospitals have expanded the use of hospital information systems (HISs), additional documentation requirements have emerged, prompting efforts to standardize formats and improve system integration [5]. Structured and interoperable records support continuity of care, and structured handoff protocols improve communication quality [6,7].

XSL•FO
RenderX

Moreover, higher perceived nursing information system quality is linked to greater use, satisfaction, and retention, while electronic record use embedded in routine workflows reduces documentation workload and intention to leave [8,9]. These findings highlight the need for documentation tools that are structured, interoperable, and well-integrated into daily clinical practice.

Many hospitals have introduced structured documentation frameworks to improve consistency and communication. Among these, the Focus Charting method has been particularly influential in Taiwan. Originating in the late 1980s [10] and adopted locally in the early 1990s [11], it organizes each nursing note around a defined patient focus and follows the DART (data, action, response, and teaching) pattern to support concise and standardized recording of observations, interventions, patient responses, and teaching activities. Implementation studies have reported clearer note organization, improved execution of care plans, and enhanced interdisciplinary collaboration in intensive care settings [12]. However, structured formats alone do not resolve the documentation burden. Even with electronic templates, nurses must still enter most information manually, and flowsheets—standardized tabular records used for tracking routine, time-sequenced clinical parameters such as vital signs—may occupy a substantial portion of each shift [13]. These limitations highlight the need for intelligent support that preserves the clarity of DART records while reducing repetitive manual effort.

Automatic speech recognition (ASR) has been explored as a way to reduce manual documentation effort, with multisite evaluations reporting higher efficiency and satisfaction [14,15]. In Taiwan, a longitudinal study involving 21 nurses (using a corpus of 30,112 words) found that mean accuracy improved from 87.06% to 95.07% across 4 evaluation sessions as users developed more stable speaking speeds and volumes [16]. Furthermore, an ASR system designed for code-switching—the practice of alternating between languages, such as Mandarin and English, within a single utterance—achieved a word error rate (a standard metric representing the percentage of transcription errors) of 12.3% in intensive care settings [17].

Recent reviews suggest that clinical speech technologies are rapidly evolving toward large language model (LLM)–driven ambient clinical intelligence, with mature commercial solutions (such as Heidi, Tortus, and Rebrief.ai [Rebrief Inc]) being successfully deployed in various settings to provide summarization, repurposing, and assistive autonomy [18]. However, although these ambient scribes excel in standard clinical encounters, their direct application to Taiwanese nursing workflows presents significant challenges. First, commercial systems are primarily optimized for monolingual or standard bilingual speech and often struggle with the dense, specialized Mandarin-English code-switching and local abbreviations prevalent in Taiwanese hospital wards. Second, existing tools typically generate generalized clinical summaries (eg, subjective, objective, assessment, and plan notes), lacking the fine-grained, schema-constrained precision required for specialized nursing frameworks such as DART and its direct interoperability with local HISs. Therefore, a bespoke ASR and LLM pipeline is necessary to address these specific linguistic and structural demands. To bridge this gap, this study developed and evaluated an integrated voice-to-report system that combines a domain-adapted ASR model with an LLM to generate structured DART nursing documentation aligned with HIS fields. The system was assessed using transcription accuracy, DART classification performance, and noninferiority testing, and its feasibility and workflow impact were examined during real-world deployment.

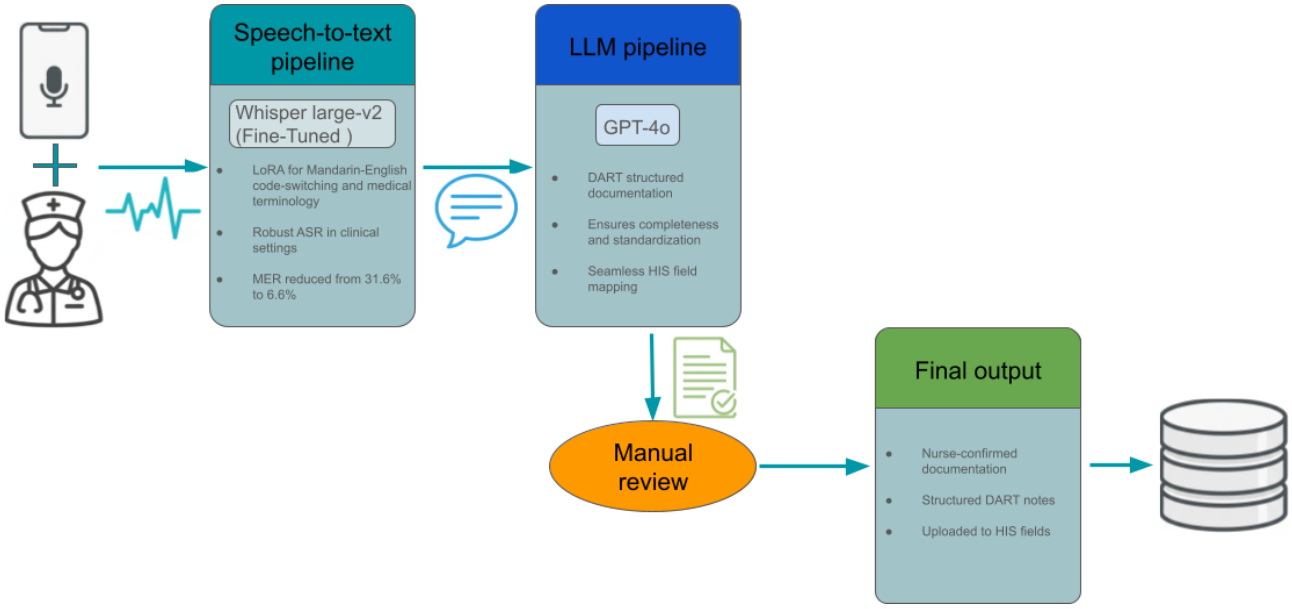
Methods

Study Design and System Overview

This development and evaluation study implemented a mobile-based nursing documentation system designed to convert voice input into structured clinical records (Figure 1). The system was deployed at China Medical University Hospital (CMUH) to evaluate its feasibility, transcription accuracy, and impact on documentation workflow. During routine care, nurses recorded observations and interventions using a mobile device. Audio input was processed by a domain-adapted Whisper ASR model trained on a Mandarin-English code-switching corpus and adapted to the medical terminology common in nursing communication. Transcribed text was subsequently processed by GPT-4o (OpenAI) using a schema-constrained prompt to generate standardized DART notes primarily in Mandarin while preserving English medical terminology. Nurses reviewed the generated content and triggered its upload to predefined HIS fields.

XSL•FO
RenderX

Figure 1. Workflow of the voice-to-DART (data, action, response, and teaching) documentation system. ASR: automatic speech recognition; HIS: hospital information system; LLM: large language model; LoRA: low-rank adaptation; MER: mixed error rate.



The system was developed in-house by the Artificial Intelligence and Robotics Innovation Center at CMUH. The codebase and intellectual property are proprietary. Due to strict patient privacy regulations and direct HIS integration, the software and raw datasets are not publicly available. However, to ensure methodological reproducibility, all model hyperparameters, evaluation frameworks, and complete schema-constrained prompt templates are fully detailed in this paper and Multimedia Appendix 1.

Dataset and Ground-Truth Preparation

The ASR model was trained and evaluated on nursing speech data from 3 complementary sources. The primary corpus, CMaiSpeech, comprised spontaneous Mandarin-English code-switched recordings from 525 nurses across multiple clinical units, characterized by frequent English medical terms and abbreviations. Table 1 summarizes the counts and durations of CMaiSpeech by category. Two synthetic corpora (approximately 9 hours) were generated using neural text-to-speech models from 1838 drug names and 1566 clinical terms to enrich underrepresented medical vocabulary. A real-world clinical dataset (totaling approximately 6 hours, 1608 utterances) captured authentic ward conversations, including vital-sign reporting and handovers. All recordings were normalized and resampled to 16-kHz pulse code modulation to ensure acoustic consistency. These datasets collectively served as the domain-adapted speech material used for training and evaluating the Whisper-based ASR model in nursing contexts.

In addition, a validation dataset of 327 nurse-authored documentation samples was constructed. Each sample included an audio recording collected during routine care, a manually transcribed reference serving as the ASR gold standard, and structured DART annotations verified by clinical nurses. This dataset was used to evaluate ASR transcription accuracy, LLM-based DART structuring, and agreement between fine-tuned and zero-shot conditions under a consistent reference standard.

Table . Counts and audio durations by category in the CMaiSpeech dataset.

Category	Item count (N=7055), n (%)	Audio length (seconds)[a]
Diastolic BP[b]	20 (0.28)	36.57
Glasgow Coma Scale eye response component	20 (0.28)	99.61
Pain index	20 (0.28)	46.05
Pulse	20 (0.28)	34.27
Systolic BP	20 (0.28)	39.34
Body temperature	20 (0.28)	39.10
Vital signs[c]	1035 (14.67)	12,350.10
Nursing notes[c]	5900 (83.65)	75,149.10

[a]Total duration of all audio samples within each category.

[b]BP: blood pressure.

[c]Categories representing full-length clinical recordings captured during routine nursing documentation activities.

XSL•FO
RenderX

ASR Model Adaptation and Evaluation

Whisper is a multilingual Transformer-based ASR model pretrained on 680,000 hours of diverse speech data [19]. Although its large-scale pretraining supports robust zero-shot generalization, accuracy declines in nursing-specific communication, where frequent Mandarin-English code-switching, dense medical terminology, and numerous drug abbreviations remain underrepresented. These discrepancies limit its direct applicability to structured nursing documentation.

Prior studies have demonstrated that domain-targeted fine-tuning markedly improves Whisper's performance in medical and mixed-language contexts. Previous studies reported gains in ASR and named entity recognition on Mandarin speech [20], and further improvements were achieved using domain-adapted fine-tuning and prompting on Mandarin-English medical data [21]. Building on these findings, Whisper (large-v2; OpenAI) was adapted to the nursing domain through low-rank adaptation–based parameter-efficient fine-tuning, which introduced trainable low-rank matrices into frozen weights to reduce computational cost while preserving performance [22]. The model was trained on a domain-specific nursing speech corpus and evaluated using mixed error rate (MER) to assess bilingual transcription accuracy relative to the baseline model.

Prior to MER calculation, a rigorous text normalization pipeline was applied to both reference and hypothesis transcripts to ensure consistent scoring. All full-width alphanumeric characters were converted to half-width equivalents, and punctuation marks were removed. Original casing for English text and medical abbreviations (eg, "SpO_2" and "BP") was strictly preserved to evaluate the model's ability to output correctly capitalized clinical terms. Crucially, numeric values, including decimals, were evaluated as independent, contiguous tokens rather than as split characters to accurately reflect their semantic weight in clinical parameters such as vital signs.

LLM Integration and Evaluation

The LLM was integrated to convert ASR transcripts into structured nursing documentation following the DART schema. GPT-4o (accessed via Azure OpenAI, application programming interface version 2024-12-01-preview), a multimodal variant of the GPT-4o family optimized for instruction following and multilingual text generation, was used for schema-based text structuring [23]. The model hyperparameters were consistently set to a temperature of 0.7, a top_p of 0.95, and a maximum of 16,000 tokens. A schema-constrained prompting strategy was applied to ensure consistent field separation and syntactic completeness, informed by prior evidence that structured prompting improves validity in clinical text generation [24]. The generation process used a dual-prompt structure: a system prompt defining the explicit task instructions and schema constraints, and a user prompt containing the ASR transcript. Two prompting configurations (a minimal version and a schema-constrained version) were compared during preliminary testing, and the schema-constrained configuration was selected for deployment. The full system prompt template is provided in Multimedia Appendix 1.

Model outputs were postprocessed to verify structural integrity and mapped to predefined HIS fields. LLM performance was evaluated on the 327-case validation dataset using field-level and macroaveraged F_1-score values across DART categories. The 95% CIs were estimated using 1000 paired bootstrap resamples at the case level. Noninferiority was tested using a predefined margin of δ=0.05, with noninferiority established when the lower bound of the ΔF_1-score CI exceeded −0.05. This 5% margin aligns with recent clinical evaluations of LLM-generated medical documentation, where a <5% variance is considered clinically acceptable for draft generation [25], and serves as a standard threshold in medical artificial intelligence performance evaluations [26]. From a clinical safety perspective, because the system captures data directly at the bedside via mobile devices, it inherently mitigates the high, undocumented risk of human memory decay and omission associated with traditional delayed documentation at the nursing station. Furthermore, because the system strictly requires human-in-the-loop verification prior to HIS submission, this minor structural variance during the drafting stage was deemed an acceptable trade-off for reduced documentation burden, while human oversight securely mitigates the risk of high-severity clinical errors.

Agreement Method

Three input conditions were compared while holding the structuring LLM constant: (1) human transcripts (reference), (2) Whisper large-v2 (zero-shot), and (3) Whisper large-v2 (fine-tuned). The primary end point was the F_1-score for DART slot classification against adjudicated references. A predefined noninferiority margin of δ=0.05 for ΔF_1-score was applied, and 95% CIs were estimated using paired bootstrap resampling. Noninferiority was established when the lower bound of the ΔF_1-score CI exceeded −0.05.

Field-level consistency was assessed because the "data," "action," "response," and "teaching" components of the DART framework represent distinct documentation intents. Labels were binarized (present-or-absent) after whitespace trimming, and both percent agreement and Cohen κ were calculated to adjust for chance agreement under skewed prevalence distributions [27].

It should be noted that the binarized F_1-score and agreement metrics were specifically chosen to evaluate the system's capability in initial structural triaging, ensuring that the LLM correctly maps spoken intents to the corresponding DART fields. Because the system operates within a strict human-in-the-loop workflow in which nurses must review and adjust the drafted text prior to HIS submission, achieving high structural accuracy significantly reduces manual documentation burden, even if minor within-field reallocations are occasionally required. To address content accuracy and penalize clinically incorrect information within these fields, the FactualCorrectness metric was used separately during the hallucination assessment.

Hallucination Assessment

Hallucination was defined as any generated content that lacked support from the corresponding reference transcript for the same case. Evaluation followed the FactualCorrectness metric from

XSL•FO
RenderX

the Retrieval-Augmented Generation Assessment (RAGAS) framework (precision mode), which decomposes model output into atomic claims and verifies each claim against the reference transcript [28,29]. A fixed Azure OpenAI model (GPT-4o, application programming interface version 2024-12-01-preview; temperature=0.7, top_p=0.95) was used as the evaluator via a LangChain wrapper to ensure consistent claim-level scoring. For each case, the 4 DART fields were concatenated into a single response, and both generated and reference texts were normalized prior to evaluation.

Mean factual-correctness precision and its complement (hallucination rate) were reported, along with the proportion of notes exhibiting hallucination under strict (<1.00) and relaxed (<0.95) thresholds. All computations were performed in Python (version 3.12; Python Software Foundation) using the RAGAS library [30].

System Use and User Feedback

To evaluate real-world feasibility and system adoption, use metrics were monitored during the initial deployment period at CMUH. Monthly use was quantified by extracting the total number of system-generated records from the application logs between March 2025 and August 2025.

Additionally, user feedback was collected to assess system acceptability and its impact on clinical workflow. Nursing staff from diverse clinical units, including wards and intensive care departments, voluntarily provided evaluations after integrating the system into their routine practice. The evaluation mechanism captured both quantitative satisfaction ratings (categorized as favorable or dissatisfied) and qualitative comments regarding user experience. Descriptive statistics were used to summarize the quantitative use and satisfaction data, while qualitative feedback underwent thematic analysis to identify common themes related to documentation efficiency, manual workload, and system accuracy.

Ethical Considerations

This study was approved by the Institutional Review Board of CMUH (CMUH110-REC2-181 and CMUH110-REC2-187). Since this was a retrospective study utilizing deidentified data, the requirement for informed consent was waived by the institutional review board. Patient privacy and data confidentiality were strictly maintained throughout the study, and no compensation was involved.

Results

Evaluation Dataset Characteristics

The evaluation dataset comprised 327 annotated nursing documentation samples, as described in the Methods section. Across transcripts, the corpus contained 12,136 Chinese characters, 1361 English words, and 1130 numeric tokens, equivalent to approximately 11.2 English words per 100 Chinese characters (Table 2, panel A). A total of 540 DART annotations were identified: 257 (47.6%) for “data,” 123 (22.8%) for “action,” 123 (22.8%) for “response,” and 37 (6.9%) for “teaching” (Table 2, panel B). The lower frequency of “teaching” entries reflects routine documentation patterns and was accounted for when estimating field-specific CIs. This dataset served as the common benchmark for evaluating both ASR and LLM outputs.

Table . Token composition and DART (data, action, response, and teaching) annotation distribution in the evaluation dataset.

Panels and categories	Distribution, n (%)	Count per sample[b], mean (SD)
Panel A: token[a] composition (n=14,627)		
Chinese characters	12,136 (83.0)	37.11 (24.31)
English words	1361 (9.3)	4.16 (5.56)
Numeric tokens	1130 (7.7)	3.46 (4.97)
Panel B: DART annotation distribution (n=540)		
Data	257 (47.6)	0.79 (0.41)
Action	123 (22.8)	0.38 (0.49)
Response	123 (22.8)	0.38 (0.49)
Teaching	37 (6.8)	0.11 (0.32)

[a]Token counts were computed from manual transcripts after text normalization; Chinese tokens represent individual characters, English tokens represent space-delimited words, and numeric tokens represent digit sequences.

[b]Values were calculated by dividing counts by the total number of samples (n=327).

ASR Model Performance

Whisper large-v2 was evaluated using MER, computed at the character level for Chinese and at the word level for English. In the zero-shot condition, MER was 44.79%. After low-rank adaptation–based fine-tuning on the nursing corpus, MER decreased to 6.67%, corresponding to an 85.11% relative error reduction (Table 3).

XSL•FO
RenderX

Table . Performance of Whisper large-v2 on the evaluation dataset.

Models	Mixed error rate (%)	Relative error reduction (%)
Zero-shot model[a]	44.79	—[b]
Fine-tuned model[c]	6.67	85.11

[a]Zero-shot: Whisper-Large-v2 inference without domain adaptation.

[b]Not applicable.

[c]Fine-tuned: Whisper-Large-v2 adapted using low-rank adaptation LoRA on the nursing speech corpus.

DART Classification Performance

DART classification performance was evaluated under 3 input conditions: human-transcribed references, fine-tuned ASR, and zero-shot ASR (Table 4). A schema-constrained prompting strategy was used for all analyses after preliminary comparison showed higher slot-level completeness and F_1-score performance than a minimal prompt.

Table . Per-field DART (data, action, response, and teaching) F_1-score performance with 95% CIs under 3 input conditions.

Fields	Zero-shot[a], F_1-score (95% CI)	Fine-tuned[b], F_1-score (95% CI)	Human transcript[c], F_1-score (95% CI)
Data	0.89 (0.85 - 0.92)	0.90 (0.87 - 0.93)	0.91 (0.89 - 0.94)
Action	0.59 (0.53 - 0.66)	0.77 (0.73 - 0.81)	0.83 (0.79 - 0.86)
Response	0.72 (0.67 - 0.77)	0.82 (0.79 - 0.85)	0.85 (0.82 - 0.88)
Teaching	0.62 (0.56 - 0.68)	0.78 (0.73 - 0.82)	0.83 (0.79 - 0.86)
Macroaverage[d]	0.70 (0.68 - 0.72)	0.82 (0.80 - 0.84)	0.85 (0.83 - 0.87)

[a]Zero-shot inference without domain adaptation.

[b]Fine-tuned using low-rank adaptation on the nursing speech corpus.

[c]Manually transcribed references used as the gold standard for automatic speech recognition evaluation.

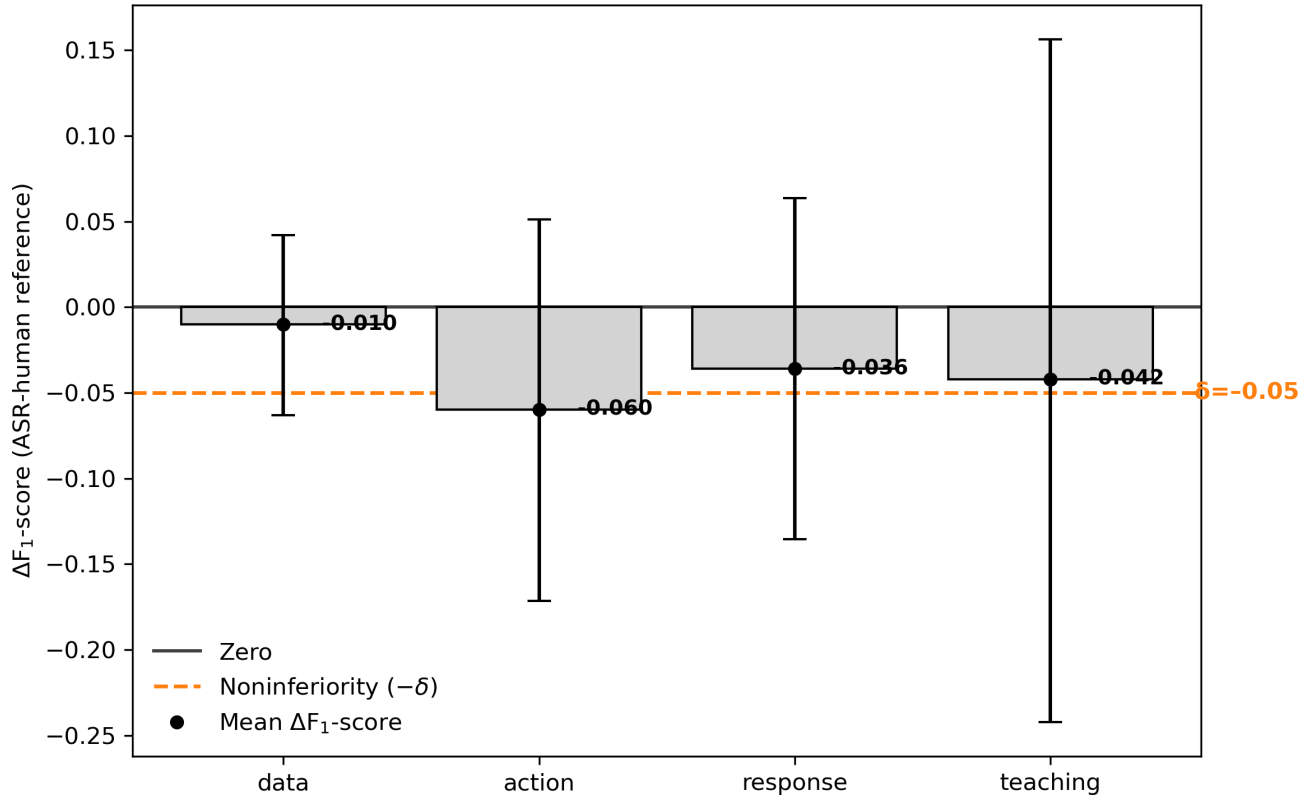
[d]Unweighted mean F_1-score across data, action, response, and teaching fields.

With this configuration, overall F_1-score values were 0.85 for human transcripts, 0.82 for fine-tuned ASR, and 0.70 for zero-shot ASR. The difference between fine-tuned ASR and human input (ΔF_1-score=−0.03) fell within the predefined noninferiority margin (δ=0.05). Field-level ΔF_1-score values ranged from −0.01 (for “data”) to −0.06 (for “action”). Zero-shot ASR showed uniformly lower performance across all fields.

Figure 2 presents field-level ΔF_1-score estimates with 95% CIs. All field means remained within the noninferiority boundary, supporting noninferiority at the macro level.

XSL•FO
RenderX

Figure 2. Field-level ΔF_1-score values (automatic speech recognition [ASR]–human reference) with 95% CIs and the noninferiority boundary (δ=−0.05).



Agreement Evaluation

Concordance was assessed across the 3 input conditions while holding the LLM constant. Agreement between classifications derived from the fine-tuned ASR and human transcripts ranged from 88.38% to 96.94%, with Cohen κ ranging from 0.66 to 0.86 across the "data," "action," "response," and "teaching" fields (Table 5), corresponding to substantial to almost perfect agreement [28]. For the "data" field, raw agreement was 95.41% with κ=0.66, a pattern consistent with imbalanced class prevalence. Overall, slot-level results were consistent with the macrolevel noninferiority findings.

Table . Agreement and Cohen kappa for DART (data, action, response, and teaching) fields comparing large language model outputs generated from fine-tuned automatic speech recognition with outputs from human transcripts.

Fields	Agreement[a] (%)	Cohen κ[b]
Data	95.41	0.66
Action	88.38	0.75
Response	89.91	0.79
Teaching	96.94	0.86

[a]Percent agreement between large language model outputs generated from fine-tuned automatic speech recognition transcripts and human transcript inputs.

[b]Chance-corrected agreement coefficient computed using the standard Cohen kappa formulation.

Hallucination Evaluation

Hallucination rates were evaluated using the FactualCorrectness metric from the RAGAS framework in precision mode (Table 6). With human transcripts, the mean hallucination rate was 2.35% (SD 9.93%; 95% CI 1.27–3.43), and hallucinations were detected in 28 (8.56%) samples (95% CI 5.99–12.10). Fine-tuned ASR yielded a mean hallucination rate of 2.51% (SD 10.86%; 95% CI 1.33–3.70), with hallucinations detected in 26 (7.95%) samples (95% CI 5.48–11.40). Zero-shot ASR showed higher hallucination rates, with a mean rate of 8.98% (SD 23.99%; 95% CI 6.36–11.60) and hallucinations detected in 65 (20%) samples (95% CI 16.01–24.69).

XSL•FO
RenderX

Table . Hallucination rates of large language model outputs across different input sources.

Input sources	Hallucination rate[a] (%), mean (SD)	Samples with hallucination[b], n (%)
Human transcript	2.35 (9.93)	28 (8.56)
Fine-tuned ASR[c]	2.51 (10.86)	26 (7.95)
Zero-shot ASR	8.98 (23.99)	65 (20.00)

[a]Computed using the FactualCorrectness metric (precision mode), representing the proportion of hallucinated tokens relative to all generated tokens.

[b]Percentage of evaluation samples containing at least 1 hallucination (n=327).

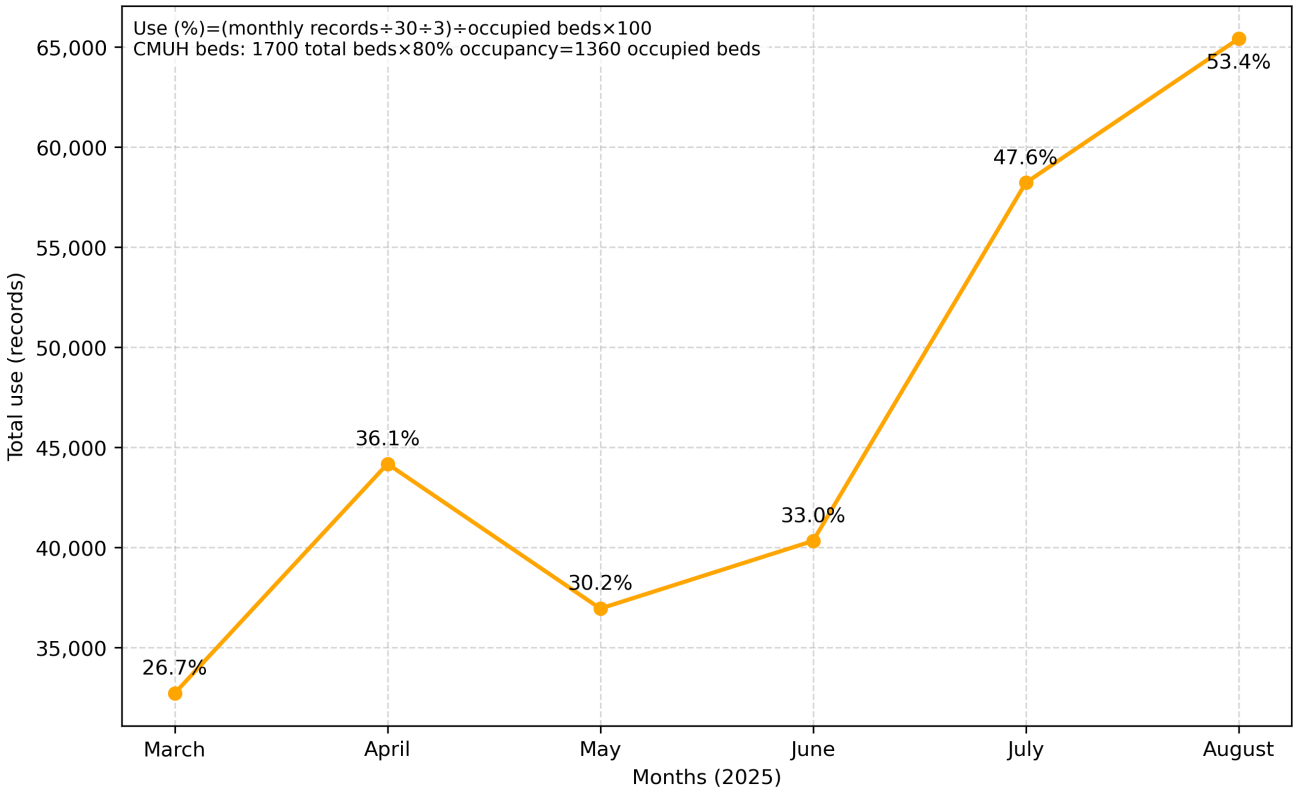
[c]ASR: automatic speech recognition.

System Use and User Feedback

Following deployment, the system was increasingly incorporated into routine documentation at CMUH. Monthly use nearly doubled between March 2025 and August 2025, rising from 32,724 to 65,417 valid DART notes generated and uploaded to the HIS through voluntary use of the system (Figure 3). This upward steady increase indicates expanded use across clinical units during the initial deployment period.

User feedback was collected from a subset of nursing staff who voluntarily provided evaluations (n=120, representing approximately 12.2% of the 982 nurses actively working in the 44 participating units across wards and intensive care departments). Among these participants, 91 (75.8%) nurses reported a favorable experience, whereas 29 (24.2%) expressed dissatisfaction. Positive feedback emphasized reductions in manual transcription burden and improvements in documentation efficiency, whereas negative feedback primarily concerned occasional recognition inaccuracies and the continued need for verification. Although the feedback sample does not represent all users, it reflects the perspectives of frontline staff actively engaged with the system. Detailed departmental participation is provided in Multimedia Appendix 2. Collectively, these findings demonstrate that the system is both technically feasible and operationally acceptable in real-world clinical environments.

Figure 3. Monthly system use at China Medical University Hospital (CMUH) from March 2025 to August 2025.



Discussion

This study designed and implemented a speech-based documentation system that integrates a domain-adapted ASR model with an LLM to generate structured nursing records and evaluated its performance in real clinical settings. The fine-tuned Whisper large-v2 model achieved a MER of 6.6%, demonstrating high accuracy for Mandarin-English nursing speech. With schema-guided LLM structuring, the system reached an F_1-score of 0.82, which was statistically noninferior to human-transcribed input, and exhibited a low hallucination frequency (2.51%). These findings indicate that domain-adapted ASR combined with LLM-based structuring can produce structurally consistent draft documentation that, alongside

XSL•FO
RenderX

routine human verification, reduces cognitive and transcription burden in routine practice.

Field deployment demonstrated that the system could be incorporated into routine clinical workflows. By enabling point-of-care mobile dictation directly at the bedside, the system effectively replaced the traditional, error-prone workflow of delayed, memory-reliant documentation at the nursing station, directly addressing the critical pain points exacerbated by high nurse-to-patient ratios. System use doubled over the 6-month observation period. Because the hospital operates at a consistently high and stable capacity, this increase in absolute volume was not driven by fluctuations in patient census. Rather, as the system was deployed as an optional tool, this steady growth reflects successful voluntary adoption, supported by continuous iterative optimizations based on user feedback. Feedback from nursing staff further supported its usability. Most respondents reported reduced manual effort and improved efficiency, while negative feedback focused on recognition accuracy and the need for verification, which aligns with the human-in-the-loop design. Participation across 6 major ward categories indicates that the system was used in a wide range of clinical settings.

Several limitations warrant consideration. First, the dataset was modest and sourced from a single institution, which may constrain generalizability. Second, user feedback was voluntary and may overrepresent individuals who were more engaged with the system, and long-term patterns of use were not assessed. Third, our evaluation of structural content accuracy and hallucination relied primarily on an LLM-as-a-judge framework (RAGAS). The primary purpose of this evaluation was to conduct a strict semantic fidelity check, ensuring that the LLM strictly formatted the ASR transcripts into the DART schema without altering, inferring, or omitting clinical facts. Although RAGAS performs robust semantic contradiction checks, traditional non-LLM baselines (eg, exact string matching) were deemed unsuitable because of the heavy code-switching and spoken-to-written paraphrasing inherent in nursing documentation. Although senior nursing staff qualitatively reviewed the generated drafts within our human-in-the-loop workflow, the lack of a formal, quantitative manual fidelity evaluation on a stratified sample of high-risk note types by nursing professionals remains a limitation. Future work should include multisite validation; evaluation of interoperability with diverse HIS environments; dedicated quantitative manual safety reviews by senior nurses to further validate that the LLM does not alter critical nursing parameters (eg, medication dosages and invasive device settings); and longitudinal assessment of impacts on documentation quality, workflow efficiency, and staff satisfaction.

Funding

The authors declare that no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Prompt templates for DART (data, action, response, and teaching) versions 1 and 2 used to guide large language model–based structuring of nursing narratives into DART records.
[DOCX File, 20 KB - nursing_v9i1e88567_app1.docx]

Multimedia Appendix 2
Distribution of user feedback by ward category.
[DOCX File, 13 KB - nursing_v9i1e88567_app2.docx]

References

1. Ma H, Chiang SC, Lin MH, et al. Understanding the factors associated with nurse employment in clinics: experiences in Taiwan. J Chin Med Assoc 2024 Jul 1;87(7):670-677. [doi: 10.1097/JCMA.0000000000001118] [Medline: 38819144]
2. Wu SF, Ching CY, Liao HC, Wang RH. Pathways among the nursing practice environment, job burnout, and job satisfaction to intention to leave: a cross-sectional study conducted in Taiwan. Rev Esc Enferm USP 2024;58:e20240025. [doi: 10.1590/1980-220X-REEUSP-2024-0025en] [Medline: 39392469]
3. Lin RT, Lin YT, Hsia YF, Kuo CC. Long working hours and burnout in health care workers: non-linear dose-response relationship and the effect mediated by sleeping hours-a cross-sectional study. J Occup Health 2021 Jan;63(1):e12228. [doi: 10.1002/1348-9585.12228] [Medline: 33957007]
4. Chen YC, Guo YL, Chin WS, Cheng NY, Ho JJ, Shiao JS. Patient-nurse ratio is related to nurses' intention to leave their job through mediating factors of burnout and job dissatisfaction. Int J Environ Res Public Health 2019 Nov 29;16(23):4801. [doi: 10.3390/ijerph16234801] [Medline: 31795420]
5. Huang FT. Caring for computers: the hidden work of clinical nurses during the introduction of health information systems in a teaching hospital in Taiwan. Nurs Rep 2021 Feb 13;11(1):105-119. [doi: 10.3390/nursrep11010011] [Medline: 34968317]

XSL•FO
RenderX

6. Hsu CN, Huang K, Lin FJ, et al. Continuity and completeness of electronic health record data for patients treated with oral hypoglycemic agents: findings from healthcare delivery systems in Taiwan. Front Pharmacol 2022;13:845949. [doi: 10.3389/fphar.2022.845949] [Medline: 35444533]
7. McCarthy S, Motala A, Lawson E, Shekelle PG. Use of structured handoff protocols for within-hospital unit transitions: a systematic review from Making Healthcare Safer IV. BMJ Qual Saf 2025 Sep 18;34(10):680-690. [doi: 10.1136/bmjqs-2024-018385] [Medline: 40306923]
8. Yang MH, Liu CY, Wang TJ, et al. A comparison of two cross-sectional studies on successful model of introducing nursing information system in a regional teaching hospital in Taiwan. Comput Inform Nurs 2022 Aug 1;40(8):571-579. [doi: 10.1097/CIN.0000000000000818] [Medline: 34740222]
9. Chang HY, Huang GL, Lotus Shyu YI, et al. Deep structure usage of electronic patient records: enhancing the influence of nurses' professional commitment to decrease turnover intention: deep structure usage and turnover. J Nurs Manag 2024;2024:5822368 [FREE Full text] [doi: 10.1155/2024/5822368] [Medline: 40224842]
10. Lampe SS. Focus charting: streamlining documentation. Nurs Manage 1985 Jul;16(7):43-46. [Medline: 3847830]
11. Wang CL. Focus charting: streamlining documentation. Hu Li Za Zhi 1993;40(3):87-92 [FREE Full text]
12. Wu CC, Tsai HW. The effectiveness of application the focus charting system in clinical nursing care plan. J Hosp Bimonthly 2016;49(5):22-33 [FREE Full text] [doi: 10.30260/HOSPITALBi.201610_49(5).0004]
13. Jacques D, Will J, Dauterman D, et al. Evaluating nurses' perceptions of documentation in the electronic health record: multimethod analysis. JMIR Nurs 2025 Apr 28;8:e69651. [doi: 10.2196/69651] [Medline: 40294588]
14. Duggan MJ, Gervase J, Schoenbaum A, et al. Clinician experiences with ambient scribe technology to assist with documentation burden and efficiency. JAMA Netw Open 2025 Feb 3;8(2):e2460637. [doi: 10.1001/jamanetworkopen.2024.60637] [Medline: 39969880]
15. Shour AR, Anguzu R, Onitilo AA. Speech recognition technology and documentation efficiency. JAMA Netw Open 2025 Mar 3;8(3):e251526. [doi: 10.1001/jamanetworkopen.2025.1526] [Medline: 40126483]
16. Lee TY, Li CC, Chou KR, et al. Machine learning-based speech recognition system for nursing documentation - a pilot study. Int J Med Inform 2023 Oct;178:105213. [doi: 10.1016/j.ijmedinf.2023.105213] [Medline: 37690224]
17. Hou SY, Wu YL, Chen KC, et al. Code-switching automatic speech recognition for nursing record documentation: system development and evaluation. JMIR Nurs 2022 Dec 7;5(1):e37562. [doi: 10.2196/37562] [Medline: 36476781]
18. Ng JJ, Wang E, Zhou X, et al. Evaluating the performance of artificial intelligence-based speech recognition for clinical documentation: a systematic review. BMC Med Inform Decis Mak 2025 Jul 1;25(1):236. [doi: 10.1186/s12911-025-03061-0] [Medline: 40598136]
19. Radford A, Kim JW, Xu T, Brockman G, McLeavey C, Sutskever I. Robust speech recognition via large-scale weak supervision. arXiv. Preprint posted online on Dec 6, 2022. [doi: 10.48550/arXiv.2212.04356]
20. Yang H, Zhang M, Tao S, Ma M, Qin Y. Chinese ASR and NER improvement based on Whisper fine-tuning. In: 25th International Conference on Advanced Communication Technology (ICACT): IEEE; 2023:213-217. [doi: 10.23919/ICACT56868.2023.10079686]
21. Yang CK, Huang KP, Lu KH, Kuan CY, Hsiao CY, Lee HY. Investigating zero-shot generalizability on Mandarin-English code-switched ASR and speech-to-text translation of recent foundation models with self-supervision and weak supervision. In: 2024 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW): IEEE; 2024:540-544. [doi: 10.1109/ICASSPW62465.2024.10626762]
22. Hu EJ, Shen Y, Wallis P, et al. LoRA: low-rank adaptation of large language models. arXiv. Preprint posted online on Jun 17, 2021. [doi: 10.48550/arXiv.2106.09685]
23. Achiam J, Adler S, Agarwal S, et al. GPT-4 technical report. arXiv. Preprint posted online on Mar 15, 2023. [doi: 10.48550/arXiv.2303.08774]
24. Shorten C, Pierse C, Smith TB, et al. StructuredRAG: JSON response formatting with large language models. arXiv. Preprint posted online on Aug 7, 2024. [doi: 10.48550/arXiv.2408.11061]
25. Lu X, Gao X, Wang X, et al. Comparison of medical history documentation efficiency and quality based on GPT-4o: a study on the comparison between residents and artificial intelligence. Front Med (Lausanne) 2025;12:1545730. [doi: 10.3389/fmed.2025.1545730] [Medline: 40438356]
26. Decroos F, Springenberg S, Lang T, et al. A deep learning approach for histopathological diagnosis of onychomycosis: not inferior to analogue diagnosis by histopathologists. Acta Derm Venereol 2021 Aug 31;101(8):adv00532. [doi: 10.2340/00015555-3893] [Medline: 34405243]
27. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960 Apr;20(1):37-46. [doi: 10.1177/001316446002000104]
28. Es S, James J, Espinosa Anke L, Schockaert S. RAGAs: automated evaluation of retrieval augmented generation. In: Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations: Association for Computational Linguistics; 2024:150-158. [doi: 10.18653/v1/2024.eacl-demo.16]
29. Factual correctness. Ragas. URL: https://docs.ragas.io/en/stable/concepts/metrics/available_metrics/factual_correctness/ [accessed 2026-05-20]

XSL•FO
RenderX

30. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [Medline: 843571]

Abbreviations

ASR: automatic speech recognition
CMUH: China Medical University Hospital
DART: data, action, response, and teaching
HIS: hospital information system
LLM: large language model
MER: mixed error rate
RAGAS: Retrieval-Augmented Generation Assessment

Edited by T Risling; submitted 27.Nov.2025; peer-reviewed by M Mcdonald, Q Ng; revised version received 02.Mar.2026; accepted 04.May.2026; published 05.Jun.2026.

Please cite as:
Su MH, Wang WC, Hsu YM, Hou SY, Chuang SJ, Chang SS
Voice-Based Structured Nursing Documentation Using Automatic Speech Recognition and Large Language Models: Development and Evaluation Study
JMIR Nursing 2026;9:e88567
URL: https://nursing.jmir.org/2026/1/e88567
doi:10.2196/88567

© Meng-Han Su, Wei-Chun Wang, Yi-Min Hsu, Shih-Yen Hou, Su-Jung Chuang, Shih-Sheng Chang. Originally published in JMIR Nursing (https://nursing.jmir.org), 5.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Barriers and Enablers for Sustaining Nurse-Led Use of Clinical Decision Support Tools for Antibiotic Stewardship: Qualitative Study

Victoria L Tiase[1], PhD, RN; Aurie D Tovar[1], MSN, RN; Natalie Henning[2], MPH; Mariana Braga[2], MPH; Keelin McHugh[3], MPH; Lynn Xu[2], MPH; Haddy Bah[4]; Alice Yuroff[3], PhD; Rachel Hess[4], MS, MD; Devin M Mann[2], MS, MD; David Feldstein[3], MD; Elizabeth R Stevens[2], MPH, PhD

[1]Department of Biomedical Informatics, University of Utah, 421 Wakara Way, Salt Lake City, UT, United States

[2]Department of Population Health, NYU Grossman School of Medicine, New York, NY, United States

[3]Department of Medicine, University of Wisconsin School of Medicine and Public Health, Madison, WI, United States

[4]Department of Population Health Sciences, University of Utah, Salt Lake City, UT, United States

Corresponding Author:
Victoria L Tiase, PhD, RN
Department of Biomedical Informatics, University of Utah, 421 Wakara Way, Salt Lake City, UT, United States

Abstract

Background: Clinical decision support (CDS) tools embedded in electronic health records in the form of integrated clinical prediction rules provide a potentially effective intervention to reduce inappropriate antibiotic prescribing for acute respiratory infections. However, their effectiveness has been limited by workflow barriers and low adoption by health care providers. Nurses are well positioned to implement evidence-based protocols using CDS tools. In a multicenter randomized controlled trial, a nurse-led implementation strategy for acute respiratory infection integrated clinical prediction rules was evaluated for use in primary care and urgent care settings.

Objective: This study aimed to examine nurse and nurse leader perspectives on the sustainability of an electronic health record–integrated CDS tool for antibiotic stewardship and explored factors influencing its potential long-term integration into ambulatory nursing practice beyond the clinical trial.

Methods: We interviewed 22 nurses and nurse leaders from 37 clinics across 3 academic medical centers that participated in the clinical trial. Two semistructured interview guides, one for nurses and one for nursing leadership, were developed to understand the barriers and facilitators to implementing a decision aid tool for nurses and to elicit challenges specific to nursing interactions with the CDS tool. Interviews were recorded and transcribed. Using thematic content analysis and iterative coding, our team collaboratively identified emerging themes related to sustainability and refined the results with consensus.

Results: Five themes emerged: (1) importance of staffing stability and capacity, (2) impact of dedicated clinic resource availability, (3) variable nurse readiness with CDS-guided clinical care, (4) influence of openness to change and a nurse-supportive clinic culture, and (5) ongoing need for training and support. Specific recommendations for future actions were also noted.

Conclusions: Our findings revealed specific barriers and facilitators to the sustainability of a CDS tool from the nursing perspective that can inform further implementation of nurse-led delegation protocols in the ambulatory setting. Future solutions should consider mapping physical workflows, scheduling specific to nurse visits, continuing education, and treating cough and sore throat as 2 distinct processes.

(JMIR Nursing 2026;9:e83567) doi:10.2196/83567

KEYWORDS

clinical decision support systems; electronic health records; nursing; antibiotic stewardship; ambulatory care facilities

Introduction

Antibiotic resistance is a growing global health concern, exacerbated by the overprescription of antibiotics for acute respiratory infections (ARIs) [1]. Despite available evidence-based guidelines and clinical decision support (CDS), physicians often prescribe antibiotics inappropriately due to perceived patient expectations, time pressures, and cognitive fatigue [2]. To address this issue, clinical prediction rule (CPR) systems have been developed and integrated as integrated clinical prediction rules (iCPRs) into electronic health records (EHRs) as CDS to assist physicians in making informed antibiotic prescribing decisions [3,4]. However, their impact has been limited by low provider adoption rates [5,6]. Evidence

XSL•FO
RenderX

suggests that registered nurses (RNs) have the potential to implement evidence-based protocols [7,8], participate in antimicrobial stewardship [9], achieve high patient satisfaction and symptom resolution [7], and demonstrate higher adoption rates of CDS tools than physician-led initiatives [10].

Building on this potential, the multicenter, step-wedge, cluster randomized controlled trial (iCPR3 RCT) evaluated whether RNs could effectively use iCPR tools to guide evidence-based care for patients presenting with ARIs in ambulatory care settings [11,12].

An early implementation assessment identified several potential barriers and facilitators to adopting the nurse-driven CDS tools [10]. As the iCPR3 RCT concluded, we sought to understand whether nursing departments would consider continuing to use iCPR tools in clinical practice after research support was withdrawn. We also aimed to identify barriers that may hinder sustained use within ARI workflows beyond the study period. By examining these perspectives, our goal was to identify strategies to sustain nurse-led use of iCPR tools in clinical practice as a means to improve patient outcomes [13].

Methods

Research Design

We used semistructured interviews and thematic analysis to examine the perspectives of RNs and nurse leaders regarding the sustainability of the iCPR tools, including aspects of implementation, usability, and perceived impact. For the purposes of this study, we defined sustainability as the extent to which nurses would continue to use a tool as part of standard clinical workflows beyond the study period and without research personnel implementation support or incentives.

Ethical Considerations

The study protocol and procedures were approved by the NYU Langone Health Institutional Review Board (NYULH study number i19-01222), which served as the study's single institutional review board. We received a waiver of written informed consent for this study. All study data reported in this manuscript are deidentified. Compensation was provided in accordance with institutional policies.

Nurse iCPR Intervention

The nurse intervention used a structured delegation protocol, wherein the specific clinical task of assessing the need for antibiotics in low-acuity ARI patients was transferred from a physician to an RN. Provider-to-nurse delegation is operationalized through protocol-driven authorization of nursing tasks by a licensed provider, with accountability retained by the delegating provider. In the iCPR3 RCT, the intervention consisted of triage followed by in-person CDS-guided RN visits for patients with low-acuity ARI symptoms. The EHR-integrated CDS included a triage note template to assess symptoms and acuity and 2 reason-for-visit-specific (ie, sore throat or cough) note templates with iCPRs that guide RNs to complete an evidence-based risk calculator as part of the visit [12]. The resulting risk score is linked to an order set for diagnostic testing, prescriptions, and patient instructions.

All tasks performed by RNs as part of the intervention—triage, symptom assessment, use of CDS tools, and execution of evidence-based care pathways—are well within the scope of nursing practice and align with their clinical training. The iCPRs used were previously found to be effective for reducing inappropriate antibiotic prescribing when used by physicians [2,4,11]. The intervention aimed to standardize care delivery and reduce unnecessary antibiotic use in outpatient settings.

The intervention was implemented as a pragmatic stepped-wedge cluster randomized trial across 43 primary and urgent care practices in 3 academic health systems affiliated with the University of Utah, the University of Wisconsin, and NYU Langone. To be eligible for participation, practices were required to have at least 1 RN full-time equivalent capable of performing triage within the EHR and conducting RN on-site visits. Nurses in the intervention group received 1 hour of online training prior to 1 hour of in-person training on how to use the iCPR tool, including background on the iCPRs, EHR-based walkthroughs of using the iCPR tool within the CDS, demonstration videos simulating the tool in a live clinical encounter, and clinical training for physical examinations [11]. Complete study procedures are reported elsewhere [12].

Study Design

We conducted remote video interviews (via Zoom and Webex) lasting approximately 30 minutes between September 2024 and January 2025. Each participant was interviewed by a trained research team member from the corresponding health system (NH, KM, AY, VT) using 1 of 2 semistructured interview guides. Our team developed these interview guides, 1 for RNs and 1 for leadership, based on a series of open-ended structured questions previously used in a validated survey to understand the barriers and facilitators to implementing a lupus decision aid tool for nurses [14]. We adapted the questions to explore relevant Consolidated Framework for Implementation Research (CFIR) constructs, as well as address conditions specific to the ambulatory care environment and the iCPR intervention; we retained all other aspects of the survey [15]. Interviews consisted of questions to understand barriers and facilitators related to tool use, clinical team interactions, what worked well, any competing priorities, strategies for mitigating issues, and future adoption sustainability (see Tables S1 and S2 in Multimedia Appendix 1). Probing questions were used to verify the interpretation of participant responses, and recommendations were solicited for continued use of the iCPR tool in clinical workflows.

Participants

We recruited a convenience sample of RNs and nurse leaders (did not have to be an RN) from primary care clinics and urgent care centers participating in iCPR3. Participants were considered eligible if they expressed familiarity with the iCPR tool and, if they were nurse participants, had completed a nurse visit using the tool. Participants were recruited via email by study personnel at each site. Recruitment procedures varied slightly by study site due to local policies and conditions. Specifically, given the proportionately large number of clinics at the University of Wisconsin (21 clinics vs 9 and 7 clinics at NYU and Utah, respectively) and with the objective of obtaining diverse

XSL•FO
RenderX

perspectives, participants were recruited as a stratified convenience sample from the top and bottom 5 performing clinics, based on the average number of nurse visits completed per week per clinic volume at randomization. In addition, University of Utah participants were offered a US $25 gift card as an incentive, whereas other study sites did not provide financial incentives due to institutional policies around employee compensation. A member of the study team verbally reviewed the study information with each participant to ensure understanding; all participants then provided verbal agreement to participate.

Thematic Analysis Procedures

We conducted a thematic content analysis to identify themes related to sustainability by extracting high-level themes from the data while examining the frequency of concepts or keywords [16]. Each interview was audio-recorded and transcribed using an automated service (Landmark Associates). One team member (NH) reviewed each file for transcription accuracy prior to analysis. Transcripts were analyzed, and a codebook was developed using Dedoose [17].

We first conducted a joint analysis of 2 interviews to allow the study team members to become accustomed to the codebook, make updates to codes, and resolve conflicts. All study team members were trained in qualitative techniques. Next, 2 team members coded each transcript independently, and then all coders (VT, NG, MB, KM, AT) met as a team to compare and resolve disagreements. During this process, we reviewed for thematic saturation, ensuring the sample met this criterion based on the diminishing need to update code definitions. Our analysis focused on the barriers and facilitators to sustainability of iCPR tool use beyond the study end date. Once all transcripts were coded, we reviewed the data code by code to ensure definitions had remained consistent. All coders then reviewed each excerpt to collectively derive emerging themes related to barriers, facilitators, and recommendations for solutions. Emerging themes were grouped into high-level categories through consensus. This study adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines for qualitative studies involving interviews and focus groups [18].

Results

Descriptive Results

We invited 56 eligible individuals to participate and conducted 22 individual interviews across the 3 sites, each lasting 15 to 45 minutes. Participants included 9 nurses (1 NYU, 8 Wisconsin) and 13 nurse leaders (4 NYU, 2 Utah, 7 Wisconsin); demographic information was not collected from either group. Due to limited tool use during the study, no nursing participants from the University of Utah elected to participate in the interviews.

Through thematic analysis, we identified 5 overarching themes related to nurses' and nurse leaders' perspectives on the implementation, effectiveness, and sustainability of the iCPR intervention as part of standard care processes. The 5 themes that emerged included: (1) importance of staffing stability and capacity; (2) impact of dedicated clinic resource availability; (3) variable nurse readiness with CDS-guided clinical care; (4) influence of openness to change and a nurse-supportive clinic culture; and (5) ongoing need for training and support.

Importance of Staffing Stability and Capacity

One of the most prominent themes concerned the need to have sufficient trained RNs staffed and available in the clinic to support the time needed to triage patients and conduct nurse visits. In addition to nurse staffing, participants emphasized concerns related to physician and support staff levels. Contributing to these concerns, many participants (20, 91%) discussed issues related to staff turnover, including the challenges of retaining staff, training new staff, and implications for using the iCPR tool workflows.

When questioned about barriers to using the iCPR tool, one nurse stated:

> *We are staffed with only three nurses, which is not a lot… One of our nurses left in May of last year to work elsewhere. Then, we hired a new nurse. Our third nurse went on maternity leave. Then when she came back, the other nurse had to leave for other reasons too. We've kind of just been at this [struggle] of not being able to consistently have a good number of staffing.* [N 3.18]

The shortage of nursing support staff often required nurses to cover multiple roles, making it challenging to prioritize tasks and manage workload effectively. One participant stated:

> *…some clinics can't handle it… they don't have enough people to even answer the phones.* [N 3.3]

With full staffing, nurses expressed confidence in executing the iCPRs:

> *We are pretty much fully staffed… it has not been a burden to participate and be a part of this survey and process… it's certainly helped our patients.* [N 3.3]

> *We could have two nurses and visits at the same time and patients aren't having to wait.* [N 3.2]

Additionally, staff turnover was linked to challenges in familiarizing and educating new staff regarding the iCPRs due to limited training time. Increasing awareness among new nurses completing orientation who lacked exposure to the study was discussed. One nurse stated:

> *Now that I'm thinking about it, there's a whole bunch of nurses that started that they weren't probably there when it first started. They don't even know about it, and I don't know if it's part of their orientation packet.* [N 1.7]

A key subtheme involved virtual triage coordination, where nurses working remotely were able to triage patients by phone and monitor clinic appointment availability. Approximately one-quarter of the participants (6, 27%) felt that this workflow disrupted nurse visit appointment availability.

> *It's a small nursing team so you have anywhere from two to three nurses working on a given day and then with the nurses that have the option to work remote now sometimes you only have one RN [registered*

XSL•FO
RenderX

> *nurse] in clinic. That impacts availability for appointments, so I think that is always a challenge.* [N 3.11]

Another participant expressed difficulties coordinating with colleagues not physically in the same place, since the phone triage nurse at Wisconsin and NYU sites was typically different from the nurse who conducts the in-person nurse visit, highlighting staffing coordination in a hybrid care setting.

> *Sometimes, if you're at home triaging the patient, and they could use a nurse visit, trying to coordinate that with colleagues can be hard if it's a busy day. You don't really know what's going on over there [in the clinic]…* [N 3.15]

Impact of Dedicated Clinic Resource Availability

The impact of clinic resource availability, including physical space for nurse visits and scheduling accessibility, was identified as a barrier to uptake of the in-person nurse visit component of the iCPR workflow. The ability to manage the patient care continuum was disrupted when examination rooms were unavailable. One nurse stated:

> *…we don't necessarily have a place to go with our patients…again, sometimes we are just running around trying to find a location for the patient to be assessed in and to complete the visit in.* [N 3.4]

Discussion of the physical clinic setup highlighted that the lack of available examination rooms posed challenges. One RN stated:

> *We've always had to kind of just find a room on the fly for any of our nurse visit types each day, whether they're iCPR or not. A day like today, all of our exam rooms are occupied by the providers that are here. Our plan is just when the patient comes, if there is an available room, just kind of quickly utilize that.* [N 3.18]

Alternatively, when the clinic set aside dedicated rooms for nurse visits and triage, it reportedly streamlined workflows and supported iCPR tool uptake.

> *Structurally, we had a room all ready to do nurse visits that we do for other things. That was easy enough to adopt this [iCPR] into that.* [N 3.3]

> *We have a designated room for nurse visits every day 'cause we have nurse visits every day, and it's an exam room at the front of the clinic, and we have everything that we need in there.* [N 3.10]

Coupled with dedicated room space, time set aside for nursing appointments facilitated iCPR care workflows. One clinic workaround included having patients with sore throats or coughs come in at prescheduled time slots so that nurses and testing resources (such as throat swabs) were available. This allowed the clinic to structure the schedule in a way that allowed nurses to complete patient care activities.

> *I know that one of the things that we tried to focus on and try to do too, as far as our department, is we tried to make sure that we had like the 11:00 and 1:00 p.m. slots available for the nursing and cough and sore throat visits…* [N 3.4]

Variable Nurse Readiness With CDS-Guided Clinical Care

Among RNs who received training, there was variability in their comfort when providing clinical care guided by the iCPR tools. This was expressed as resistance to the delegation portion of the intervention regarding adherence to the iCPRs, staying within their scope of practice, and compatibility with their clinical skills.

Some nurses felt organizational resistance to operationalizing the delegation protocol, which required nurses to practice to the full extent of their qualifications. One nurse leader expressed that the intervention enables nurses to function within the bounds of what they are licensed to do, but existing clinic workflows may not be able to accommodate them.

> *It's asking our nurses to work at the top of our license. The problem is that it takes—the way that they've worked in urgent care for so many years, it causes a change in that workflow…Some of the leaders have been more willing to facilitate that process, and some have not.* [L 2.14]

Nurses expressed discomfort and uncertainty when triaging patients with cough-related symptoms, particularly when required to perform comprehensive respiratory assessments such as auscultating lung sounds. This hesitation often stemmed from fear of missing critical findings and potentially causing harm, which made them reluctant to fully engage with the iCPRs designed for these cases.

> *…there's a lot of nurses who do not like the cough part. They don't like to listen to lungs, and the reason is they're afraid they're going to miss something and harm patients.* [L 3.20]

Educating nurses was highlighted as a potential means of addressing these barriers. One nurse leader shared:

> *Training of new people training is huge, so I think and many people have to be hands on… mixture of hands-on CBT [Computer Based Training] all of that to reach the nurses…but I think training is key.* [L 3.20]

Conversely, nurses reported a strong sense of satisfaction with the sore throat iCPR, noting its simplicity and efficiency, allowing patient assessments to be completed in as little as 15 minutes.

> *The sore throat one we'll use constantly. That is very much something that has been, in my opinion, and I think our providers, everybody really, has been very successful.* [N 3.3]

Influence of Openness to Change and a Nurse-Supportive Clinic Culture

Nurse Factors

A supportive clinic culture, openness to change, and positivity toward nurse support influenced the use and sustainability of

XSL•FO
RenderX

the iCPR-guided nurse delegation workflow. Within clinic culture, the perceptions of support from providers and patients emerged as subthemes, and leadership support was seen as crucial in facilitating the usage of the iCPR tool. Participants reported that leaders who actively listen to nurses and maintain ongoing dialogue foster a supportive environment that encourages new practices. Nurse leaders agreed that supporting staff was a success factor.

> *I think just making sure that staff are feeling confident and comfortable with the process, hearing where there's issues…I think just hearing what the staff is experiencing, trying to support them through any issues that are coming up, and making sure that they're confident in their training to conduct the visits.* [L 3.11]

> *Ongoing discussions with the nursing team to make sure that they feel supported.* [L 3.1]

The overall clinic culture was perceived to be particularly critical. One nurse described the culture of being open to novel processes.

> *…I think that our culture, to use that term for sure, is exceptional in that way. As far as being open to adopting things. Moving forward with nursing skills and nursing involvement in clinic practice is certainly something we focus on as a group and with our leadership.* [N 3.1]

Another participant reported the benefit of a shared value system within the clinic:

> *We have an excellent connection between each other. I think we all work for each other and with each other. There is a shared value system there that was very apparent and that is healthy in many ways… If my boss wasn't that interested in this, I think that we as a group would still be very interested together. We would have just done it by ourselves.* [N 3.1]

Moreover, nurses expressed aspects of professional fulfillment due to spending more time with patients:

> *Talking to the patients is great. Sometimes it's nice, the ones that you talk to, then you get to do a nurse visit.* [N 1.7]

> *Once again, I think it's great for us. It's great for the patient. I think being accessible and giving our patients what they need is—just been a really nice experience.* [N 3.2]

Provider Factors

Factors related to providers were found to be an essential subtheme of the clinic culture. Overall, the provider's acceptability of the processes was viewed as necessary for the sustainability of the iCPR3 tool.

> *We'd have to get the… medical director on board to really get the providers engaged with this. I do feel like providers are one of the bigger barriers. The nursing teams can do this work, but unless we work in collaboration, it'd be pretty hard to have it be sustainable.* [L 2.14]

When medical directors showed interest in teaching and experimenting with new workflows that incurred time savings, the tool's adoption by RNs was supported.

> *If they perceive, or if it's a reality that this program takes visits away from them, rather than gives them more time to see other patients. Then that could be a deterrent and also just their own understanding of how nursing practice can be used. I think in ambulatory it's sometimes disinterested.* [L 1.13]

> *When the doctors like to teach…then it's very helpful. Some doctors… don't even want to do the study…but the ones that want to teach then, that's very helpful.* [N 1.7]

Engagement from a medical director through collaboration was deemed important for facilitating clinic change.

> *It's doable, it's just, again, it's change management. It's really having everybody on board. Probably something that would help it move along better would be all the leaders and medical director meeting and deciding this is the right process….* [L 2.14]

Patient Factors

Patient engagement varied by location and was viewed as tightly connected to the clinic culture. Some nurses reported challenges in communicating the iCPR outcome to patients, especially when the patient had certain expectations connected to receiving antibiotics as part of the visit.

> *We have a low compliance rate in our clinic. When they come to see us, sometimes even when you're like, "You're not gonna see a doctor. You're just gonna see a nurse." They just want to come in and get antibiotics… When they come in and see us, I'm really like, "You're not gonna [get antibiotics]," but sometimes does not go over well with our patient population.* [N 3.10]

A number of issues also described a conflict between the delegation process and the patient's preference for provider type.

> *We do get some that are not interested in having the nurse visit type because they are looking to discuss other remedies or other treatment options, or their complexity medically is sometimes a barrier.* [N 3.18]

> *I think our patient population is different than other clinics. I think that that has made it a little bit harder… We do get pushback from our patients. They're like, 'Well, you're not a doctor,'…*

> *[N 3.10]*

However, some patients appreciated the extra access to appointments and were willing to have a nurse visit and were satisfied with the nurse visit.

> *I feel like patients are very agreeable and willing to come in to see a nurse, after we've explained exactly what's gonna happen… Patients have been really*

XSL•FO
RenderX

> *appreciative. No cost, they get in the same day, which is really what they're looking for…* [N 3.2]

Clinic Communication

General clinic communication was the final subtheme related to clinic culture. A nurse leader expressed interest in having more usage data about the study's progress to communicate with the staff for positive reinforcement.

> *…if people know the why we're doing it that helps…It really comes down to great patient care, antibiotic stewardship, how many visits we did. All of that. I think if we have that data, then as we train, we can bring that forward and show people why we're doing it.* [L 3.20]

Nurse participants also reported satisfaction connected to the receipt of feedback on the iCPR use.

> *I think it's been very positive and we—you guys send out the statistics and it reinforces that we are doing a good job and that always obviously makes me feel good that you are making a difference and… because obviously it's helping everybody. It's helping our patients. It's helping us.* [N 3.2]

Group communication and decision-making as a team were noted as key to the sustainability of clinic initiatives.

> *Again, it's about everyone deciding together that this is the right path and the way we want to go. Then moving it forward as a group so that it's not over here at this clinic but not at this clinic…It's just making sure you've got both nursing and operations, if we're going to really operationalize something and put it into permanent status.* [L 2.14]

Another participant emphasized the importance of teamwork among colleagues, noting the ability to share experiences and seek advice from others to navigate challenges.

> *My co-workers, my colleagues, us working together was another really important piece… Have you done this? Did you run into that?* [N 3.3]

In addition, communication tools were reported to provide support beyond programmatic issues and were seen as helpful in solving acute technology issues.

> *The Webex group was helpful troubleshooting technology issues. They were really good with the patient care part, but, if a technology issue came up, the Webex group was really helpful.* [N 3.6]

Ongoing Need for Training and Support

Participants felt that additional training, specifically hands-on training, would enhance the program's sustainability.

> *… Can we retrain? Can we have champions at the site who can help? Training of new people training is huge, so I think and many people have to be hands-on. Is it a mixture of hands-on CBT [computer-based training]… I think training is key.* [N 3.2]

Another participant reported that recurrent training would help with maintaining skills:

> *Offering skills' refreshers every once in a while if this is gonna be an ongoing thing would be helpful too.* [N 3.15]
>
> *I think education and I think not just a one and done education…at the time of hire, at the time of roll out or sustainability, but then I almost think something yearly just for people to review, ask questions, do hands on skills again* [L 3.20]

In the context of simulation training, a playground or safe, interactive environment was described to bridge this gap, providing a way for learners to practice skills without real-world consequences. One nurse leader was enthusiastic about ways to integrate hands-on training with the iCPR intervention.

> *I would say the hands-on training was fantastic for the nurses. They really appreciated it. It made them feel much more comfortable… in a setting where they can use the playground and practice the swabbing and stuff… making sure we can do a real-world, full-picture visit for practice would be really helpful.* [L 3.6]

Both nurse and nurse leader participants considered nursing leadership support to be part of the sustainability of this program. One of the nurse leaders described their approach as:

> *Just making sure the nurses are comfortable and have the training and education that they need. If the cough part is a barrier, working to get them hands-on training, making sure they've got swabs that they need, stethoscopes that they need, and just being support for them. If they don't have a buddy in the clinic that day, maybe I step in and do it with them.* [L 3.6]

Recommendations Based on Participant Feedback

Participants provided several recommendations to enhance the sustainability of the iCPR tool in nursing practice (Table 1).

XSL•FO
RenderX

Table . Recommendations.

Recommendation	Representative quote
Identify clinic champions	"have champions at the site." "we're going to need a provider champion."
Educate patients	Educate the patients more on "This is what we're doing."
Organize frequent lunch and learns	"circling back to another … skill driven practice type session might be helpful just to make sure … we still feel confident that we can conduct it and not feel too flustered."
Perform triage in advance	"When the first person that gets the [triage] call says, "Oh, they have a cough or a sore throat." That's very helpful because right away it's identified …"
Map physical workflows	"… in the future … you can see the layout of how things are, then you can maybe give a suggestion or tell us something to make a different flow that works somewhere else."
Incorporate into routine practice and policy	"… has to make it a delegation process … more than anything."

Discussion

Principal Findings

The examination of the perceptions of nurses and nurse leaders familiar with the implementation of a CDS-guided nurse intervention for ARIs revealed several barriers and facilitators to the sustainability of the tools' use beyond the conclusion of the study. Factors perceived as contributing to sustainability emerged within 5 themes, including (1) the importance of staffing stability and capacity; (2) the impact of dedicated clinic resource availability; (3) variable nurse readiness with CDS-guided clinical care; (4) the influence of openness to change and a nurse-supportive clinic culture; and (5) the ongoing need for training and support. By touching on the characteristics of the intervention, clinic setting, and individuals, these themes highlight several CFIR domain constructs and strategies at the individual and organizational levels to be considered when seeking to develop a sustainable CDS-guided intervention [15,19]. The most pertinent constructs identified in this research are local attitudes and conditions, partnerships and connections, structural characteristics such as physical and work infrastructure, organizational culture, available resources and space, the roles of innovation deliverers (nurses) and recipients (patients), as well as strategies for tailoring and adapting the use of the iCPR tool.

Individual-Level Strategies

Our analysis highlighted that nurses expressed lower confidence when using the iCPR tool for conditions requiring more complex or subjective clinical assessments, such as listening to breath sounds, compared to workflows perceived as more straightforward, such as visual inspections for sore throat evaluations. Indeed, several participants indicated a preference for sustaining use of the sore throat iCPR workflow only. This gap in confidence suggests a need for targeted skill development. Simulation-based training offers a valuable strategy to build these skills in a realistic, low-risk environment. Regular simulations can help nurses practice nuanced assessments, receive real-time feedback, and strengthen clinical judgment. Incorporating iCPR-related scenarios into onboarding and ongoing education can reinforce competence and confidence. Fostering a culture of continuous learning, with structured supervision and feedback, is essential for supporting effective and sustained implementation [20]. In alignment with CFIR individual constructs [15], the identification of subject matter experts could further assist and support access to information and knowledge.

Organizational-Level Strategies

As seen with many health care interventions, staffing levels were perceived as a key factor influencing long-term sustainability [21]. Inadequate ambulatory clinic nurse staffing impedes staff mastery of new procedures, restricts training time, and hinders comprehension of essential tools, such as EHR CDS modules [22]. Moreover, the increased complexity of ambulatory nursing requires nurse leaders to proactively adjust outpatient staffing models to meet evolving demands [23]. Staffing models were not adjusted to accommodate iCPR3; instead, participating units were only required to have at least one nurse. This model worked well for some clinics but not all. In settings with limited staffing, nurses may be forced to prioritize urgent clinical tasks over engaging with new initiatives, making it difficult to implement interventions effectively. Our findings suggest that a "one size fits all" approach may be inadequate. Instead, staffing models should be tailored to the needs of each clinic, particularly when new interventions are introduced. Due to national nursing shortages, calibrating nurse staffing is more important than ever before [24]. Adequate staffing enables nurses to focus on their patient care activities [24,25], allowing for better adherence to the CPRs. Study participants highlighted encountering challenges integrating new hires, resulting from turnover, and upskilling nurses who have transferred from other specialty areas. Prior to intervention implementation, analyses of turnover rates, percentage of new hires, and overall patient volume should be considered to ensure appropriate staffing availability, aligning with the CFIR construct of general staffing levels to support the functional performance of the inner setting [16]. Strategic planning should incorporate anticipated staffing fluctuations and allow flexibility in response to sudden staffing interruptions. The use of alternative staffing models, such as virtual nursing resources, needs to be considered. While an efficient use of nursing resources in some settings [26], hybrid work environments may contribute to barriers to care coordination and in-person visits.

XSL•FO
RenderX

Availability of clinic resources, especially physical space, was emphasized as a barrier more than expected. Success with the iCPR workflow hinged on ready access to supplies, designated rooms for nurse consultations, a supportive physical layout, and timely provider availability. Compared to inpatient hospital care, the outpatient environment is known to often lack infrastructure support [27]. When nurses had to search for appropriate locations to assess patients upon arrival, both efficiency and patient experience suffered. In contrast, clinics that reserved rooms specifically for nurse visits reported smoother workflows and better outcomes. As per the CFIR physical infrastructure construct, the impact of the layout and configuration of the space affects functional performance [15]. In nonambulatory clinics, similar issues have been previously observed, indicating that both ambulatory and nonambulatory settings face challenges related to resources. Despite progress in CDS tools, considerable barriers remain, especially regarding workflow integration and resource allocation [27]. Future implementations should therefore include a thorough assessment of space and workflow needs, ensuring that dedicated nurse-visit rooms and adaptable layouts are built into the design from the onset [28].

Clinics with positive experiences using the iCPR tools often reported a strong organizational culture characterized by teamwork, openness to change, and leadership that valued nurse autonomy, clinical expertise, and collaborative practice. Supportive leadership and a team-based approach further enabled the effective coordination between nurses and providers required to implement and sustain a new CDS intervention. As recognized by updated CFIR constructs, local attitudes and beliefs influence future use [15]. Participants also emphasized the importance of involving patients in shared decision-making, underscoring the need to respect patient preferences and set clear expectations during care encounters. This finding suggests that fostering a culture that empowers nurses, supports interdisciplinary collaboration, and centers patient engagement may be key elements to the successful nurse adoption of CDS tools in the ambulatory care setting [29].

Recognizing that the iCPR RCT took place in ambulatory settings connected to academic medical centers, there are insights that may translate to nonacademic centers. While community centers may be less resourced for experimentation and have underdeveloped nursing platforms, nurses in nonacademic settings often work at the top of their license out of necessity for patient throughput [30]. Prior to implementing the recommended strategies that connect to our findings, it will be important to compare workload patterns and assess the need for delegation protocols prior to implementing a similar intervention.

Limitations

Several limitations should be noted. This study was conducted within academic medical centers, which are typically characterized by greater infrastructural resources and institutional support, potentially limiting the generalizability of findings for community or resource-constrained settings. Additionally, we did not account for the differences in institutional policies, patient populations, or resource constraints as part of our examination. While thematic saturation was achieved, the final sample size was relatively small, and the sample skewed toward nurses at one site and toward nurse leaders rather than frontline nursing staff, which may have introduced hierarchical bias in perceptions of implementation. However, given low intervention engagement during the RCT at the Utah and NYU sites and that Wisconsin had the most nurses using the tool, we believe the sample was a representative of the iCPR RCT site variability. Participant selection may have been subject to response bias, as we used a convenience sample and did not stratify responses based on participants' familiarity with technology, experience, or frequency of tool usage. Because use of the intervention tools was part of the inclusion criteria, we did not assess the perspectives of nurses who did not perform a nurse visit or who did not participate in the RCT. In addition, participants were recruited and interviewed by known study personnel at their corresponding sites, which may have led to social desirability bias and impacted candor in participant responses. Lastly, this analysis does not include an examination of tool utilization or patient-level outcomes and their influence on perceived sustainability.

Conclusion

This qualitative analysis underscores the multifaceted nature of sustaining a CDS-guided nurse iCPR intervention for ARIs in ambulatory care settings after a formal study period. The findings reveal that sustainability hinges not only on the design of the intervention but also on contextual factors such as staffing adequacy, resource availability, nurse confidence, organizational culture, and ongoing training. Importantly, variability in nurse comfort with CDS tools and differential success across clinic settings suggest that tailored implementation strategies are essential. Future efforts should prioritize clinic-specific readiness assessments, targeted skill development, and leadership engagement to enhance the long-term viability of nurse-led CDS interventions in clinical workflows.

Acknowledgments

The authors wish to acknowledge health care professionals for their professional dedication to patients and their contributions to this project.

Funding

This work was funded by the National Institute of Allergy and Infectious Diseases (5R01AI108680; principal investigator DMM). The funder had no role in the content or the decision to publish this manuscript.

XSL•FO
RenderX

Data Availability

The datasets generated during this study are not publicly available due to privacy restrictions but may be made available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: VLT, ADT, NH, MB, KM, LX, HB, AY, RH, DMM, DF, ERS
Data acquisition: VLT, ERS, NH, MB, KM, AY, ADT
Design: VLT, ADT, NH, MB, KM, LX, HB, AY, RH, DMM, DF, ERS
Methodology: VLT, ADT, NH, MB, KM, LX, HB, AY, RH, DMM, DF, ERS
Formal analysis: VLT, ERS, NH, MB, KM, AY, ADT
Interpretation: VLT, ERS, NH, MB, KM, AY, ADT
Investigation: VLT, ADT, NH, MB, KM, LX, HB, AY
Data curation: VLT, ERS, NH, MB, KM, ADT
Writing – original draft: VLT, ADT, ERS
Writing – review & editing: VLT, ADT, NH, MB, KM, LX, HB, AY, RH, DM, DF, ERS
Supervision: DM, DF, ERS, RH
Project administration: VLT, ERS
Funding acquisition: DM, DF, RH

Conflicts of Interest

None declared.

Multimedia Appendix 1
iCPR survey tool questions.
[DOCX File, 18 KB - nursing_v9i1e83567_app1.docx]

References

1. Stevens ER, Feldstein D, Jones S, et al. Ambulatory antibiotic prescription rates for acute respiratory infection rebound two years after the start of the COVID-19 pandemic. PLoS One 2024;19(6):e0306195. [doi: 10.1371/journal.pone.0306195] [Medline: 38917147]
2. Feldstein DA, Hess R, McGinn T, et al. Design and implementation of electronic health record integrated clinical prediction rules (iCPR): a randomized trial in diverse primary care settings. Implement Sci 2017 Mar 14;12(1):37. [doi: 10.1186/s13012-017-0567-y] [Medline: 28292304]
3. Mishuris RG, Palmisano J, McCullagh L, et al. Using normalisation process theory to understand workflow implications of decision support implementation across diverse primary care settings. BMJ Health Care Inform 2019 Oct;26(1):e100088. [doi: 10.1136/bmjhci-2019-100088] [Medline: 31630113]
4. Mann DM, Kannry JL, Edonyabo D, et al. Rationale, design, and implementation protocol of an electronic health record integrated clinical prediction rule (iCPR) randomized trial in primary care. Implement Sci 2011 Sep 19;6:109. [doi: 10.1186/1748-5908-6-109] [Medline: 21929769]
5. Mann D, Hess R, McGinn T, et al. Impact of clinical decision support on antibiotic prescribing for acute respiratory infections: a cluster randomized implementation trial. J Gen Intern Med 2020 Nov;35(Suppl 2):788-795. [doi: 10.1007/s11606-020-06096-3] [Medline: 32875505]
6. Mann D, Hess R, McGinn T, et al. Adaptive design of a clinical decision support tool: what the impact on utilization rates means for future CDS research. Digit Health 2019;5:2055207619827716. [doi: 10.1177/2055207619827716] [Medline: 30792877]
7. Shaw RJ, McDuffie JR, Hendrix CC, et al. Effects of nurse-managed protocols in the outpatient management of adults with chronic conditions: a systematic review and meta-analysis. Ann Intern Med 2014 Jul 15;161(2):113-121. [doi: 10.7326/M13-2567] [Medline: 25023250]
8. Wieben AM, Alreshidi BG, Douthit BJ, et al. Nurses' perceptions of the design, implementation, and adoption of machine learning clinical decision support: a descriptive qualitative study. J Nurs Scholarsh 2024 Jun 19;57(1):82-94. [doi: 10.1111/jnu.13001] [Medline: 38898636]
9. Luo Y, He L, Li W, Zhao J, Deng R, Zong Z. Promoting nurse participation in hospital antimicrobial stewardship: a realist review. Int J Nurs Stud 2025 Sep;169:105142. [doi: 10.1016/j.ijnurstu.2025.105142] [Medline: 40561812]
10. Stevens ER, Xu L, Kwon J, et al. Barriers to implementing registered nurse-driven clinical decision support for antibiotic stewardship: retrospective case study. JMIR Form Res 2024 May 23;8:e54996. [doi: 10.2196/54996] [Medline: 38781006]
11. Dauber-Decker KL, Feldstein D, Hess R, et al. Snowball group usability testing for rapid and iterative multisite tool development: method development study. JMIR Form Res 2025 Feb 18;9:e55316. [doi: 10.2196/55316] [Medline: 39964400]

XSL•FO
RenderX

12. Stevens ER, Agbakoba R, Mann DM, et al. Reducing prescribing of antibiotics for acute respiratory infections using a frontline nurse-led EHR-Integrated clinical decision support tool: protocol for a stepped wedge randomized control trial. BMC Med Inform Decis Mak 2023 Nov 14;23(1):260. [doi: 10.1186/s12911-023-02368-0] [Medline: 37964232]
13. McCullagh L, Mann D, Rosen L, Kannry J, McGinn T. Longitudinal adoption rates of complex decision support tools in primary care. Evid Based Med 2014 Dec;19(6):204-209. [doi: 10.1136/ebmed-2014-110054] [Medline: 25238769]
14. Qu H, Hu X, Singh JA. Factors influencing implementation of a computerized, individualized, culturally tailored lupus decision aid in lupus clinics: a qualitative semi-structured interview study. Clin Rheumatol 2019 Oct;38(10):2793-2801. [doi: 10.1007/s10067-019-04643-w] [Medline: 31222574]
15. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
16. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
17. Dedoose. URL: www.dedoose.com [accessed 2026-02-19]
18. Booth A, Hannes K, Harden A, Noyes J, Harris J, Tong A. COREQ (consolidated criteria for reporting qualitative studies). In: Moher D, Altman DG, Schulz KF, Simera I, Wager E, editors. Guidelines for Reporting Health Research A User's Manual: John Wiley & Sons; 2014:214-226. [doi: 10.1002/9781118715598.ch21]
19. Cowie J, Nicoll A, Dimova ED, Campbell P, Duncan EA. The barriers and facilitators influencing the sustainability of hospital-based interventions: a systematic review. BMC Health Serv Res 2020 Jun 28;20(1):588. [doi: 10.1186/s12913-020-05434-9] [Medline: 32594912]
20. Powell BJ, Proctor EK, Glass JE. A systematic review of strategies for implementing empirically supported mental health interventions. Res Soc Work Pract 2014 Mar 1;24(2):192-212. [doi: 10.1177/1049731513505778] [Medline: 24791131]
21. Haas SA, Vlasses F, Havey J. Developing staffing models to support population health management and quality outcomes in ambulatory care settings. Nurs Econ 2016;34(3):126-133. [Medline: 27439249]
22. Aiken LH, Lasater KB, Sloane DM, et al. Physician and nurse well-being and preferred interventions to address burnout in hospital practice: factors associated with turnover, outcomes, and patient safety. JAMA Health Forum 2023 Jul 7;4(7):e231809. [doi: 10.1001/jamahealthforum.2023.1809] [Medline: 37418269]
23. Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. JAMA 2002;288(16):1987-1993. [doi: 10.1001/jama.288.16.1987] [Medline: 12387650]
24. Adynski GI, Sherwood G, Ikharo E, Tran A, Jones CB. Outpatient nurse staffing relationship with organizational, nurse and patient outcomes: a scoping review. Int J Nurs Stud Adv 2022 Dec;4:100064. [doi: 10.1016/j.ijnsa.2022.100064] [Medline: 38745630]
25. Khairat S, Morelli J, Qin Q, et al. Effect of virtual nursing implementation on emergency department efficiency and quality of care. Am J Emerg Med 2025 May;91:59-66. [doi: 10.1016/j.ajem.2025.02.024] [Medline: 39987629]
26. Webster JS, King HB, Toomey LM, et al. Understanding quality and safety problems in the ambulatory environment: seeking improvement with promising teamwork tools and strategies. In: Advances in Patient Safety: New Directions and Alternative Approaches (Volume 3: Performance and Tools): Agency for Healthcare Research and Quality (US); 2008:211-223 URL: https://www.ncbi.nlm.nih.gov/books/NBK43683/ [accessed 2026-02-19]
27. Shi Y, Amill-Rosario A, Rudin RS, et al. Barriers to using clinical decision support in ambulatory care: do clinics in health systems fare better? J Am Med Inform Assoc 2021 Jul 30;28(8):1667-1675. [doi: 10.1093/jamia/ocab064] [Medline: 33895828]
28. Nursing Informatics Scope and Standards of Practice, 2nd edition: American Nurses Association; 2015. URL: https://www.homeworkforyou.com/static_media/uploadedfiles/1729564993_4793663__166..pdf [accessed 2026-02-19]
29. Fusiak J, Sarpari K, Ma I, Mansmann U, Hoffmann VS. Practical applications of methods to incorporate patient preferences into medical decision models: a scoping review. BMC Med Inform Decis Mak 2025 Mar 3;25(1):109. [doi: 10.1186/s12911-025-02945-5] [Medline: 40033306]
30. Hlebichuk J, Lancaster RJ, Vizgirda V, Quinlan S. Top-of-license practice for registered nurses: a scoping review. Nurs Outlook 2025;73(3):102397. [doi: 10.1016/j.outlook.2025.102397] [Medline: 40250044]

Abbreviations

ARI: acute respiratory infection
CDS: clinical decision support
CFIR: Consolidated Framework for Implementation Research
CPR: clinical prediction rules
EHR: electronic health record
iCPR: integrated clinical prediction rules
RCT: randomized controlled trial
RN : Registered Nurse

XSL•FO
RenderX

Edited by E Borycki; submitted 04.Sep.2025; peer-reviewed by AN Okere, K Connell; revised version received 13.Jan.2026; accepted 20.Jan.2026; published 10.Mar.2026.

Please cite as:
Tiase VL, Tovar AD, Henning N, Braga M, McHugh K, Xu L, Bah H, Yuroff A, Hess R, Mann DM, Feldstein D, Stevens ER
Barriers and Enablers for Sustaining Nurse-Led Use of Clinical Decision Support Tools for Antibiotic Stewardship: Qualitative Study
JMIR Nursing 2026;9:e83567
URL: https://nursing.jmir.org/2026/1/e83567
doi:10.2196/83567

© Victoria L Tiase, Aurie D Tovar, Natalie Henning, Mariana Braga, Keelin McHugh, Lynn Xu, Haddy Bah, Alice Yuroff, Rachel Hess, Devin M Mann, David Feldstein, Elizabeth R Stevens. Originally published in JMIR Nursing (https://nursing.jmir.org), 10.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association Between Electronic Health Record–Based Nursing Workload and Turnover: Retrospective Cohort Study

Linlin Xia[1], MS; Daphne Lew[2], MPH, PhD; Lindsay Tessmer[3], BSN, RN; Elise Eiden[4], MSW; Sunny Lou[4], MD, PhD; Thomas Kannampallil[1,4,5], PhD

[1]Division of Computational and Data Sciences, Washington University in St. Louis, St. Louis, MO, United States

[2]Institute for Informatics, Data Science and Biostatistics, Washington University School of Medicine in St Louis, St Louis, MO, United States

[3]Barnes-Jewish College, St. Louis, MO, United States

[4]Department of Anesthesiology, Washington University School of Medicine in St. Louis, Washington University in St. Louis, 660 South Euclid Avenue, Campus Box 8054, St. Louis, MO, United States

[5]Department of Computer Science and Engineering, Washington University in St Louis, St Louis, MO, United States

Corresponding Author:
Thomas Kannampallil, PhD
Division of Computational and Data Sciences, Washington University in St. Louis, St. Louis, MO, United States

Abstract

Background: Nurse turnover remains a major challenge for health systems, yet objective, scalable measures of workload that predict turnover are limited. Electronic health record (EHR) audit logs offer a potential data source to quantify nursing work patterns.

Objective: This study aimed to evaluate the association between EHR-derived measures of nursing workload and turnover among inpatient nurses.

Methods: We analyzed work-related activities of staff nurses from medical and surgical inpatient units at a large academic medical center from January 1, 2022, to December 31, 2022. Data included deidentified demographics (age, sex, years since licensure, and service group), shift characteristics (number of shifts worked, proportion of night shifts, timing, and location), and measures of work activities derived from EHR audit logs. Audit logs were used to develop measures related to nurse workload, including (1) nurse activities (information review, medication administration, alert management, navigation, documentation, and communication), (2) patient load (based on the number of unique patient charts accessed), and (3) cognitive load (based on the number of patient switches). For nurses who left, we excluded the 6 weeks immediately preceding termination (washout period) and measured workload during the preceding 6-week period; for those who remained, a random 6-week working period was selected. Associations between workload measures and turnover were assessed using mixed-effects logistic regression, adjusting for demographics and shift-related characteristics.

Results: Among 432 nurses (n=363, 84% female; median age 27, IQR 23-36 years), contributing 6812 shifts and approximately 13 million audit log actions, 84 (19%) left the institution in 2022. A higher proportion of medication administration actions was associated with higher odds of turnover (odds ratio [OR] 2.20, 95% CI 1.36-3.54), whereas greater alert engagement (OR 0.48, 95% CI 0.32-0.72) and more years since licensure (OR 0.57, 95% CI 0.38-0.84) were associated with lower odds of turnover.

Conclusions: EHR-derived workload measures, particularly greater medication administration burden and lower alert engagement, were independently associated with the risk of nurse turnover. Patterns of EHR use may help identify nurses at higher risk of leaving an institution and can potentially inform targeted workforce retention strategies.

(JMIR Nursing 2026;9:e89645) doi:10.2196/89645

KEYWORDS

electronic health records; nursing workload; clinical informatics; nurse turnover; workforce retention

Introduction

Nurse turnover has increased exponentially over the past decade, posing a threat to health system stability. In the United States, annual turnover rates rose from approximately 18% in 2016 to approximately 29% immediately following the COVID-19 pandemic [1-3]. Globally, the scale of the problem is equally alarming, with similarly high rates of nurses reporting intentions to leave their organizations [4-6]. High nursing turnover destabilizes care teams, increases medication errors, reduces quality of care, and imposes substantial operational and financial burdens on health care systems [3].

Despite the considerable challenges associated with nurse turnover, its underlying drivers remain poorly understood. Prior

XSL•FO
RenderX

work has often linked nurse turnover to factors such as demographics, organizational dynamics [7,8], and workload [1,2,9], primarily relying on observational methods such as surveys, interviews, and focus groups [10-12]. For example, Zheng et al [6] used responses to questionnaires to assess factors such as professional fulfillment, emotional exhaustion, and self-efficacy as contributors to turnover. Other studies have used similar methods to characterize the contributors to nurse turnover and have found high perceived workload, low workplace commitment, poor team collaboration, and lack of recognition from supervisors to be key contributors to nurse burnout and intentions to leave [13,14]. Although these studies provide considerable insights, they depend heavily on self-reported assessments, require significant human effort, and are subject to recall bias, response bias, and limited scalability [1].

Electronic health record (EHR) audit logs offer a scalable approach to quantify nursing workload, capturing time-stamped data on EHR use, including documentation, information review, medication administration, messaging activities, and even indicators of cognitive load such as patient switches and on-screen interruptions [15-17]. However, despite the availability of detailed EHR interaction data from audit logs, little is known about how measures of day-to-day nursing work activities relate to turnover.

In this study, we address this gap by investigating the association between EHR-based measures of nursing workload and subsequent turnover among inpatient nurses at a large academic medical center. By linking nurse work activity metrics with real-world employment outcomes, we aim to provide new insights into how daily clinical work patterns may relate to nurses' decisions to leave their roles.

Methods

Study Setting, Participants, and Inclusion Criteria

This retrospective cohort study included all staff nurses who worked at least 1 clinical shift on a medical or surgical inpatient unit at a large academic medical center in St. Louis, Missouri, between January 1, 2022, and December 31, 2022. Temporary or visiting nursing staff and individuals not officially employed by the hospital were excluded.

Ethical Considerations

This study was approved by the Washington University Institutional Review Board (IRB#202306138) with a waiver of informed consent. All data were deidentified prior to analysis, and no individually identifiable information was available to the study team. Appropriate measures were taken to protect participant privacy and maintain data confidentiality throughout the study. Participants were not directly involved in the research and did not receive any compensation.

Data Sources

We obtained deidentified data on eligible nurses from three primary sources: (1) employment records, which provided demographic and employment details (age, sex, years since nursing licensure, duration of employment, and turnover status [ie, whether the nurse left the hospital system in 2022]); (2) nurse shift records, which included daily punch-in and punch-out times, nursing unit (eg, general medicine unit, postprocedure unit, or general surgical unit), level of care (eg, acute, intermediate care, or observation), and clinical service group (medical or surgical); and (3) EHR-based audit logs extracted from Epic (Verona, Wisconsin) Clarity tables (ACCESS_LOG), which provide time-stamped records of nurses' EHR interactions and the corresponding patient charts.

Experimental Design

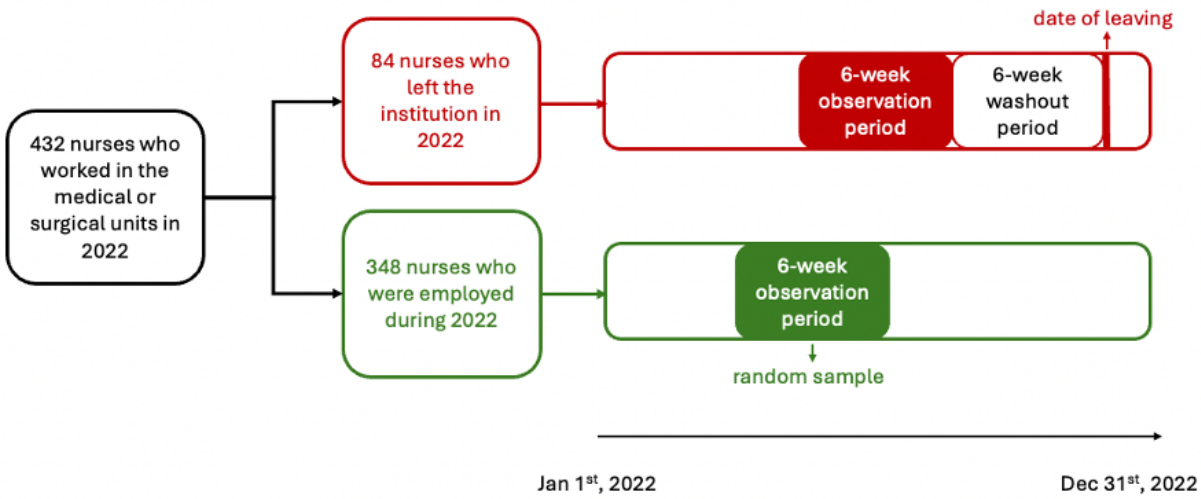
To examine the relationship between EHR-based measures of nurse workload and turnover, we classified nurses into 2 groups: those who left the institution during the 2022 calendar year ("left" group) and those who remained employed throughout the year ("stay" group).

To ensure temporal comparability between groups, we constructed a standardized 6-week observation window for each nurse. The choice of the 6-week period was based on discussions with nursing collaborators across various units rather than on specific empirical evidence.

For nurses who left, we imposed a 6-week washout period before their termination date to avoid capturing behavioral changes occurring after the decision to leave. We then analyzed shifts during the 6-week window immediately preceding the washout period (Figure 1; the red vertical bar marks the date the nurse left, and the filled red bar shows the observation window).

XSL•FO
RenderX

Figure 1. Overall study design. Nurses were grouped into those who were employed at the end of the calendar year (2022) and those who left the institution. For those who left the institution, the date of leaving was used to define a 6-week washout period, and the 6 weeks prior to the washout period were used as the "data period" to calculate the workload metrics. For those who stayed, a random 6-week period during the calendar year was selected for workload metric calculation.



For nurses who remained employed, we randomly selected a 6-week working period during the year-long study period (Figure 1; the filled green bar shows the observation window). This approach ensured consistent observation lengths while minimizing temporal bias and reducing the likelihood that observed behaviors reflected impending departure rather than typical workload.

Data Processing

Each audit log entry included a time stamp, user identifier, patient identifier, and the specific EHR action performed, allowing us to quantify both the frequency and the type of EHR interactions for each nurse during each shift. Audit log events were linked to each nurse's shift by matching time stamps to staffing punch-in and punch-out records, which defined the temporal boundaries for measurement for each shift. To characterize individual nurse workload patterns, we derived shift-level metrics and then summarized these across shifts within each nurse's observation window (ie, 6 weeks).

We first used punch-in and punch-out time records to calculate overall schedule-related metrics for each nurse, including the total number of shifts worked during the 6-week observation period and the proportion of shifts occurring during overnight hours (defined as shifts with punch-in or punch-out times between 7 PM and 7 AM; Figures S1 and S2 in Multimedia Appendix 1). These variables were included to adjust for potential differences in workload and turnover risk associated with shift patterns.

Workload metrics included three domains: (1) EHR-based nursing activities, (2) patient load (with the number of unique patient charts accessed per shift used as a proxy), and (3) cognitive effort (with patient-switching actions used as a proxy) [18,19].

To characterize the composition of EHR-based nursing activities, we relied on prior time-and-motion studies that provided detailed insights into nursing workflows in the EHR [20]. The categorization framework was developed through an iterative process involving audit log experts and clinical domain experts (coauthors LT and SL) familiar with clinical and nursing workflows. Audit log actions were initially mapped to functional task types using vendor documentation and institutional data dictionaries and then refined through multiple rounds of structured clinician review to ensure face validity and interpretability.

Specifically, we categorized EHR actions into 6 EHR-based nursing activities: information review (eg, viewing patient reports, storyboard, and clinical notes); medication administration (eg, barcode scanning, barcode processing, and reviewing administration data); documentation (eg, interacting with flow sheets and clinical forms); workflow navigation (eg, accessing the inpatient system list and loading patient lists); active alert management (eg, acknowledging best practice advisories, reviewing medication warnings, or taking recommended actions); and communication (eg, secure chat communication; refer to Table 1 for category descriptions). Table S3 in Multimedia Appendix 1 provides the complete actions of audit log mapping to activity categories.

XSL•FO
RenderX

Table . Nurse workload categories, their description, and associated example metrics from audit logs[a].

Categories	Description	Example metrics (from audit logs)
Information review	Encompasses nurse activities involving the examination and evaluation of patient-specific data and associated clinical documentation	• Report with patient data viewed • Storyboard viewed • Flow sheet viewed
Medication administration	Involves clinical workflow and procedural tasks required for the administration of medications	• Barcode scanned • MAR[b] barcode processed • MAR administration viewed
Documentation	Activities focused on entering, updating, or interacting with clinical documentation	• Flow sheet accepted • Flow sheet data copied forward • Inpatient work list task edited
Navigation	Actions related to moving through the Epic system to access different functions or patient records	• Visit a loaded Navigator template • Inpatient system list accessed • Radar dashboard accessed
Active alerts	Interactions with system-generated alerts and clinical decision support advisories	• Best practice advisories acknowledged • Medication warning displayed • Patient chart advisories viewed
Communication	Activities using secure messaging apps to facilitate communication and care coordination among the health care team	• Secure chat conversation opened • Secure chat conversation created • Secure chat activity accessed

[a]A full list of the audit logs for each category is provided in Multimedia Appendix 1.

[b]MAR: Medication Administration Record.

We calculated the number of patient switches per shift as a proxy for cognitive effort, reflecting the frequency with which nurses transitioned between patient charts during active EHR use [15,17]. Consistent with prior work, a patient switch was defined as a transition from one patient's chart to another within 5 minutes, to reflect active transitions rather than idle time [17].

Additional workload measures included the total number of EHR actions per shift, the number of unique patient charts accessed per shift (ie, patient load), and time spent using the EHR per shift. Time spent using the EHR, measured in minutes, was calculated from audit log data by summing the time intervals between successive audit log actions, excluding intervals exceeding 5 minutes, which were considered periods of inactivity [18,19].

To summarize these workload indicators at the individual nurse level, we aggregated each nurse's total task-based EHR actions, unique patient charts accessed, and patient switches across the entire 6-week observation period. Specifically, we calculated the proportion of each nurse's EHR actions falling into each activity category by dividing task-specific action counts by the nurse's total EHR actions. For shift-level workload frequency measures, we computed the median value of each metric across all observed shifts for each nurse. Together, these task-based EHR activity proportions and shift-level workload frequency measures served as the primary exposure variables, capturing individual differences in overall EHR-based workload.

Covariates in the regression models included demographic characteristics and shift-related patterns. Demographic characteristics included age, sex, years since licensure, and primary clinical service group. Shift-based factors included the total number of shifts worked and the proportion of night shifts.

All data processing and feature engineering were conducted in Python (version 3.9.10; Python Software Foundation) using the following libraries: Pandas (version 2.2.2) for data manipulation, Matplotlib (version 3.10.3) and Seaborn (version 0.13.2) for visualization, and Scikit-learn (version 1.6.1) and Statsmodels (version 0.14.4) for data preparation [21-26].

Outcome

The primary outcome was nurse turnover, defined as a binary variable indicating whether the nurse left the organization during the 2022 calendar year.

Statistical Analysis

Descriptive statistics were calculated as medians and IQRs or frequencies and percentages. Univariable comparisons between nurses who stayed and those who left were conducted using the Wilcoxon rank-sum test for continuous variables and the chi-square test for categorical variables. Because task-specific action categories are compositional and collectively sum to approximately 100%, we conducted a collinearity analysis using variance inflation factors to assess potential multicollinearity among these variables (Table S1 in Multimedia Appendix 1). On the basis of this assessment and input from clinician experts, the workflow navigation category demonstrated substantial collinearity with other task categories, particularly information review, and was therefore excluded from multivariable modeling to improve model stability and interpretability; information review was retained given its clinical relevance.

XSL•FO
RenderX

To assess the association between EHR-based nursing workload measures and turnover, we used a mixed-effects logistic regression model with random intercepts for nursing units to account for the contribution of each unit's culture and work environment to turnover. The statistical model adjusted for demographic factors (age, sex, years since licensure, and service group), shift characteristics (number of shifts worked and proportion of night shifts), and workload intensity (total EHR actions, patient switches, and number of patients cared for). All continuous predictors were standardized using IQR scaling prior to modeling. Thus, reported odds ratios (ORs) and 95% CIs represent the association with a 1-IQR increase (from the 25th to the 75th percentile) in each continuous predictor.

To assess robustness, we conducted a sensitivity analysis using a shorter 2-week washout period prior to each nurse's departure. Statistical analyses were performed using R (R Foundation for Statistical Computing) with the lme4 package [27].

Results

Sample Characteristics

The analytic sample included 432 inpatient nurses who collectively contributed 6812 shifts and more than 13.8 million EHR audit log actions during the study period (Table 2). Of these nurses, 84 (19%) left the institution (Figure 1). The sample was predominantly female (84%), with a median age of 27 (IQR 23-36) years. In total, 60% (n=259) of the nurses worked on surgical services and 40% (n=173) worked on medical services.

XSL•FO
RenderX

Table . Study cohort characteristics stratified by turnover status (N=432).

Turnover status in 2022	All participating nurses	Nurses who stayed (n=348)	Nurses who left (n=84)	*P* value
Number of shifts	6812	5543	1269	N/A[a]
Number of audit log actions	13,807,571	11,228,809	2,578,762	N/A
Age (years), median (IQR)	27 (23-36)	28 (23-37)	26 (24-31)	.58
Sex, n (%)				.67
Female	364 (84)	295 (85)	69 (82)	
Male	68 (16)	53 (15)	15 (18)	
Clinical service group, n (%)				.60
Surgical unit	259 (60)	206 (59)	53 (63)	
Medicine unit	173 (40)	142 (41)	31 (37)	
Years since licensure, median (IQR)	2 (1-6)	3 (1-9)	1 (1-3)	.01[b]
Shift characteristics, median (IQR)				
Shift counts	17 (12-20)	17 (12-20)	17 (13-19)	.12
Proportion of night shifts (%)	10 (0-94)	11 (0-94)	7 (0-94)	.75
EHR[c] workload intensity, median (IQR)				
EHR time per shift time (%)	33.87 (28.67-39.09)	33.87 (28.57-39.52)	33.85 (28.93-37.99)	.82
EHR time per shift (minute)	260 (216-302)	261 (215-305)	257 (217-289)	.54
Patient charts accessed per shift	6 (5-9)	6 (5-9)	6 (5-9)	.82
Patient switches per shift	61 (44-82)	60 (44-85)	64 (46-80)	.50
EHR actions per shift	2088 (1464-2723)	2122 (1430-2751)	1964 (1566-2595)	.99
EHR task composition (percentage of total EHR actions), median (IQR)				
Documentation	5.96 (4.75-8.10)	5.91 (4.69-8.15)	6.16 (5.02-7.94)	.38
Medication administration	8.75 (6.65-11.98)	8.54 (6.44-11.69)	9.61 (7.41-12.49)	.02[b]
Navigation	9.41 (5.09-15.50)	9.31 (4.86-15.66)	9.84 (5.71-14.41)	.82
Information review	69.55 (57.30-74.58)	69.58 (57.72-74.84)	69.42 (56.16-74.02)	.67
Active alert	0.16 (0.09-0.27)	0.17 (0.09-0.31)	0.14 (0.06-0.19)	.01[b]
Communication	0.73 (0.36-1.38)	0.71 (0.36-1.35)	0.83 (0.43-1.38)	.31

[a]N/A: not applicable.

[b]Indicates statistical significance at the $P<.05$ level.

[c]EHR: electronic health record.

Nurses who left the organization were younger on average (median 26 vs 28 years), although this difference was not statistically significant. They had substantially fewer years since licensure (median 1, IQR 1 - 3 years vs median 3, IQR 1 - 9 years). Sex distribution and clinical service group were similar between groups.

Univariable Analysis

Shift patterns did not differ meaningfully between nurses who stayed and those who left, with a comparable number of shifts worked during the observation window (median 17 in both groups) and similar proportions of night shifts.

Overall EHR workload intensity—including median EHR time per shift, proportion of shift time spent in the EHR, total EHR actions, number of unique patient charts accessed, and number of patient switches—was nearly identical between groups (all $P>.30$).

However, 2 EHR task composition measures differed significantly. Nurses who left had a higher proportion of medication administration actions (median 9.61%, IQR 7.41% - 12.49% vs median 8.54%, IQR 6.44% - 11.69%; $P=.02$) and a lower proportion of active alert engagement (median 0.14%, IQR 0.06% - 0.19% vs median 0.17%, IQR 0.09% - 0.31%; $P=.01$). No significant differences were

XSL•FO
RenderX

observed across other EHR activity categories, including documentation, information review, navigation, or communication (Table 2).

Multivariable Analysis

In mixed-effects logistic regression models adjusting for demographic characteristics, shift patterns, and workload intensity, task composition emerged as the primary workload-related predictor of turnover (Table 3). An increase in the proportion of medication administration actions from the 25th percentile (6.65%) to the 75th percentile (11.98%) was associated with more than twice the odds of leaving the institution (adjusted OR 2.20, 95% CI 1.36 - 3.54). Conversely, an increase in active alert engagement from the 25th percentile (0.09%) to the 75th percentile (0.27%) was associated with substantially lower odds of turnover (adjusted OR 0.48, 95% CI 0.32 - 0.72).

Table . Multivariable analysis showing the association between nursing workload components, demographics, work-related aspects, and turnover.

Variables	Scaled odds ratio (95% CI)	*P* value
Age (years; as of 2022)	1.04 (0.61 - 1.79)	.88
Sex (male vs female)	0.79 (0.39 - 1.59)	.50
Clinical service group (surgical vs medicine)	1.05 (0.53 - 2.07)	.90
Years since licensure	0.57 (0.38 - 0.84)	<.01[a]
Number of shifts worked	0.79 (0.56 - 1.12)	.18
Proportion of night shifts	0.89 (0.47 - 1.67)	.72
Median EHR[b] actions per shift	1.43 (0.80 - 2.54)	.23
Median patient charts accessed per shift	1.07 (0.87 - 1.32)	.53
Median patient switches per shift	0.93 (0.64 - 1.35)	.70
Information review (percentage of total EHR actions)	0.93 (0.44 - 1.96)	.84
Medication administration (percentage of total EHR actions)	2.20 (1.36 - 3.54)	<.01[a]
Documentation (percentage of total EHR actions)	0.95 (0.70 - 1.31)	.77
Active alert (percentage of total EHR actions)	0.48 (0.32 - 0.72)	<.01[a]
Communication (percentage of total EHR actions)	0.96 (0.61 - 1.51)	.86

[a]Indicates statistical significance at the $P<.05$ level.

[b]EHR: electronic health record.

Years since licensure also remained significant; an increase from the 25th (1 year) to the 75th percentile (6 years) was associated with lower odds of leaving (adjusted OR 0.57; 95% CI 0.38-0.84). No other demographic variables, shift characteristics, or workload intensity measures demonstrated significant associations with turnover.

Sensitivity Analysis

A sensitivity analysis using a 2-week washout period before departure produced results consistent with the primary analysis, supporting the robustness of the associations identified (Table S2 in Multimedia Appendix 1).

Discussion

In this study, we investigated the association between EHR-based nursing workload and the likelihood of leaving the institution. We found that fewer years since nursing licensure and performing a higher fraction of medication administration actions were associated with higher odds of turnover. We also found that an increased fraction of active EHR alerts was associated with lower odds of turnover. Taken together, these findings have implications for understanding and potentially redesigning EHR-based nursing workload.

There are several possible explanations for the finding that a higher fraction of medication administration actions was associated with higher odds of turnover. The association of medication administration and documentation with chronic stress and burden has been well-documented [28]. Medication administration is a complex, multistep process involving identifying patients due for medications, prioritizing schedules, accessing centralized dispensing systems, and individually retrieving medications. These tasks may also include secondary nurse verification or additional preparation steps, most of which are not captured in EHR audit logs. Balancing complex medication schedules and performing detailed documentation can lead to mental fatigue and a persistent sense of “catching up” on charting. Over time, this may contribute to emotional exhaustion and increase the likelihood of turnover intention or actual departure [29,30]. During medication workflows, nurses may also encounter various interruptions—from unit alarms to emergency scenarios and interprofessional requests—which further elevate workload and burden [31]. Additionally,

XSL•FO
RenderX

increased medication administration burden may serve as a proxy for higher patient complexity, with sicker patients requiring more frequent interventions, although we could not directly measure patient complexity in this study.

We also found that a higher proportion of active alert engagement was associated with lower odds of turnover. For this study, active alerts were defined as audit log events requiring explicit acknowledgment, often surfacing high-priority safety checks or patient status cues that help structure nurses' decision-making. Nurses engaged with a median of 1.5 active alerts per shift (IQR 1.0 - 3.0), indicating that these events are infrequent and unlikely to represent a significant burden. Rather, they may function as a helpful tool that enhances situational awareness, reinforces clinical priorities, and reduces cognitive load by externalizing key reminders. Therefore, nurses who engage with these alerts may experience greater clarity, support, and alignment with expected workflows, mitigating the sense of chaotic work environments that contribute to turnover. Although we could not differentiate alert types, the small absolute variation in alert fractions across clinicians (0.09% to 0.27%) likely reflects the low frequency of these high-salience events rather than differences in overall workload. Differences in alert engagement may signal how effectively nurses interact with digital workflow supports. Importantly, our model adjusted for demographics, shift type, clinical service, and overall EHR activity, making it less likely that this association reflects underlying workload differences alone. Future work characterizing alert content, cognitive load, and perceived usefulness may clarify whether active alert engagement functions as an indicator of supportive digital environments that promote clinician engagement and retention.

This study has several strengths. It was a longitudinal study that included all nurses at a large academic medical center over 1 year (432 nurses, 6812 shifts, >13 million audit log actions), providing a detailed and objective assessment of EHR-based workload. Our audit log measures captured granular shift-level behaviors, offering a scalable framework to quantify workload patterns that may influence clinician retention. Consistent with prior time-and-motion studies, we found that nurses spent approximately 34% of each shift on EHR activities, including documentation, chart review, and medication management, highlighting the substantial cognitive and administrative demands of modern nursing work [20].

This study also has several limitations. First, this was a single-center study of specific inpatient settings over 1 year. Although we had a large sample, findings may not generalize to other institutions or regions. In addition, the cohort was relatively young (median age 27 years) and in an early career stage (median 2 years since licensure), which may not reflect the broader nursing workforce. Turnover drivers in this population may differ from those in more experienced nurses (eg, career development vs long-term occupational strain), which may influence the observed associations.

Second, audit logs do not capture all aspects of nursing work—such as direct patient care, interpersonal interactions, or work with other technologies (eg, medication-dispensing systems). Thus, our workload measures reflect EHR-based workload only. However, prior research demonstrates that audit logs offer an objective measure of cognitive and administrative workload, capturing a critical aspect of modern clinical practice that can influence clinician engagement and retention [15,32].

Third, our dataset was limited to nurses directly employed by the institution; in other words, the service time only reflected their tenure as employees. If a contingent or agency nurse transitioned to direct employment, their prior service as a nonemployee was not reflected. Furthermore, agency or travel nurses who were not direct employees of the health system were not included in our sample.

Fourth, patient complexity and exact nurse-patient assignments could not be fully accounted for. Patient load was derived from EHR audit logs and operationalized as the number of unique patient charts accessed per nurse, rather than staffing assignment data. Although we adjusted for patient load using this measure and included unit type as a covariate, these proxies may not fully capture patient complexity or staffing patterns. Because nurses may access charts for patients who are not formally assigned to them, this measure may not precisely reflect true patient assignments, introducing potential measurement imprecision. Medication administration–related EHR activity may reflect higher patient complexity, medication burden, or assignment to higher-intensity units rather than workload strain alone, and residual confounding by unmeasured clinical factors (eg, patient acuity) remains possible. Direct measures of patient acuity (eg, case-mix index or high-risk medication exposure) were not available in this dataset.

Fifth, our analysis focused on nurses who fully left the organization and did not include those who transferred internally to other units. To avoid mixing data from prior assignments, our observation window was restricted to each nurse's current unit. Although internal transfers were relatively few, they may still create administrative burden, incur training costs, and pose temporary staffing challenges, which warrants further study alongside complete organizational turnover. Additionally, turnover was defined based on recorded separation from employment, and the reason for leaving (eg, voluntary resignation, termination, or contract completion) was not available, which may introduce heterogeneity in the outcome.

Sixth, workload metrics were aggregated at the nurse level using medians and proportions across shifts to align exposure measurement with a nurse-level outcome of turnover. Although this approach reduces noise in shift-to-shift variability and reflects overall workload exposure during the observation period, it may obscure within-nurse variability, including high-burden shifts. As such, infrequent but extreme workload experiences may not be fully captured. Multilevel shift-based modeling could potentially help in characterizing such patterns.

Finally, we applied a 6-week washout period, based on input from clinical collaborators, to account for the time nurses typically take to decide to leave the organization. We acknowledge that individual decision timelines vary; however, sensitivity analyses using a shorter, 2-week washout period confirmed the robustness of our findings.

XSL•FO
RenderX

Despite these limitations, this study demonstrates a scalable, data-driven approach to quantifying EHR-based nursing workload and its association with turnover. Findings highlight the potential of granular workload measures to identify patterns of digital work allocation that may relate to clinician retention. Future work should integrate direct measures of patient complexity and qualitative assessments of unit culture, interpersonal dynamics, and institutional factors, as well as validate these measures across diverse health systems, to better understand how digital workload interfaces with nurse engagement, burnout, and workforce stability.

Funding

The authors declared that no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Association between electronic health record–based nursing workload and turnover.
[DOCX File, 71 KB - nursing_v9i1e89645_app1.docx]

References

1. Woodward KF, Willgerodt M. A systematic review of registered nurse turnover and retention in the United States. Nurs Outlook 2022;70(4):664-678. [doi: 10.1016/j.outlook.2022.04.005] [Medline: 35798584]
2. Rutledge DN, Douville S, Winokur EJ. Chronic fatigue predicts hospital nurse turnover intentions. J Nurs Adm 2022 Apr 1;52(4):241-247. [doi: 10.1097/NNA.0000000000001139] [Medline: 35348490]
3. Falatah R. The impact of the coronavirus disease (COVID-19) pandemic on nurses' turnover intention: an integrative review. Nurs Rep 2021 Oct 19;11(4):787-810. [doi: 10.3390/nursrep11040075] [Medline: 34968269]
4. Dewanto A, Wardhani V. Nurse turnover and perceived causes and consequences: a preliminary study at private hospitals in Indonesia. BMC Nurs 2018;17(Suppl 2):52. [doi: 10.1186/s12912-018-0317-8] [Medline: 30574016]
5. Labrague LJ, De Los Santos JA, Falguera CC, et al. Predictors of nurses' turnover intention at one and five years' time. Int Nurs Rev 2020 Jun;67(2):191-198. [doi: 10.1111/inr.12581] [Medline: 32202329]
6. Zheng J, Feng S, Feng Y, Wang L, Gao R, Xue B. Relationship between burnout and turnover intention among nurses: a network analysis. BMC Nurs 2024 Dec 18;23(1):921. [doi: 10.1186/s12912-024-02624-2] [Medline: 39696402]
7. Bae SH. Noneconomic and economic impacts of nurse turnover in hospitals: a systematic review. Int Nurs Rev 2022 Sep;69(3):392-404. [doi: 10.1111/inr.12769] [Medline: 35654041]
8. Hu H, Wang C, Lan Y, Wu X. Nurses' turnover intention, hope and career identity: the mediating role of job satisfaction. BMC Nurs 2022 Feb 10;21(1):43. [doi: 10.1186/s12912-022-00821-5] [Medline: 35144604]
9. Pressley C, Garside J. Safeguarding the retention of nurses: a systematic review on determinants of nurse's intentions to stay. Nurs Open 2023 May;10(5):2842-2858. [doi: 10.1002/nop2.1588] [Medline: 36646646]
10. Adomat R, Hicks C. Measuring nursing workload in intensive care: an observational study using closed circuit video cameras. J Adv Nurs 2003 May;42(4):402-412. [doi: 10.1046/j.1365-2648.2003.02632.x] [Medline: 12752885]
11. Ampt A, Westbrook J, Creswick N, Mallock N. A comparison of self-reported and observational work sampling techniques for measuring time in nursing tasks. J Health Serv Res Policy 2007 Jan;12(1):18-24. [doi: 10.1258/135581907779497576] [Medline: 17244393]
12. Ivziku D, Ferramosca FM, Filomeno L, Gualandi R, De Maria M, Tartaglini D. Defining nursing workload predictors: a pilot study. J Nurs Manag 2022 Mar;30(2):473-481. [doi: 10.1111/jonm.13523] [Medline: 34825432]
13. Phillips C. Relationships between workload perception, burnout, and intent to leave among medical-surgical nurses. Int J Evid Based Healthc 2020 Jun;18(2):265-273. [doi: 10.1097/XEB.0000000000000220] [Medline: 32141948]
14. Diehl E, Rieger S, Letzel S, et al. The relationship between workload and burnout among nurses: the buffering role of personal, social and organisational resources. PLoS One 2021;16(1):e0245798. [doi: 10.1371/journal.pone.0245798] [Medline: 33481918]
15. Lou SS, Lew D, Harford DR, et al. Temporal associations between EHR-derived workload, burnout, and errors: a prospective cohort study. J Gen Intern Med 2022 Jul;37(9):2165-2172. [doi: 10.1007/s11606-022-07620-3] [Medline: 35710654]
16. Lou SS, Lew D, Xia L, Baratta L, Eiden E, Kannampallil T. Secure messaging use and wrong-patient ordering errors among inpatient clinicians. JAMA Netw Open 2024 Dec 2;7(12):e2447797. [doi: 10.1001/jamanetworkopen.2024.47797] [Medline: 39630450]
17. Lew D, Baratta LR, Xia L, et al. Association of EHR-integrated secure messaging use with clinician workload and attention switching. J Gen Intern Med 2025 Jul;40(10):2240-2247. [doi: 10.1007/s11606-025-09466-x] [Medline: 40085321]

XSL•FO
RenderX

18. Lou SS, Kim S, Harford D, et al. Effect of clinician attention switching on workload and wrong-patient errors. Br J Anaesth 2022 Jul;129(1):e22-e24. [doi: 10.1016/j.bja.2022.04.012] [Medline: 35568509]
19. Xia L, Lew D, Baratta L, Eiden E, Lou SS, Kannampallil T. Association between conversational multitasking and clinician work behaviors at a large US health care system: cohort study. J Med Internet Res 2025 Sep 9;27:e72768. [doi: 10.2196/72768] [Medline: 40925645]
20. Yen PY, Kellye M, Lopetegui M, et al. Nurses' time allocation and multitasking of nursing activities: a time motion study. AMIA Annu Symp Proc 2018;2018:1137-1146. [Medline: 30815156]
21. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825-2830. [doi: 10.5555/1953048.2078195]
22. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with Python. Presented at: 9th Python in Science Conference; Jun 28 to Jul 3, 2010. [doi: 10.25080/Majora-92bf1922-011]
23. Waskom ML. seaborn: statistical data visualization. J Open Source Softw 2021;6(60):3021. [doi: 10.21105/joss.03021]
24. Package overview. Pandas. URL: https://pandas.pydata.org/docs/getting_started/overview.html [accessed 2025-06-18]
25. McKinney W. Data structures for statistical computing in Python. Presented at: 9th Python in Science Conference; Jun 28 to Jul 3, 2010. [doi: 10.25080/Majora-92bf1922-00a]
26. Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng 2007;9(3):90-95. [doi: 10.1109/MCSE.2007.55]
27. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw 2015;67(1):1-48. [doi: 10.18637/jss.v067.i01]
28. Gesner E, Dykes PC, Zhang L, Gazarian P. Documentation burden in nursing and its role in clinician burnout syndrome. Appl Clin Inform 2022 Oct;13(5):983-990. [doi: 10.1055/s-0042-1757157] [Medline: 36261113]
29. Friganović A, Selič P, Ilić B, Sedić B. Stress and burnout syndrome and their associations with coping and job satisfaction in critical care nurses: a literature review. Psychiatr Danub 2019 Mar;31(Suppl 1):21-31. [Medline: 30946714]
30. Strudwick G, Jeffs L, Kemp J, et al. Identifying and adapting interventions to reduce documentation burden and improve nurses' efficiency in using electronic health record systems (The IDEA Study): protocol for a mixed methods study. BMC Nurs 2022 Aug 4;21(1):213. [doi: 10.1186/s12912-022-00989-w] [Medline: 35927701]
31. Sassaki RL, Cucolo DF, Perroca MG. Interruptions and nursing workload during medication administration process [Article in English, Portuguese]. Rev Bras Enferm 2019 Aug 19;72(4):1001-1006. [doi: 10.1590/0034-7167-2018-0680] [Medline: 31432958]
32. Kannampallil T, Adler-Milstein J. Using electronic health record audit log data for research: insights from early efforts. J Am Med Inform Assoc 2022 Dec 13;30(1):167-171. [doi: 10.1093/jamia/ocac173] [Medline: 36173351]

Abbreviations

EHR: electronic health record
OR: odds ratio

Edited by E Borycki; submitted 15.Dec.2025; peer-reviewed by Q Zhou, V Tiase; revised version received 23.Mar.2026; accepted 21.Apr.2026; published 12.Jun.2026.

Please cite as:
Xia L, Lew D, Tessmer L, Eiden E, Lou S, Kannampallil T
Association Between Electronic Health Record–Based Nursing Workload and Turnover: Retrospective Cohort Study
JMIR Nursing 2026;9:e89645
URL: https://nursing.jmir.org/2026/1/e89645
doi:10.2196/89645

© Linlin Xia, Daphne Lew, Lindsay Tessmer, Elise Eiden, Sunny Lou, Thomas Kannampallil. Originally published in JMIR Nursing (https://nursing.jmir.org), 12.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Nurse-Led Ambient AI Scribe for Patient Safety Incident Investigation Reports (Project NARRATE): Retrospective Pre-Post Comparative Document-Quality Study

Kai Yunn Teo[*], MPSHQ; Liwen Huang[*], MPSHQ; Kelly Chai Yuen Woh[*], BSc; Shin Yuh Ang[*], MBA; Jade Gaik Nai Ng, MHlthServMgnt

Department of Nursing Administration, Singapore General Hospital, Outram Road, Singapore, Singapore, Singapore

[*]these authors contributed equally

Corresponding Author:
Kai Yunn Teo, MPSHQ
Department of Nursing Administration, Singapore General Hospital, Outram Road, Singapore, Singapore, Singapore

Abstract

Background: Patient safety investigation reports support organizational learning only when they are complete, usable, and sufficiently detailed. Conventional free-text reports are often inconsistent and may omit information needed for review and learning. Project NARRATE (Nursing AI-Refined for Accurate Transcription of Events) is a nursing-led ambient artificial intelligence workflow that uses prompts aligned with the World Health Organization Minimal Information Model for Patient Safety Incident Reporting and Learning Systems, Situation-Background-Assessment-Recommendation output, and visible missing-information cues to support structured supervisor reporting.

Objective: This study aimed to compare the completeness and narrative quality of conventional and NARRATE-period supervisor investigation reports for falls and medication administration-related incidents.

Methods: We conducted a retrospective pre-post document-quality study at a tertiary academic medical center in Singapore. We reviewed 150 deidentified completed supervisor investigation reports: 75 conventional reports from June to August 2025 and 75 confirmed NARRATE reports from January to March 2026. NARRATE use was voluntary, and recorded use represented approximately 40% of eligible postimplementation reports. Two blinded reviewers rated reports using a World Health Organization (WHO)–aligned completeness checklist and an adapted 8-domain Physician Documentation Quality Instrument (PDQI). Report-level comparisons were adjusted for repeated reports by the same supervisor using random-intercept linear mixed-effects models. A stratified 60-report plain-paragraph rerating examined whether visible structure influenced ratings.

Results: All 150 reports were analyzed. Unadjusted mean WHO total completeness was 11.81 (SD 3.39) for conventional reports and 13.61 (SD 2.54) for NARRATE reports; the unadjusted difference was 1.80 points, and the cluster-adjusted mean difference was 1.95 (95% CI 0.91 - 3.00; P<.001). The adapted PDQI mean was 3.61 (SD 0.52) and 4.13 (SD 0.34), respectively; the unadjusted difference was 0.52 points, and the cluster-adjusted mean difference was 0.53 (95% CI 0.37 - 0.69; P<.001). In the plain-paragraph sensitivity analysis, the completeness advantage remained (adjusted mean difference 1.70, 95% CI 0.27 - 3.14; P=.02), as did the adapted PDQI mean advantage (adjusted mean difference 0.25, 95% CI 0.06 - 0.43; P=.009). Explanation, organization, and comprehensibility remained significantly higher after deformatting; actions were borderline (P=.05), and synthesis, internal consistency, and fairness/balance were not statistically significant.

Conclusions: Among voluntary early adopters, NARRATE use was associated with more complete reports and higher adapted PDQI mean scores after accounting for supervisor clustering. Because recorded use represented approximately 40% of eligible postimplementation reports and users self-selected, findings may reflect adopter and supervisor characteristics. Results support the structured workflow as a whole, not any single AI component, and do not demonstrate downstream patient-safety effects. Confirmatory evaluation under broader adoption with a concurrent, reliably classified comparison group is needed.

(JMIR Nursing 2026;9:e100775) doi:10.2196/100775

KEYWORDS

artificial intelligence; ambient AI scribe; nursing; patient safety; incident reporting; documentation quality; learning health system

XSL•FO
RenderX

Introduction

Background

Patient safety incident reporting systems support learning only when reports contain enough structured, interpretable information to guide investigation and improvement [1-3]. The World Health Organization (WHO) Minimal Information Model for Patient Safety Incident Reporting and Learning Systems (MIM PS) specifies the elements needed for this purpose, including what happened, why it happened, and what was done in response [2,3].

In practice, many incident reports fall short. Prior studies have found incomplete capture of causal factors, latent conditions, patient impact, and follow-up actions, limiting both case review and system-level learning [4,5]. These shortcomings are not merely documentation defects; they weaken the ability of reporting systems to generate meaningful analysis and improvement [4,5].

Reporting alone therefore does not guarantee learning [6-8]. Interventions that improve the completeness, structure, and usability of investigation narratives may strengthen the value of incident reporting for review and action.

Rationale for the Present Intervention

Ambient AI documentation tools offer one possible response. These scribes can convert spoken narrative into structured draft notes and may reduce documentation burden while improving consistency. Early evaluations of the in-house SingHealth ambient scribe program suggest acceptable note quality and reduced documentation time in outpatient care [9,10]. Whether similar tools improve nursing incident investigation reports remains unclear. To our knowledge, no prior peer-reviewed study has evaluated an ambient AI scribe explicitly aligned with the advanced WHO MIM PS for this purpose.

Study Context

Project NARRATE (Nursing AI-Refined for Accurate Transcription of Events) was developed to improve the quality of completed nursing incident-investigation reports. Built on SingHealth Note Buddy, it prompts supervisors to narrate key WHO MIM PS elements and generates a Situation-Background-Assessment-Recommendation (SBAR)–structured draft. When a safety-critical element is missing, the draft retains a visible missing-information cue for supervisor completion before final submission. The workflow was intended to support more complete, organized, and review-ready documentation while reducing manual reformatting.

We focused on completed supervisor investigation reports rather than initial frontline incident narratives because they contain the postinvestigation information most relevant to learning, including contributing factors, findings, and proposed actions [2,11].

Study Objectives and Hypothesis

This study compared the completeness and narrative quality of completed nursing supervisor investigation reports before and after implementation of the NARRATE workflow. We hypothesized that NARRATE-period reports would score higher on WHO MIM PS–aligned completeness and on an adapted 8-domain Physician Documentation Quality Instrument (PDQI)–based narrative-quality assessment [12]. The primary objective was WHO total completeness; the secondary objective was the adapted PDQI mean; additional objectives included domain-specific and interrater reliability analyses.

Methods

Study Design

We conducted a retrospective pre-post comparative document-quality study, supplemented by a format-blinded plain-text rerating sensitivity analysis in a stratified subsample to assess whether recognizable SBAR structure influenced ratings. The unit of analysis was the completed supervisor investigation report for a nursing-originated patient safety incident. The exposure was documentation method and period: conventional manual documentation before implementation versus NARRATE-supported documentation after implementation. The primary outcome was WHO total completeness, and the secondary outcome was the adapted PDQI mean. We report the study using the SQUIRE (Standards for Quality Improvement Reporting Excellence) 2.0, supplemented by relevant STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) elements.

Setting

The study was conducted at Singapore General Hospital, a tertiary academic medical center in Singapore, within a nursing-led patient safety documentation improvement initiative.

Intervention: NARRATE Supervisor Workflow

NARRATE stands for Nursing AI-Refined for Accurate Transcription of Events. The intervention used SingHealth Note Buddy, an in-house ambient AI scribe operating within approved institutional digital infrastructure. Note Buddy combines automatic speech recognition with a generative large language model, processes audio in real time, and does not retain audio recordings [10]. Users review and edit draft notes before manually transferring verified content into downstream systems [10].

For Project NARRATE, the team configured a supervisor prompt template aligned with the advanced WHO MIM PS. Supervisors narrated incident context, sequence, outcome, causes, contributing and mitigating factors, and actions; the system generated an SBAR-structured draft [13]. When key elements were missing, the draft retained a visible missing-information placeholder. Supervisors were required to verify, edit, and complete the draft before transfer into the incident reporting system. The workflow treated AI output as documentation support rather than an autonomous investigation conclusion.

AI Governance and Safety Checks

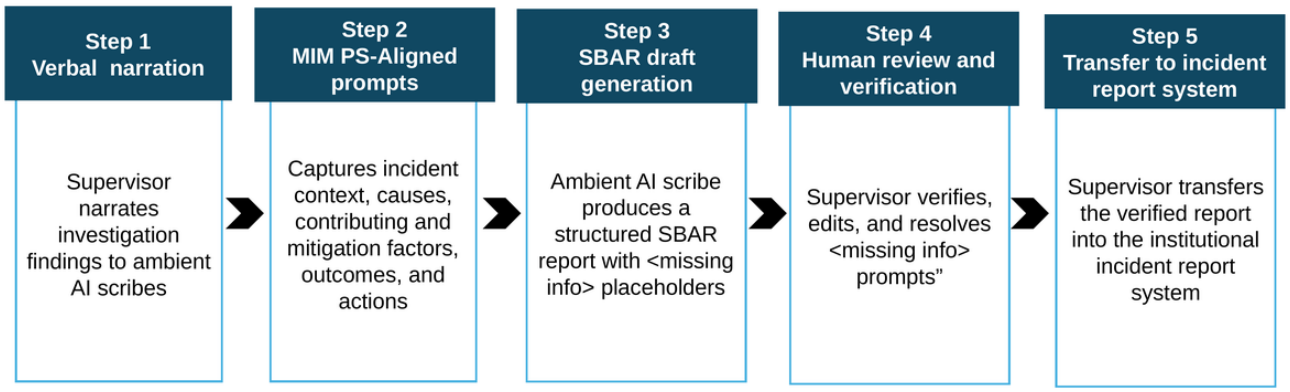
During the study period, NARRATE operated as a stable, postlaunch workflow. The supervisor prompt template was fixed before the study and unchanged throughout, and the underlying

XSL•FO
RenderX

ambient-scribe model did not change during the January to March 2026 window. The underlying platform maintained system-level audit logs of user access and draft-generation activity; these platform logs were not analyzed for this study. Every draft required mandatory supervisor verification and editing, including correction of transcription errors and resolution of all missing-information cues, before transfer into the institutional incident-reporting system.

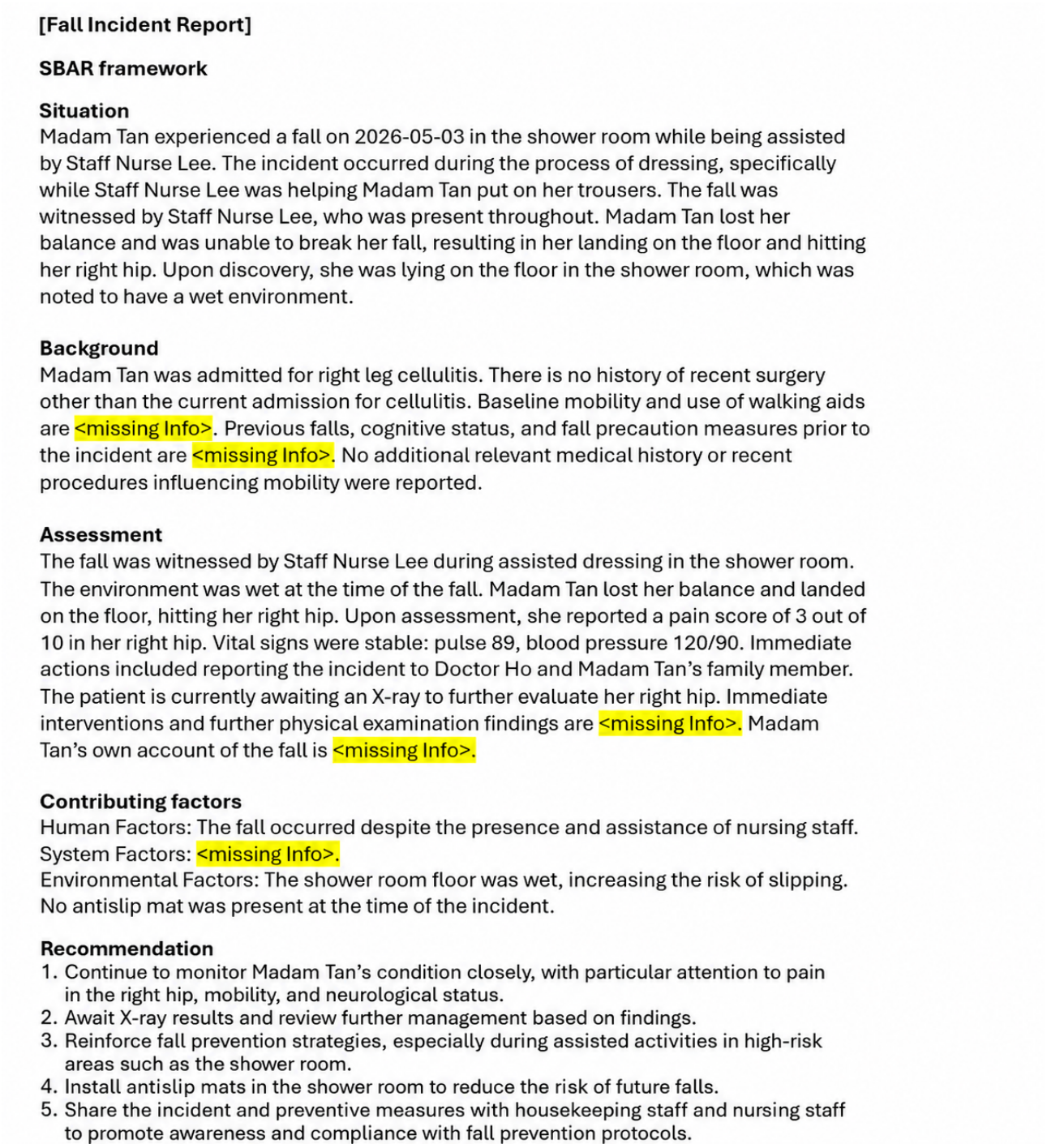
Figure 1 illustrates the supervisor workflow from verbal narration through human verification to final report submission. Figure 2 provides an illustrative example of NARRATE-generated SBAR output with missing-information placeholders.

Figure 1. Project NARRATE (Nursing AI-Refined for Accurate Transcription of Events) AI-assisted supervisor investigation documentation workflow. The workflow moves from verbal narration to World Health Organization Minimal Information Model for Patient Safety Incident Reporting and Learning Systems (MIM PS)–aligned prompts, AI-generated Situation-Background-Assessment-Recommendation (SBAR) draft output, human verification, and transfer to the institutional incident reporting system. Supervisors verify all AI-generated content and resolve any missing-information placeholders before final submission.



XSL•FO
RenderX

Figure 2. Illustrative sample of NARRATE (Nursing AI-Refined for Accurate Transcription of Events)-generated Situation-Background-Assessment-Recommendation (SBAR) output with missing-information placeholders. This example, based on a fictitious fall incident, shows how the AI-generated draft note retains explicit missing-information markers for safety-critical elements requiring supervisor verification. Supervisors are expected to review and address each placeholder before transferring the verified report into the institutional incident-reporting system. All patient and unit details are fictitious.



[Fall Incident Report]

SBAR framework

Situation

Madam Tan experienced a fall on 2026-05-03 in the shower room while being assisted by Staff Nurse Lee. The incident occurred during the process of dressing, specifically while Staff Nurse Lee was helping Madam Tan put on her trousers. The fall was witnessed by Staff Nurse Lee, who was present throughout. Madam Tan lost her balance and was unable to break her fall, resulting in her landing on the floor and hitting her right hip. Upon discovery, she was lying on the floor in the shower room, which was noted to have a wet environment.

Background

Madam Tan was admitted for right leg cellulitis. There is no history of recent surgery other than the current admission for cellulitis. Baseline mobility and use of walking aids are <missing Info>. Previous falls, cognitive status, and fall precaution measures prior to the incident are <missing Info>. No additional relevant medical history or recent procedures influencing mobility were reported.

Assessment

The fall was witnessed by Staff Nurse Lee during assisted dressing in the shower room. The environment was wet at the time of the fall. Madam Tan lost her balance and landed on the floor, hitting her right hip. Upon assessment, she reported a pain score of 3 out of 10 in her right hip. Vital signs were stable: pulse 89, blood pressure 120/90. Immediate actions included reporting the incident to Doctor Ho and Madam Tan's family member. The patient is currently awaiting an X-ray to further evaluate her right hip. Immediate interventions and further physical examination findings are <missing Info>. Madam Tan's own account of the fall is <missing Info>.

Contributing factors

Human Factors: The fall occurred despite the presence and assistance of nursing staff.
System Factors: <missing Info>.
Environmental Factors: The shower room floor was wet, increasing the risk of slipping. No antislip mat was present at the time of the incident.

Recommendation

1. Continue to monitor Madam Tan's condition closely, with particular attention to pain in the right hip, mobility, and neurological status.
2. Await X-ray results and review further management based on findings.
3. Reinforce fall prevention strategies, especially during assisted activities in high-risk areas such as the shower room.
4. Install antislip mats in the shower room to reduce the risk of future falls.
5. Share the incident and preventive measures with housekeeping staff and nursing staff to promote awareness and compliance with fall prevention protocols.

Comparator: Conventional Documentation

In the conventional process, supervisors documented investigation findings manually in free text after reviewing the incident and gathering information from relevant staff or records. Conventional documentation did not include embedded voice-to-text prompts, SBAR-structured AI output, or visible missing-information cues.

Sample, Eligibility, and Selection

We analyzed 150 deidentified completed supervisor investigation reports: 75 conventional reports from June to August 2025 and 75 NARRATE-period reports from January to March 2026. Eligible reports were nursing-originated investigation reports with sufficient narrative text to support scoring. The study focused on falls and medication administration-related incidents.

XSL•FO
RenderX

We excluded duplicates, administrative entries without substantive investigation narrative, and reports outside the study windows. NARRATE operated through the institutional ambient-scribe platform, which is technically separate from the incident-reporting system, and its use could not be mandated within the existing reporting workflow. Adoption during the postimplementation period was therefore voluntary (opt-in), and recorded NARRATE use represented approximately 40% of eligible reports. Supervisors who used NARRATE recorded the corresponding report number in a brief operational record maintained for a separate reporting-time initiative rather than for the present analysis. Entries in this record allowed confirmed NARRATE-generated reports to be identified, but the record was not designed as an exhaustive NARRATE-use registry. Absence from the record therefore could not confirm nonuse. We did not construct a postimplementation non-NARRATE comparison group because a complement-defined group could have contained unrecorded NARRATE reports and introduced exposure misclassification.

From the confirmed NARRATE-documented reports, we drew 75 at random, and we drew 75 conventional reports at random from eligible preimplementation reports. Random selection was applied within each pool to avoid investigator selection of individual reports, but it cannot remove adopter self-selection. Early adopters may differ from nonadopters in documentation practice or motivation. We therefore interpret the comparison as the NARRATE workflow as delivered under real-world voluntary uptake rather than under randomized assignment.

All 150 selected reports were analyzable, with no missing report-level or item-level data.

Sample size was not determined a priori; it was fixed by the number of eligible reports available within the institutional study windows. Accordingly, we frame this power calculation as a minimum detectable effect analysis for a fixed sample rather than as a sample-size justification. Using G*Power 3.1, a sample of 150 reports (75 per group) was sufficient to detect a standardized mean difference of d=0.46 with approximately 80% power at a 2-sided α of .05 [14]; both primary outcomes exceeded this minimum detectable effect. Domain-level and subgroup analyses were exploratory.

Outcomes and Measurement

Completeness Checklist Development and Content Validity

We assessed the primary outcome, report completeness, using a graded checklist derived from WHO technical guidance and aligned with the advanced WHO MIM PS [2,3]. The checklist grouped items into description (what happened and patient impact), explanation (suspected causes, contributing factors, and mitigating factors), and actions (immediate, corrective, and proposed follow-up actions).

Each item was rated 0="absent," 1="partial or insufficient," 2="clear and sufficient for review and learning," or 9="not applicable." Not-applicable items were excluded from the relevant denominator. We calculated WHO total completeness from applicable items and calculated domain subscores as the mean of items within each domain. The full rubric is provided in Multimedia Appendix 1.

Content validity was established through structured review by 3 experts with operational responsibility in patient safety incident review: a Nursing Division patient safety officer, an Office of Patient Safety and Quality manager, and a senior principal pharmacist with medication safety expertise.

Across 2 rounds, panelists evaluated item relevance and contextual appropriateness, and items were refined between rounds. After the second round, all retained items achieved an item-level content validity index of 1.00 [15,16]. The checklist was finalized after the NARRATE prompt template had been locked; the resulting intervention-outcome alignment is addressed in the *Limitations*.

Adapted 8-Domain PDQI-Based Narrative Quality Assessment

We assessed narrative quality using an adapted PDQI approach derived from the PDQI [12]. For this incident-investigation context, 8 domains were scored: thoroughness, usefulness, organization, comprehensibility, succinctness, synthesis, internal consistency, and fairness/balance. Each domain was rated from 1 to 5, and the adapted PDQI mean was calculated as the average of the 8 domain scores. Accuracy was excluded a priori because the ambient AI workflow did not retain an audio record or other source transcript for independent audit, so reviewers could not compare the final report with the original verbal narration or otherwise verify factual fidelity beyond the report text itself. This exclusion matters because ambient AI outputs can contain transcription errors that may propagate into the final note, occasionally with potential for harm [17]; reviewers therefore judged the quality of the final written narrative, not the truth of the underlying investigation. NARRATE drafts were always reviewed and edited by a supervisor before submission. The full rubric is provided in Multimedia Appendix 2.

Reviewer Training, Blinding, and Scoring

Two reviewers from the Nursing Safety and Quality team independently assessed all 150 deidentified reports. Neither reviewer was involved in NARRATE design, prompt authoring, user training, or implementation. Reports were deidentified and masked by an independent data handler. Direct identifiers were removed from the reviewer-facing reports, while a pseudonymous supervisor code was retained separately for clustered analysis. The 2 reviewers had no role in data preparation or access to the linkage information.

Before formal scoring, the reviewers received written rubric guidance and completed iterative calibration on 10 reports outside the study sample; this process also informed refinement of the rubrics in Multimedia Appendices 1 and 2. All calibration-informed rubric refinement, including finalization of the checklist and its item-level content validity index, was completed before any of the 150 analytic reports were rated, so postcalibration rubric drift did not affect the rated sample.

Residual unblinding remained possible because NARRATE-period reports could retain recognizable SBAR structure. This risk was addressed in the plain-text sensitivity

XSL•FO
RenderX

analysis and discussed in the *Limitations*. For the primary analysis, each report was scored independently by both reviewers, and composite scores were calculated as the mean of the 2 ratings after interrater reliability assessment. Individual discrepancies were not adjudicated before the main analysis so that ICC estimates reflected independent ratings.

Statistical Analysis

We prespecified comparisons of preimplementation versus postimplementation reports on the primary outcome (WHO total completeness) and secondary outcome (adapted PDQI [8-domain] mean), with domain-level comparisons treated as exploratory. We reported means and SDs for continuous scores and frequencies with percentages for categorical characteristics.

Interrater reliability for composite scores was assessed using 2-way random-effects, absolute-agreement intraclass correlation coefficients (ICCs), following the framework described by McGraw and Wong [18]. Single-measure ICCs represented the reliability of 1 reviewer's score, and average-measure ICCs represented the reliability of the mean across the 2 reviewers.

To account for supervisors contributing multiple reports, all report-level comparisons were refitted using linear mixed-effects models with documentation method as a fixed effect and a random intercept for the documenting supervisor. Models used restricted maximum likelihood estimation; 95% CIs and 2-sided *P* values used Wald normal-theory inference. Supervisors were nested within documentation method because none contributed reports in both study periods. Adjusted mean differences are reported as NARRATE minus conventional. Unadjusted group means, SDs, and Cohen *d* values calculated using the pooled SD were retained as descriptive standardized effect-size summaries, with approximate 95% CIs derived from the independent-groups variance estimator [19,20].

Because Levene tests had indicated unequal variance for the 2 headline outcomes, exchangeable generalized estimating equations with robust SEs were also fitted as robustness checks; conclusions were unchanged. Statistical significance was evaluated at a 2-sided α of .05. We performed 11 prespecified exploratory domain comparisons in the full-sample analysis (3 WHO subscores and 8 adapted PDQI domains). These comparisons were not adjusted for multiplicity, so the family-wise risk of type I error is increased, and domain-level findings should be interpreted as hypothesis-generating. Original reliability and descriptive analyses were conducted using IBM SPSS Statistics for Windows (version 26.0; IBM Corp); the clustered reanalysis was conducted in Python 3.12 using statsmodels 0.14.6.

Plain-Text Rerating Sensitivity Analysis

Because recognizable SBAR structure could partially unblind reviewers, we conducted a format-blinded sensitivity analysis on a stratified subsample of 60 reports (30 conventional and 30 NARRATE, balanced by incident type) converted to a uniform plain-paragraph format. SBAR labels, section headers, horizontal rules, and other structural cues were removed while preserving report content; an automated audit confirmed that no residual structural markers remained. The rerating used the same 2 reviewers as the main analysis. To prevent reviewers from linking reports to their original SBAR-formatted versions, the data handler (who did not participate in rating) prepared the deformatted set, removed all structural cues, and renumbered and reordered the reports before rerating. Both reviewers were blinded to documentation method and independently rescored all 60 reports using the same WHO completeness checklist and adapted PDQI (8-domain) rubric. Periods were compared using averaged rater scores and the same random-intercept mixed-model approach, with documenting supervisor as the clustering unit; Cohen *d* values were retained as unadjusted descriptive effect sizes.

Ethical Considerations

We analyzed deidentified existing incident-report text generated through routine institutional reporting for this retrospective quality and safety evaluation; the analytic dataset contained no patient or staff identifiers. The retrospective analysis was reviewed under the institutional Quality Assurance/Service Improvement pathway and was determined not to require review by the SingHealth Centralized Institutional Review Board (institutional reference SHS-RSH-CIRB-4234); this determination for retrospective analysis of existing deidentified records was endorsed by the Research Office, Singapore General Hospital, on May 8, 2026. Individual consent was not required because the study used deidentified existing records and involved no direct interaction with patients or staff. We reported the project in accordance with the SQUIRE 2.0 guidelines for quality-improvement reporting, supplemented by relevant STROBE elements; a completed SQUIRE 2.0 checklist is provided in Checklist 1.

Results

Report Sample

A total of 150 nursing-originated supervisor investigation reports were included: 75 conventional and 75 NARRATE reports, covering falls and medication administration-related incidents across both study periods. No cases were excluded. The operational record and associated discussion history identified the documenting supervisor for each selected report. The 75 conventional reports were contributed by 54 supervisors: 39 contributed 1 report, 10 contributed 2, 4 contributed 3, and 1 contributed 4 reports. The 75 NARRATE reports were contributed by 61 supervisors: 49 contributed 1 report, 10 contributed 2, and 2 contributed 3 reports. No supervisor contributed reports in both periods. Sample characteristics by documentation method are summarized in Table 1. The 2 groups were comparable in report length (conventional mean 777, SD 475 words; NARRATE mean 733, SD 254 words; *P*=.49), indicating that NARRATE-period gains were not attributable to longer reports.

XSL•FO
RenderX

Table . Sample characteristics by documentation method (N=150)[a].

Characteristic	Conventional (n=75)	NARRATE[b] (n=75)
Incident type, n (%)		
Fall	41 (55)	45 (60)
Medication administration-related	34 (45)	30 (40)
Sampling period	June-August 2025	January-March 2026
Report length (words), mean (SD)	777 (475)	733 (254)

[a]Report length was calculated as the word count of the event description. Harm-severity tier, ward type, and reporter seniority were not available in the deidentified dataset. Groups did not differ significantly in report length (*P*=.49).

[b]NARRATE: Nursing AI-Refined for Accurate Transcription of Events.

Interrater Reliability

Table 2 summarizes interrater reliability for the composite scores. Agreement was moderate for WHO total completeness and somewhat stronger for the adapted PDQI (8-domain) mean. For WHO total completeness, Cronbach α was 0.829; the single-measure ICC was 0.689 (95% CI 0.577 - 0.772; *P*<.001), and the average-measure ICC was 0.816 (95% CI 0.732 - 0.872). For the adapted PDQI (8-domain) mean, Cronbach α was 0.853; the single-measure ICC was 0.744 (95% CI 0.664 - 0.808; *P*<.001), and the average-measure ICC was 0.854 (95% CI 0.798 - 0.894). Across the 9 WHO items (1350 ratings), reviewers agreed exactly on 957 (70.9%) ratings, differed by 1 category on 285 (21.1%), and differed by more than 1 category on 108 (8%). Larger disagreements clustered in the explanation domain, especially contributing factors (38/150, 25.3%), mitigating factors (23/150, 15.3%), and suspected cause (15/150, 10%), whereas description items showed high agreement (5/150, 3.3% differing by >1 category). To confirm that the explanation-domain advantage was not driven by a single reviewer, we examined each reviewer separately: both scored NARRATE-period explanation content higher than conventional content (reviewer 1, mean difference 0.49, *d*=0.81, *P*<.001; reviewer 2, mean difference 0.17, *d*=0.26, *P*=.11). The direction was therefore consistent across reviewers, although the effect was larger and statistically significant for 1 reviewer and smaller and nonsignificant for the other, consistent with the greater interpretive difficulty of explanation items.

Table . Interrater reliability for composite report-quality scores (N=150).

Outcome	Cronbach α	ICC, single-measure (95% CI)[a]	ICC, average-measure (95% CI)[a]
WHO[b] total completeness	0.829	0.689 (0.577 - 0.772)	0.816 (0.732 - 0.872)
Adapted PDQI[c] (8-domain) mean	0.853	0.744 (0.664 - 0.808)	0.854 (0.798 - 0.894)

[a]ICC: intraclass correlation coefficient; ICCs are 2-way random-effects, absolute-agreement. *P*<.001 for all ICCs.

[b]WHO: World Health Organization.

[c]PDQI: Physician Documentation Quality Instrument.

Primary and Secondary Outcomes

NARRATE reports had higher WHO total completeness scores than conventional reports. The observed mean was 13.61 (SD 2.54) for NARRATE reports and 11.81 (SD 3.39) for conventional reports, corresponding to an unadjusted difference of 1.80 points. After accounting for multiple reports contributed by the same supervisor, the cluster-adjusted mean difference was 1.95 points (95% CI 0.91 - 3.00; *P*<.001). The corresponding unadjusted descriptive standardized effect size was Cohen *d*=0.60 (95% CI 0.27 - 0.93).

NARRATE reports had higher adapted PDQI (8-domain) mean scores. The observed mean was 4.13 (SD 0.34) for NARRATE reports and 3.61 (SD 0.52) for conventional reports, corresponding to an unadjusted difference of 0.52 points. After accounting for multiple reports contributed by the same supervisor, the cluster-adjusted mean difference was 0.53 (95% CI 0.37 - 0.69; *P*<.001) points. The corresponding unadjusted descriptive standardized effect size was Cohen *d*=1.18 (95% CI 0.83 - 1.53). Exchangeable generalized estimating equation robustness analyses yielded similar estimates for WHO total completeness (1.99, 95% CI 0.94 - 3.04; *P*<.001) and adapted PDQI mean (0.53, 95% CI 0.37 - 0.70; *P*<.001). Table 3 summarizes group-level results.

XSL•FO
RenderX

Table . Overall report completeness and narrative quality by documentation method (N=150).

Outcome	Conventional, mean (SD)[a]	NARRATE[b], mean (SD)[a]	Adjusted mean difference (95% CI)[a]	*P* value[a]	Cohen *d*[c] (95% CI)
WHO[d] total completeness	11.81 (3.39)	13.61 (2.54)	1.95 (0.91-3.00)	<.001	0.60 (0.27-0.93)
Adapted PDQI[e] (8-domain) mean	3.61 (0.52)	4.13 (0.34)	0.53 (0.37-0.69)	<.001	1.18 (0.83-1.53)

[a]Mean differences, 95% CIs, and *P* values are from restricted maximum likelihood random-intercept linear mixed-effects models with documentation method as a fixed effect and documenting supervisor as a random intercept.

[b]NARRATE: Nursing AI-Refined for Accurate Transcription of Events.

[c]Cohen *d* was calculated from unadjusted group means using the pooled SD; 95% CIs for *d* were derived from the independent-groups variance estimator.

[d]WHO: World Health Organization.

[e]PDQI: Physician Documentation Quality Instrument.

Exploratory WHO Completeness Domain Analysis

Exploratory WHO domain analysis suggested that the overall completeness improvement was driven mainly by explanation and actions (Table 4). In cluster-adjusted models, the description subscore increased only slightly from 1.86 (SD 0.32) to 1.94 (SD 0.18; adjusted mean difference 0.08, 95% CI −0.01 to 0.16; *P*=.07; *d*=0.30), suggesting a ceiling effect. Explanation increased from 1.00 (SD 0.58) to 1.33 (SD 0.51; adjusted mean difference 0.34, 95% CI 0.15 - 0.52; *P*<.001; *d*=0.60), and actions increased from 1.27 (SD 0.43) to 1.43 (SD 0.31; adjusted mean difference 0.19, 95% CI 0.06 - 0.32; *P*=.004; *d*=0.43). Cohen *d* values are unadjusted descriptive effects. These subscore findings should be interpreted as exploratory.

Table . World Health Organization minimal information model for patient safety incident reporting and learning systems–aligned completeness subscores by documentation method (N=150)[a].

WHO[b] domain[c]	Conventional, mean (SD)[d]	NARRATE[e], mean (SD)[d]	Adjusted mean difference (95% CI)[d]	*P* value[d]	Cohen *d*[f] (95% CI)
Description	1.86 (0.32)	1.94 (0.18)	0.08 (−0.01 to 0.16)	.07	0.30 (−0.03 to 0.62)
Explanation	1.00 (0.58)	1.33 (0.51)	0.34 (0.15 to 0.52)	<.001	0.60 (0.28 to 0.93)
Actions	1.27 (0.43)	1.43 (0.31)	0.19 (0.06 to 0.32)	.004	0.43 (0.11 to 0.76)

[a]Subscore values are means of constituent items within each domain.

[b]WHO: World Health Organization.

[c]Domain-level analyses were exploratory.

[d]Mean differences, 95% CIs, and *P* values are from restricted maximum likelihood random-intercept linear mixed-effects models clustered by documenting supervisor.

[e]NARRATE: Nursing AI-Refined for Accurate Transcription of Events.

[f]Cohen *d* values are unadjusted descriptive effects. Domain-level analyses were exploratory.

Exploratory Adapted PDQI (8-Domain) Domain Analysis

Exploratory adapted PDQI (8-domain) analysis showed higher unadjusted scores in the NARRATE group across all 8 measured domains (Table 5). In cluster-adjusted models, differences were statistically significant for thoroughness, usefulness, organization, comprehensibility, synthesis, internal consistency, and fairness/balance. Succinctness improved numerically from 4.31 (SD 0.80) to 4.47 (SD 0.57) but was not statistically significant (adjusted mean difference 0.12, 95% CI −0.12 to 0.37; *P*=.31; *d*=0.23). The largest adjusted absolute increase was in organization, and the largest unadjusted standardized effects were for organization (*d*=2.10), comprehensibility (*d*=1.33), and fairness/balance (*d*=1.32). Fairness/balance nevertheless remained the lowest-rated adapted PDQI domain in both groups.

XSL•FO
RenderX

Table . Adapted Physician Documentation Quality Instrument (8-domain) narrative quality domain scores by documentation method (N=150).

Adapted PDQI[a] (8-domain) domain[b]	Conventional, mean (SD)[c]	NARRATE[d], mean (SD)[c]	Adjusted mean difference (95% CI)[c]	*P* value	Cohen *d*[e] (95% CI)
Thorough	3.61 (1.08)	4.21 (0.85)	0.66 (0.32 to 1.00)	<.001	0.62 (0.30 to 0.95)
Useful	3.70 (1.09)	4.27 (0.88)	0.61 (0.27 to 0.95)	<.001	0.57 (0.25 to 0.90)
Organized	3.59 (0.52)	4.65 (0.48)	1.04 (0.86 to 1.23)	<.001	2.10 (1.70 to 2.50)
Comprehensible	4.01 (0.49)	4.57 (0.35)	0.56 (0.41 to 0.72)	<.001	1.33 (0.97 to 1.68)
Succinct	4.31 (0.80)	4.47 (0.57)	0.12 (−0.12 to 0.37)	.31	0.23 (−0.09 to 0.55)
Synthesized	3.43 (1.14)	3.96 (0.87)	0.54 (0.20 to 0.88)	.002	0.53 (0.20 to 0.85)
Internal consistency	3.47 (0.42)	3.63 (0.34)	0.17 (0.04 to 0.31)	.01	0.43 (0.11 to 0.76)
Fair/balanced	2.79 (0.39)	3.31 (0.41)	0.54 (0.40 to 0.67)	<.001	1.32 (0.97 to 1.68)

[a]PDQI: Physician Documentation Quality Instrument.

[b]Domain-level analyses were exploratory.

[c]Mean differences, 95% CIs, and *P* values are from restricted maximum likelihood random-intercept linear mixed-effects models clustered by documenting supervisor.

[d]NARRATE: Nursing AI-Refined for Accurate Transcription of Events.

[e]Cohen *d* was calculated from unadjusted group means using the pooled SD; 95% CIs were derived from the independent-groups variance estimator.

Plain-Text Rerating Sensitivity Analysis

After reports were re-rendered into a uniform plain-paragraph format and a stratified subsample was rerated, the completeness advantage remained in the cluster-adjusted analysis (adjusted mean difference 1.70, 95% CI 0.27 - 3.14; *P*=.02; unadjusted *d*=0.59, 95% CI 0.08 - 1.11), as did the explanation subscore (adjusted mean difference 0.25, 95% CI 0.05 - 0.45; *P*=.02; *d*=0.63). The actions subscore still favored NARRATE but was borderline (adjusted mean difference 0.16, 95% CI −0.0007 to 0.32; *P*=.05; *d*=0.50). The adapted PDQI (8-domain) mean remained higher in NARRATE reports but attenuated to a moderate effect (adjusted mean difference 0.25, 95% CI 0.06 - 0.43; *P*=.009; *d*=0.64, 95% CI 0.12 - 1.16). Among PDQI domains, organization (adjusted mean difference 0.43, 95% CI 0.19 - 0.67; *P*<.001; *d*=0.92) and comprehensibility (adjusted mean difference 0.45, 95% CI 0.21 - 0.69; *P*<.001; *d*=0.93) remained higher after deformatting, whereas synthesis, internal consistency, and fairness/balance were not statistically significant. These results suggest that the completeness finding and part of the organization and comprehensibility advantage were robust to removal of structural formatting, whereas several other narrative-quality differences were more format-dependent. Cluster-adjusted mean differences, clustered *P* values, and unadjusted effect sizes are summarized in Table 6; corresponding full-sample results are reported in Tables 3-5.

XSL•FO
RenderX

Table . Format-blinded plain-text sensitivity-analysis results (n=60)[a].

Outcome	Adjusted mean difference (95% CI)	Cohen *d*[b] (95% CI)	*P* value[c]
WHO[d] total completeness	1.70 (0.27 to 3.14)	0.59 (0.08 to 1.11)	.02
Description subscore	0.13 (−0.004 to 0.26)	0.48 (−0.03 to 0.99)	.06
Explanation subscore	0.25 (0.05 to 0.45)	0.63 (0.11 to 1.15)	.02
Actions subscore	0.16 (−0.001 to 0.32)	0.50 (−0.01 to 1.01)	.05
Adapted PDQI[e] (8-domain) mean	0.25 (0.06 to 0.43)	0.64 (0.12 to 1.16)	.009
Thorough	0.26 (−0.13 to 0.65)	0.30 (−0.21 to 0.81)	.19
Useful	0.26 (−0.13 to 0.65)	0.30 (−0.21 to 0.81)	.19
Organized	0.43 (0.19 to 0.67)	0.92 (0.39 to 1.45)	<.001
Comprehensible	0.45 (0.21 to 0.69)	0.93 (0.40 to 1.46)	<.001
Succinct	0.13 (−0.16 to 0.41)	0.18 (−0.32 to 0.69)	.38
Synthesized	0.21 (−0.03 to 0.45)	0.39 (−0.12 to 0.90)	.09
Internal consistency	0.07 (−0.14 to 0.29)	0.20 (−0.30 to 0.71)	.51
Fair/balanced	0.13 (−0.09 to 0.36)	0.30 (−0.21 to 0.81)	.25

[a]Positive values favor NARRATE. Corresponding full-sample results are reported in Tables 3-5.

[b]Cohen *d* was calculated from unadjusted group means using the pooled SD; 95% CIs were derived from the independent-groups variance estimator.

[c]*P* values are from restricted maximum likelihood random-intercept linear mixed-effects models clustered by documenting supervisor.

[d]WHO: World Health Organization.

[e]PDQI: Physician Documentation Quality Instrument.

Discussion

Principal Findings

In this retrospective pre-post comparative document-quality study of 150 completed supervisor investigation reports, NARRATE-period reports were more complete and had higher adapted PDQI (8-domain) mean scores than conventional reports. The plain-paragraph sensitivity analysis refined that picture: the completeness advantage and explanation-domain gain persisted after structural formatting was removed, the actions-domain gain attenuated, and organization and comprehensibility remained higher but with smaller effects. By contrast, synthesis, internal consistency, and fairness/balance no longer differed significantly. The most defensible interpretation is therefore improved completeness, explanation content, and some aspects of narrative organization and clarity, with several other narrative-quality differences partly mediated by visible format.

Description was already well documented in conventional reports, which likely limited room for improvement. By contrast, explanation and action content require supervisors to organize findings, identify contributing factors, and specify follow-up, making them more plausible targets for structured prompting [21,22].

Interpretation of WHO Completeness Findings

The completeness findings support prompt-guided documentation aligned to the advanced WHO MIM PS. NARRATE did not appear merely to lengthen or polish reports; it improved the explanatory and action-oriented elements most relevant to review and learning [2,3]. This matters because reports that omit causes or contributing factors are weaker starting points for investigation and root-cause analysis [2,4,23]. By prompting these elements and making omissions visible through missing-information cues, NARRATE may help turn routine supervisor reports into more usable review documents.

The actions-domain gain should still be interpreted cautiously. The effect was modest in the full sample and only borderline in the plain-text analysis, suggesting improved documentation of actions more than proof of stronger action planning itself. Future iterations could prompt more explicitly for system-level actions, accountability, and timelines.

Interpretation of Adapted PDQI (8-Domain) Findings

The adapted PDQI findings require domain-specific interpretation rather than a single claim of globally improved narrative quality [12].

In the full sample, organization, comprehensibility, and fairness/balance showed the largest effects, consistent with the visible SBAR format, which may reduce reviewer cognitive load [13,24,25]. The plain-paragraph sensitivity analysis showed that this pattern was only partly format-driven: organization and comprehensibility remained significantly higher after SBAR labels and structural cues were removed, although both attenuated substantially. By contrast, synthesis, internal consistency, and fairness/balance were no longer significant, and the adapted PDQI (8-domain) mean decreased from a large to a moderate effect.

The narrative-quality findings therefore appear to reflect a mixed mechanism. Some gains, particularly in organization and

XSL•FO
RenderX

comprehensibility, seem to extend beyond recognition of the formatted template, whereas a substantial portion of the synthesis, internal consistency, and fairness/balance advantage is likely format-mediated. This does not make structure unimportant in practice; structured presentation may still improve readability and navigation even if it should not be interpreted here as an independent content gain.

Succinctness was the only PDQI domain without a significant full-sample gain, which is plausible because SBAR summaries and completion prompts can lengthen reports. Even so, succinctness remained high in the NARRATE period, suggesting that improved completeness did not come at the cost of markedly poorer brevity.

Fairness/balance had the lowest mean PDQI (8-domain) domain score in both groups and became non-significant after deformatting. This suggests that fairness in incident reporting depends less on formatting than on investigation depth, avoidance of blame, and adequate attention to system factors [26,27]. The lower variance of adapted PDQI (8-domain) mean scores in the NARRATE group (SD 0.34 vs 0.52) is also a mixed finding: more standardized documentation may improve consistency, but it may also compress idiosyncratic detail that remains useful for review.

Interrater Reliability and Measurement Implications

Interrater reliability was somewhat stronger for adapted PDQI (8-domain) mean than for WHO total completeness, but both composites were within an acceptable range for group-level comparisons. Larger WHO disagreements clustered in the explanation items, which require more interpretive judgment than the more factual description items. These findings support continued use of detailed anchors and suggest that future studies could consider adjudication or consensus scoring for borderline causal items.

Comparison With Prior Work

These findings extend prior concerns that incomplete incident reports limit patient safety learning [4,5]. Rather than evaluating a generic transcription tool, this study examined a workflow that combined voice-to-text capture, WHO MIM PS-aligned prompts, visible omission cues, and SBAR output. The results suggest that prompt-guided narration may improve the capture of learning-relevant elements, although they remain documentation-quality outcomes rather than evidence of reduced harm.

Implications for Nursing Practice, Leadership, and Safety Governance

For nursing leaders, the practical implication is that AI-enabled documentation may be more valuable as a way to improve the quality of incident-review inputs than as a pure efficiency tool [3,4,10,23,28]. When designed to improve the structure, completeness, and usability of investigation narratives, workflows like NARRATE may help supervisors produce reports that are easier to review and more useful for learning.

Implementation should nevertheless preserve professional accountability. Supervisors must verify AI-generated content, correct transcription errors, resolve missing-information placeholders, and ensure that recommendations are clinically and operationally appropriate. Governance should also address confidentiality, responsible AI use, prompt versioning, audit trails, user training, and equity of use across roles and settings [29]. Because accuracy was not assessed, these findings indicate more complete and better-organized reports, not more factually correct ones. No draft-to-final accuracy audit was performed, and accuracy assurance in this workflow rests on mandatory supervisor verification of every draft. Ambient-scribe transcription errors can propagate to the final note, sometimes with potential for moderate-to-severe harm [17]. Completeness and organization gains therefore do not substitute for accuracy assurance and should not be taken as evidence that AI-generated incident narratives can be relied upon without source-data verification in higher-stakes investigations.

Strengths

This study has several strengths: a guidance-informed completeness framework aligned to WHO reporting principles, expert content-validity review of the completeness checklist, a 150-report analytic sample, blinded dual-reviewer scoring with quantified interrater reliability, and parallel examination of both total scores and domain-specific patterns.

Limitations

This study has several limitations. The retrospective pre-post design cannot establish causality, and secular changes cannot be entirely excluded. No additional reporting-quality training or audit-feedback initiative was conducted in the interval, and the institutional incident-reporting system was unchanged; the postimplementation window represented a stable postlaunch period with a fixed prompt template.

NARRATE use was voluntary, and recorded use represented approximately 40% of eligible postimplementation reports. Because the operational record was maintained for another purpose and was not exhaustive, a reliably classified postimplementation non-NARRATE group could not be constructed: absence from the record could not confirm nonuse, and a complement-defined group could have included unrecorded NARRATE reports. The NARRATE group therefore reflects self-selected early adopters, who may differ from nonadopters in motivation, experience, or documentation practice; observed differences may consequently overestimate effects under mandated or universal adoption.

Although the mixed models accounted for correlation among multiple reports from the same supervisor, no supervisor contributed reports in both periods. Documentation method was therefore completely nested within supervisor pool, and clustering adjustment cannot eliminate confounding by different supervisor composition. The comparison cannot fully separate workflow effects from adopter or supervisor characteristics and should be interpreted as performance under real-world voluntary uptake rather than as an unbiased causal contrast. Residual unblinding was possible because NARRATE reports could retain recognizable SBAR structure; the plain-text analysis suggests that some narrative-quality effects were format-sensitive. Because we assessed final supervisor-edited text, we cannot separate the independent effects of transcription,

XSL•FO
RenderX

prompting, missing-information cues, and human editing; effects should be attributed to the workflow as a whole.

The prompt template and completeness checklist were both aligned with the advanced WHO MIM PS, creating intervention-outcome alignment. Because the checklist was finalized after the NARRATE prompt template was locked, this criterion-contamination risk is directional: the checklist could in principle have been shaped by what the prompt already produced. To limit this risk, the content-validity panel evaluated items against WHO MIM PS source guidance rather than the specific NARRATE prompt wording, the checklist graded sufficiency rather than simple field completion, and the adapted PDQI findings partly persisted after deformatting. Nonetheless, confirmatory work using a completeness instrument developed independently of the NARRATE prompts is needed.

The adapted PDQI underwent content adaptation but not full construct validation, and the study assessed documentation quality rather than action quality, timeliness, recurrence, factual accuracy, or patient-safety outcomes. No draft-to-final accuracy audit was performed; accuracy assurance rested on mandatory supervisor verification of every draft. Granular case-mix variables were unavailable, and the study was conducted at a single Singapore tertiary center, limiting causal inference and generalizability.

Future Work

Future work should test whether improved completeness translates into stronger action planning, faster review, better aggregation of contributing factors, and measurable organizational learning. Confirmatory studies should include concurrent, reliably classified nonadopter comparison groups or phased implementation designs that permit within-supervisor comparisons, helping distinguish workflow effects from adopter and supervisor characteristics. Larger multicenter studies could assess implementation fidelity, adoption, and equity across roles and settings. Prospective studies should retain auditable source material so that documentation quality and draft-to-final accuracy can be evaluated together.

Conclusions

Project NARRATE is a nursing-led application of ambient AI documentation support to incident-investigation reporting. Among voluntary early adopters in this pre-post study, NARRATE-supported reports were associated with greater completeness and higher adapted PDQI (8-domain) mean scores after accounting for repeated reports by supervisors, and completeness gains persisted after structural formatting was removed. The findings do not establish that NARRATE alone caused these differences because adopter self-selection and nonoverlapping supervisor pools remain potential explanations. The strongest evidence supports gains in completeness, explanation content, and selected aspects of narrative organization and clarity, whereas several other narrative-quality differences were format-sensitive. Broader prospective evaluation with concurrent controls is required, with AI-generated documentation remaining subject to human verification and governance.

Acknowledgments

The authors acknowledge the nursing supervisors and frontline nurses who contributed to the implementation of Project NARRATE (Nursing AI-Refined for Accurate Transcription of Events) and provided feedback that informed iterative workflow refinement. The authors also thank the expert panel members from Nursing Safety and Quality, the Office of Patient Safety and Quality, and the Medication Safety Committee for their review of the incident report completeness checklist. The authors further acknowledge the support of the SingHealth Digital Empowerment Office, SGH Artificial Intelligence and Automation, and the SingHealth Office of Insights and Analytics in enabling the broader implementation and evaluation of the NARRATE workflow.

The authors used generative AI tools for language editing, organization, formatting support, and manuscript-review suggestions during manuscript preparation. ChatGPT (GPT-5.6 Thinking; OpenAI) supported drafting, editing, and formatting; Claude (Sonnet 5; Anthropic) supported independent review suggestions. The authors reviewed, revised, and verified all scientific content and take full responsibility for the final manuscript.

Funding

This work received no external funding.

Data Availability

The data analyzed in this study were derived from deidentified patient safety incident-investigation reports and are not publicly available because of confidentiality and patient safety governance restrictions. Aggregated, nonidentifiable data may be made available from the corresponding author on reasonable request and subject to institutional approval.

Authors' Contributions

Conceptualization: KYT
Data interpretation: LH
Intervention development: KCYW
Intervention implementation: LH
Investigation: KYT

XSL•FO
RenderX

Methodology: KYT, KCYW, LH
Project administration: KYT, KCYW
Project governance: SYA, JGNN
Supervision: KYT, SYA, JGNN
Writing – original draft: KYT
Writing – review & editing: KYT, KCYW, LH, SYA, JGNN
All authors reviewed and approved the final manuscript and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Advanced WHO Minimal Information Model for Patient Safety Incident Reporting and Learning Systems–aligned graded completeness checklist, adapted from WHO technical guidance and the MIM PS User Guide [2,3].
[DOCX File, 43 KB - nursing_v9i1e100775_app1.docx]

Multimedia Appendix 2
Eight-domain narrative-quality rubric adapted from the Physician Documentation Quality Instrument-9 described by Stetson et al [12].
[DOCX File, 44 KB - nursing_v9i1e100775_app2.docx]

Checklist 1
SQUIRE 2.0 checklist.
[DOCX File, 30 KB - nursing_v9i1e100775_app3.docx]

References

1. Griffeth EM, Gajic O, Schueler N, Todd A, Ramar K. Multifaceted intervention to improve patient safety incident reporting in intensive care units. J Patient Saf 2023 Oct 1;19(7):422-428. [doi: 10.1097/PTS.0000000000001151] [Medline: 37466643]
2. Patient safety incident reporting and learning systems: technical report and guidance. : World Health Organization; 2020 URL: https://iris.who.int/server/api/core/bitstreams/49947589-0a06-4708-8cc9-379bc8d2f9f2/content [accessed 2026-07-25]
3. Minimal information model for patient safety incident reporting and learning systems: user guide. : World Health Organization; 2016 URL: https://iris.who.int/server/api/core/bitstreams/a13b925e-fe8e-49d3-a31e-7e7a7d88e646/content [accessed 2026-07-25]
4. Yao B, Kang H, Wang J, Zhou S, Gong Y. Toward reporting support and quality assessment for learning from reporting: a necessary data elements model for narrative medication error reports. AMIA Annu Symp Proc 2018;2018:1581-1590. [Medline: 30815204]
5. Scott J, Dawson P, Heavey E, et al. Content analysis of patient safety incident reports for older adult patient transfers, handovers, and discharges: do they serve organizations, staff, or patients? J Patient Saf 2021 Dec 1;17(8):e1744-e1758. [doi: 10.1097/PTS.0000000000000654] [Medline: 31790011]
6. Stavropoulou C, Doherty C, Tosey P. How effective are incident‐reporting systems for improving patient safety? A systematic literature review. Milbank Q 2015 Dec;93(4):826-866. [doi: 10.1111/1468-0009.12166] [Medline: 26626987]
7. Mitchell I, Schuster A, Smith K, Pronovost P, Wu A. Patient safety incident reporting: a qualitative study of thoughts and perceptions of experts 15 years after “To Err is Human”. BMJ Qual Saf 2016 Feb;25(2):92-99. [doi: 10.1136/bmjqs-2015-004405] [Medline: 26217037]
8. Macrae C. The problem with incident reporting. BMJ Qual Saf 2016 Feb;25(2):71-75. [doi: 10.1136/bmjqs-2015-004732] [Medline: 26347519]
9. Ong JHW, Tung JYM, Sng GGR, et al. A pilot study using ambient artificial intelligence scribes in clinical documentation in a urology outpatient clinic. BJU Int 2025 May 20;136(3):415-417. [doi: 10.1111/bju.16784] [Medline: 40390678]
10. Tan JYE, Rafi IBM, Sng GGR, et al. Impact of an ambient AI scribe among clinicians and patients: real-world prospective observational time-motion study. JMIR Med Inform 2026 Mar 31;14:e85580. [doi: 10.2196/85580] [Medline: 41915701]
11. Lea W, Budworth L, O’Hara J, Vincent C, Lawton R. Investigators are human too: outcome bias and perceptions of individual culpability in patient safety incident investigations. BMJ Qual Saf 2026 Feb 19;35(3):159-168. [doi: 10.1136/bmjqs-2024-017926] [Medline: 39929715]
12. Stetson PD, Bakken S, Wrenn JO, Siegler EL. Assessing electronic note quality using the Physician Documentation Quality Instrument (PDQI-9). Appl Clin Inform 2012;3(2):164-174. [doi: 10.4338/aci-2011-11-ra-0070] [Medline: 22577483]
13. Haig KM, Sutton S, Whittington J. SBAR: a shared mental model for improving communication between clinicians. Jt Comm J Qual Patient Saf 2006 Mar;32(3):167-175. [doi: 10.1016/s1553-7250(06)32022-3] [Medline: 16617948]

XSL•FO
RenderX

14. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
15. Lynn MR. Determination and quantification of content validity. Nurs Res 1986;35(6):382-385. [Medline: 3640358]
16. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health 2007 Aug;30(4):459-467. [doi: 10.1002/nur.20199] [Medline: 17654487]
17. Anderson TN, Mohan V, Dorr DA, Ratwani RM, Biro JM, Gold JA. Evaluating the quality and safety of ambient digital scribe platforms using simulated ambulatory encounters. Mayo Clin Proc Digit Health 2025 Dec;3(4):100292. [doi: 10.1016/j.mcpdig.2025.100292] [Medline: 41234546]
18. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1(1):30-46. [doi: 10.1037/1082-989X.1.1.30]
19. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Lawrence Erlbaum Associates; 1988. [doi: 10.4324/9780203771587]
20. Hedges LV, Olkin I. Statistical Methods for Meta-Analysis: Academic Press; 1985.
21. Peerally MF, Carr S, Waring J, Dixon-Woods M. The problem with root cause analysis. BMJ Qual Saf 2017 May;26(5):417-422. [doi: 10.1136/bmjqs-2016-005511] [Medline: 27340202]
22. Hibbert PD, Thomas MJW, Deakin A, et al. Are root cause analyses recommendations effective and sustainable? An observational study. Int J Qual Health Care 2018 Mar 1;30(2):124-131. [doi: 10.1093/intqhc/mzx181] [Medline: 29346587]
23. Yao B, Kang H, Gong Y. Data quality assessment of narrative medication error reports. Stud Health Technol Inform 2019 Aug 9;265:101-106. [doi: 10.3233/SHTI190146] [Medline: 31431584]
24. Sweller J. Cognitive load during problem solving: effects on learning. Cogn Sci 1988 Apr;12(2):257-285. [doi: 10.1207/s15516709cog1202_4]
25. Paas F, Renkl A, Sweller J. Cognitive load theory and instructional design: recent developments. Educ Psychol 2003 Jan 1;38(1):1-4. [doi: 10.1207/S15326985EP3801_1]
26. Boysen PG. Just culture: a foundation for balanced accountability and patient safety. Ochsner J 2013;13(3):400-406. [Medline: 24052772]
27. Cooper J, Edwards A, Williams H, et al. Nature of blame in patient safety incident reports: mixed methods analysis of a national database. Ann Fam Med 2017 Sep;15(5):455-461. [doi: 10.1370/afm.2123] [Medline: 28893816]
28. Perkins SW, Muste JC, Alam T, Singh RP. Improving clinical documentation with artificial intelligence: a systematic review. Perspect Health Inf Manag 2024;21(2):1d. [Medline: 40134899]
29. Guidance on the use of AI-enabled ambient scribing products in health and care settings. NHS England. 2025. URL: https://www.england.nhs.uk/long-read/guidance-on-the-use-of-ai-enabled-ambient-scribing-products-in-health-and-care-settings/ [accessed 2026-07-18]

Abbreviations

ICC: intraclass correlation coefficient
MIM PS: Minimal Information Model for Patient Safety Incident Reporting and Learning Systems
NARRATE: Nursing AI-Refined for Accurate Transcription of Events
PDQI: Physician Documentation Quality Instrument
SBAR: Situation-Background-Assessment-Recommendation
SQUIRE: Standards for Quality Improvement Reporting Excellence
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
WHO: World Health Organization

Edited by E Borycki, J Sarvestan; submitted 08.May.2026; peer-reviewed by DL Rosa, F Yousefi, J Wolfe; revised version received 18.Jul.2026; accepted 20.Jul.2026; published 10.Aug.2026.

Please cite as:
Teo KY, Huang L, Woh KCY, Ang SY, Ng JGN
Nurse-Led Ambient AI Scribe for Patient Safety Incident Investigation Reports (Project NARRATE): Retrospective Pre-Post Comparative Document-Quality Study
JMIR Nursing 2026;9:e100775
URL: https://nursing.jmir.org/2026/1/e100775
doi:10.2196/100775

© Kai Yunn Teo, Liwen Huang, Kelly Chai Yuen Woh, Shin Yuh Ang, Jade Gaik Nai Ng. Originally published in JMIR Nursing (https://nursing.jmir.org), 10.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons

XSL•FO
RenderX

Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Care Partner Perspectives on Hypertension Management in Alzheimer Disease and Related Dementias in an Online Forum: Qualitative Thematic Analysis

Jisook Ko[1], PhD, RN; Kylie Meyer[2], PhD; Jing Wang[3], MPH, PhD

[1]College of Nursing, University of Utah, 10 South 2000 East, Salt Lake City, UT, United States

[2]School of Nursing, Case Western Reserve University, Cleveland, OH, United States

[3]Florida State University, Tallahassee, FL, United States

Corresponding Author:
Jisook Ko, PhD, RN
College of Nursing, University of Utah, 10 South 2000 East, Salt Lake City, UT, United States

Abstract

Background: Family care partners manage care for individuals with Alzheimer disease and related dementias (ADRD), including hypertension management crucial for preventing cognitive decline. Nurses frequently interact with these caregiver-patient dyads but have limited evidence about caregivers' real-world experiences of managing this comorbidity. Understanding these experiences is essential for developing nurse-led interventions that support both patients and caregivers.

Objective: This study aimed to explore ADRD care partners' perceptions, experiences, and challenges related to hypertension management. A qualitative thematic analysis of care partner posts from an online forum focused on ADRD was conducted.

Methods: We analyzed 300 posts from ALZConnected, collected using the keywords "hypertension" and "blood pressure." Two independent researchers used reflexive thematic analysis, following the Braun and Clarke framework until data saturation was achieved.

Results: Among posts from predominantly family care partners (10,427/11,092, 94%, mostly adult children), 7 themes emerged: Pharmacotherapy (155/300, 51.7%), addressing medication efficacy, polypharmacy, and side effects; Patient Health Care Implementation Strategies (n=125, 41.7%), encompassing care coordination and provider interactions; Comorbid Care Management (n=136, 45.3%), highlighting caregiving demands, behavioral symptoms of ADRD, and care coordination challenges as barriers to hypertension control; Hypertension Management and Outcomes (n=71, 23.7%), including hypertensive crises; Managing Behavioral Symptoms (n=38, 12.7%), particularly care resistance; Peer Medical Advice Exchange (n=9, 3%); and Caregiver Burden and Self-Care for Hypertension (n=25, 8.3%), revealing care partners' own hypertension attributed to caregiving stress.

Conclusions: Managing comorbid hypertension in ADRD presents challenges spanning medication management, behavioral symptoms, and health care coordination. Caregivers reported developing hypertension themselves, highlighting bidirectional health impacts. Health care providers should implement dyadic interventions addressing both patient hypertension management and caregiver health within the care partnership.

(JMIR Nursing 2026;9:e85818) doi:10.2196/85818

KEYWORDS

Alzheimer disease; dementia; care partner; hypertension; online health communities

Introduction

Background

Alzheimer disease and related dementias (ADRD) affect approximately 7 million Americans aged 65 years and older, projected to reach 13.8 million by 2050 [1,2]. While antiamyloid therapies represent a significant advance as disease-modifying treatments, their clinical impact and accessibility remain subjects of ongoing debate; thus, prevention continues to focus on managing chronic conditions, particularly hypertension, which contributes to cerebrovascular disease, mixed pathology, and potential acceleration of cognitive decline [3-6]. Hypertension serves both as a vascular risk factor for ADRD development and as a highly prevalent comorbidity, with 83% of patients with ADRD experiencing hypertension [7]. Midlife hypertension shows consistent associations with later cognitive impairment and dementia [4,8].

Family care partners provide the majority of care for persons living with ADRD, yet they typically lack formal training in managing complex health needs such as hypertension [9,10]. This gap is particularly concerning because hypertension management encompasses multiple demanding tasks: medication

XSL•FO
RenderX

administration, dietary oversight, exercise encouragement, and blood pressure (BP) monitoring, all requiring knowledge and skills that care partners may not possess.

Nearly half of caregivers providing complex care, regardless of care recipient condition, worry about making mistakes, suggesting low self-efficacy [10]. Research demonstrates that low self-efficacy when managing complex tasks, such as hypertension regimens, is associated with adverse caregiver health outcomes, including depression [11]. Behavioral and psychological symptoms of dementia (BPSD), such as refusal of care, can interfere with care delivery and exacerbate caregiver burden [12-14]. These patterns are consistent with predictions from the stress and coping model [15,16], which posits that caregiver outcomes depend on the dynamic interplay between primary stressors (care demands), secondary stressors (role strain, competing responsibilities), appraisal processes (perceived burden and self-efficacy), and coping resources (social support, problem-solving strategies). Pearlin et al [16] adapted this framework specifically for dementia caregiving, emphasizing how chronic care demands can cascade into secondary stress proliferation, affecting caregiver health and well-being. Within this framework, managing hypertension in ADRD represents a significant primary stressor, one that, when compounded by low caregiver self-efficacy and BPSD, may contribute to the broader pattern of caregiver health decline documented in the literature.

Within this framework, managing hypertension in ADRD represents a significant primary stressor. Nearly half of caregivers providing complex care, regardless of care recipient condition, worry about making mistakes, suggesting low self-efficacy [10]. According to the stress and coping model, low self-efficacy when providing complex care, such as hypertension management, among ADRD caregivers contributes to adverse caregiver health outcomes, such as depression [15]. BPSD, such as refusal of care, can interfere with care delivery and exacerbate caregiver burden [11,12,17].

Care partners performing health management tasks report reduced social participation [13], while those with greater social support experience lower burden [14,18]. Online support communities provide caregivers with accessible platforms for social support, including information seeking, emotional support, and shared decision-making [19-21]. A meta-analysis showed positive impacts of internet-based group support on caregiver social support and self-efficacy [22]. Specifically, this meta-analysis synthesized findings from internet-based group support programs—structured, moderated online interventions designed to deliver peer connection, psychoeducation, and coping skills to dementia care partners via videoconferencing or asynchronous message boards—and found improvements in perceived social support, self-efficacy, and reductions in burden across studies. A scoping review of online support forums further identified ALZConnected [23]—the forum examined in this study—as one of the largest and most active peer-to-peer support communities for dementia care partners, with users exchanging practical advice, emotional support, and shared coping strategies asynchronously [24]. Despite these benefits, peer advice exchanged in unmoderated or lightly moderated forums may include medically inaccurate information, highlighting the need for nurses to understand what care partners encounter and discuss in these spaces.

The psychosocial burdens described above do not occur in isolation from physical health. A growing body of evidence demonstrates that the chronic stress of ADRD caregiving has direct cardiovascular consequences. Older adults who provide care to a spouse or family member with dementia have significantly elevated rates of hypertension compared with noncaregiving peers [25,26]. Longitudinal studies indicate that caregiving stress accelerates the development of hypertension and increases the risk of incident cardiovascular disease, even after adjusting for age and baseline health status [16,26]. Two intersecting pathways explain this elevated risk. First, caregivers may develop hypertension de novo as a physiological response to chronic stress, through sustained activation of the hypothalamic-pituitary-adrenal axis and sympathetic nervous system. Second, caregivers who had pre-existing hypertension often experience worsening BP control as caregiving demands displace time and energy for their own health behaviors, including medication adherence, dietary management, and medical appointments [25]. Despite this dual burden—managing their care recipient's hypertension while their own BP goes unmanaged—caregiver cardiovascular health is rarely addressed in clinical encounters focused on the person with ADRD. Nurses are well positioned to bridge this gap, yet they need a clearer understanding of what care partners experience in managing hypertension across the dyad to design effective, targeted interventions.

Objectives

To our knowledge, no study has analyzed online community health data to explore the dual challenges of providing hypertension management to persons living with ADRD while managing caregivers' own hypertension-related health needs. This study examined 2 interconnected aspects: (1) caregivers' perceptions and challenges in managing hypertension for individuals with ADRD and hypertension and (2) caregivers' own hypertension management challenges contributing to their chronic disease burden. Findings will inform nursing interventions supporting caregiver-patient dyads.

Methods

Study Design

This exploratory qualitative study used reflexive thematic analysis, following the Braun and Clarke (2006) 6-phase framework [27] to explore ADRD caregivers' experiences with hypertension management. This inductive approach allowed themes to emerge from the data rather than imposing pre-existing theoretical categories.

Study Setting and Data Source

Data were extracted from ALZConnected [23], a moderated online community platform established by the Alzheimer's Association. The forum provides message boards where care partners and persons across the ADRD continuum—including those with mild cognitive impairment and established dementia—discuss various topics related to dementia care. This forum represents the heterogeneity of ADRD subtypes rather

XSL•FO
RenderX

than reflecting a particular disease experience. While users must register to post content, all posts are publicly accessible without authentication. At the time of data extraction, 11,092 registered users contributed to the forum, with 94% (n=10,427) self-identifying as family caregivers (n=7,543, 68% adult children; n=1,331, 12% partners/spouses; n=776, 7% other relatives; n=776, 7% grandchildren/siblings);2% (n=222) as professional caregivers; and <1% (n=111) as persons with dementia; and 3% (n=333) did not specify a role.

Data Collection

Posts were extracted via web scraping between May 21 and 28, 2019. We used the search terms “hypertension” OR “ blood pressure” to identify potentially relevant posts, yielding an initial set of 1357 results. Posts were subsequently reviewed to confirm that each referenced both a care recipient with ADRD and the management of hypertension or blood pressure; posts discussing only one of these topics were excluded. From this filtered dataset, we systematically analyzed posts until reaching data saturation at 300 posts.

Ethical Considerations

This study was determined to be exempt from full ethics review by the institutional review board at the University of Texas Health Science Center at Houston (HSC20200784N), as it used deidentified, publicly available online forum data. All forum posts are publicly accessible without authentication, and usernames are pseudonyms; no personally identifiable information was collected or retained.

Data Analyses

Analysis followed a systematic 6-phase process:

- *Phase 1—Familiarization*: 2 trained individuals independently read all extracted posts (n=1357) to gain data immersion.
- *Phase 2—Initial coding*: using manual coding, researchers independently applied open coding to identify meaningful units within posts. An initial codebook containing 14 codes was developed after coding the first 50 posts.
- *Phase 3—Code refinement*: researchers independently coded an additional 100 posts, refining code definitions and identifying 9 additional codes, resulting in 23 total codes.
- *Phase 4—Theme development*: researchers met to review all codes and organize them into potential themes through iterative discussion and consensus building.
- *Phase 5—Theme refinement*: continuing independent coding, researchers achieved saturation at 300 posts, defined as the point at which no new codes emerged across 3 consecutive 25-post batches. All 300 posts were then reviewed to ensure thematic coherence and internal consistency.
- *Phase 6—Theme finalization*: the coding team, including a senior qualitative researcher (with more than 10 years’ experience), met to finalize theme names, definitions, boundaries, and relationships. Disagreements were resolved through discussion until consensus was reached.

Each post could be assigned multiple codes reflecting different dimensions of hypertension management experience. Final themes were organized hierarchically with primary themes and constituent subcodes (Table 1).

XSL•FO
RenderX

Table . Full list of all codes generated from the data.

Code name (number of posts)	Description
Theme 1: Pharmacotherapy (155/300, 51.7%)	
Medication in Hypertension Management (n=94)	Management of hypertension with medication(s). Includes both proper and poor management of hypertension with medication.
Medication Management (n=26)	Caregiver opinion or evaluation of the list of medications a loved one was taking or was prescribed. Includes caregivers having added or removed medications with or without physician approval.
Hypertension Medication Side Effects (n=24)	Side effects presumed from a hypertension medication, such as fainting, dizziness, and falls.
Medication Administration (n=11)	Logistics of medication administration, including ways that caregivers have gotten care recipients to take their medications, such as through the use of pill dispensers or crushing medication into food.
Theme 2: Patient Health Care Implementation Strategies (125/300, 41.7%)	
Care Coordination (n=57)	Coordination of care by the caregiver for an individual with ADRD[a]. Includes topics such as medicolegal issues, finding housing care, and decisions about hiring home care services to ease caregiver burden.
Interaction With Health Care Team (n=30)	Caregiver opinions on their interactions with, or decisions made by, medical personnel (any medical professional or hired caregivers) in charge of or familiar with their loved one's medical needs and care.
Responsibility (n=20)	Where the hypertension care management responsibility falls in terms of who is responsible and how much they are responsible for, such as if the individual with ADRD manages it themselves or if this is put on a caregiver.
Blood Pressure Evaluation (n=10)	Evaluating a loved one's blood pressure by either a physician or themselves as the caregiver. Includes confirmation or reevaluation of a previous hypertension diagnosis or taking a blood pressure reading and evaluating what to do.
Palliative Care Versus Health Management (n=8)	Decision of entering palliative care and how to address hypertension in this stage, if at all. Caregivers also discuss whether or not to continue hypertension medications when in palliative care or hospice.
Theme 3: Comorbid Care Management (136/300, 45.3%)	
Multiple Chronic Conditions Management (n=112)	Having a hypertension diagnosis along with another definitive disease diagnosis, including but not limited to ADRD, physical illness, and mental illness.
Caregiver's Theories About ADRD and Hypertension Relationships (n=13)	Speculation or confirmation of the link between ADRD and hypertension as perceived by the caregiver, author of the post. Discussed either as passing speculation or with supporting articles and reasoning.
ADRD as a Barrier to Hypertension Management (n=11)	How an ADRD diagnosis or its symptoms, such as forgetfulness or anosognosia, affect the care management of hypertension.
Theme 4: Hypertension Management and Outcomes (71/300, 23.7%)	
Hypertensive Episodes (n=23)	Particular occasions, time points, isolated situations, or environments that evoked a temporary elevation in blood pressure. A chronic hypertension diagnosis was not necessary for an episode to occur.
Consequences of Noncompliance (n=12)	Effects of not complying with proper hypertension care, or the consequences of poor or noncompliant hypertension management, such as further disease or adverse health events.
Possible Hypertension (n=11)	Appearance or mention of signs and symptoms of hypertension without simultaneous mention of a concrete diagnosis. Assumed to have hypertension but is not confirmed through the post.
Hypertension as a Side Effect (n=10)	Treatment, medication, or underlying comorbidities that may elicit hypertension as a side effect, such as certain oils or over-the-counter therapies.
Hypertension Exacerbation (n=8)	Situations or stressors, usually chronic, that result in the exacerbation of preexisting hypertension. Does not relate to posts for an undiagnosed person having a hypertensive episode.

XSL•FO
RenderX

Code name (number of posts)	Description
Hypertension to Normotension/Hypotension (n=7)	A person going from explicitly diagnosed hypertension to a state of normal blood pressure or even hypotension. Causes for this can include ADRD progression or as a result of or side effect of medications given.
Theme 5: Managing Behavioral Symptoms (38/300, 12.7%)	
Resistance to Care (n=31)	ADRD individual being resistant to care, creating complications for the caregiver when trying to provide care to their loved one. Anosognosia or disease progression leads to ADRD individuals resisting medical care, medications, appointments, and general management of their care.
Strategies for Compliance (n=7)	Strategies that caregivers employ, create, or discuss in an attempt to get their loved one to comply with care. Strategies often include “fiblets” or white lies to achieve compliance or cooperation in tasks such as visiting the doctor.
Theme 6: Peer Medical Advice Exchange in Online Forum (9/300, 3%)	
Unqualified Medical Advice in a Forum (n=9)	Forum user or caregiver who either takes medical advice from another user on the forum or gives their own medical advice within the forum with intentions to guide or assist another caregiver in the decision-making process.
Theme 7: Caregiver Burden and Self-Care for Hypertension (25/300, 8.3%)	
Caregivers With Hypertension (n=25)	Caregivers discuss their own hypertension, whether this is diagnosed, chronic, acute, spontaneous, or a concern for development in the future. Many caregivers attribute high blood pressure, whether diagnosed or situational, to the stress and demands of the caregiving role.

[a]ADRD: Alzheimer disease and related dementias.

Results

Overview of Thematic Analysis

Through reflexive thematic analysis of 300 posts from the ALZConnected online community forum, we identified 7 major themes reflecting the multifaceted nature of hypertension management in persons with ADRD. These themes captured both care partners’ experiences of managing their care recipients’ hypertension and their own hypertension-related challenges. Table 1 presents the themes, their constituent codes, and frequency distributions. In the narrative that follows, codes within each theme are discussed according to three criteria: (1) frequency—codes appearing in a greater number of posts are given priority to reflect the breadth of care partner experience, (2) prominence—codes that were most central to the theme’s conceptual meaning receive fuller treatment regardless of raw count, and (3) richness of illustrative quotations—codes supported by particularly vivid or representative posts are highlighted to preserve the depth of care partner voice. Where codes were closely related in content or addressed the same underlying experience, they were analytically subordinated under a higher-order code and discussed together to avoid redundancy; this is noted explicitly within each theme narrative below.

Theme 1: Pharmacotherapy (n=155, 51.7%)

Overview

This theme captured medication-related aspects of hypertension management, including efficacy concerns, polypharmacy challenges, side effects, and administration difficulties. Four codes comprised this theme: Medication in Hypertension Management (n=94), Medication Management (n=26), Hypertension Medication Side Effects (n=24), and Medication Administration (n=11).

Medication in Hypertension Management

Caregivers frequently discussed inadequate BP control despite medication use, expressing frustration with treatment failures:

> *She (care recipient) has had very high blood pressure for about 10 years that medication is not controlling as well as it should.*

These posts revealed caregivers’ concern about persistent hypertension despite adherence to prescribed regimens.

Medication Management

Caregivers independently evaluated and modified medication regimens, sometimes without provider consultation. Caregivers expressed skepticism about polypharmacy, perceiving that providers prescribed medications to treat side effects of other medications rather than addressing root causes:

> *My mom was on ten plus meds when she was diagnosed with ALZ/Dementia. I went through them and discovered her doctor was just prescribing meds, many of which were dealing with symptoms of previous meds....Anyway, I had her go cold turkey on the meds [so] we could ease her off of the meds [and] that needed to be done slowly (i.e. Prozac).*

This quote illustrates care partners’ autonomous decision-making regarding complex medication regimens, including the abrupt discontinuation of some medications. Notably, the post does not indicate whether medical supervision was involved, and such decisions may reflect distrust of or limited access to provider guidance rather than the absence of oversight.

XSL•FO
RenderX

Theme 2: Patient Health Care Implementation Strategies (n=125, 41.7%)

Overview

This theme encompassed caregivers' roles in coordinating health care, navigating the health care system, and managing relationships with providers. Five codes comprised this theme: Care Coordination (n=57), Interaction With Health Care Team (n=30), Responsibility (n=20), Blood Pressure Evaluation (n=10), and Palliative Care Versus Health Management (n=8).

Care Coordination Complexity

Caregivers described extensive coordination responsibilities spanning medicolegal arrangements (power of attorney, advance directives), specialist appointments, home care services, and housing transitions. Posts revealed the logistical burden of orchestrating multidisciplinary care:

> *Seven months ago my 66 yo SIL (sister in-law) came to a new state to live with my husband and me after a hospitalization due to a perfect storm of heavy drinking/high blood pressure/dementia. We established her with a new primary care, neurologist and psychiatrist.*

This example demonstrates the extensive coordination required following acute events, including establishing care in new locations and navigating insurance coverage for multiple services.

BP Evaluation

Caregivers discussed conducting home BP monitoring and interpreting readings, often without clear guidance on appropriate parameters or when to seek medical intervention:

> *I've been checking Dad's blood pressure at home like the doctor suggested, but I'm confused about what's actually concerning. This morning it was 158/92, this afternoon 142/88. His doctor said to 'keep an eye on it' but didn't tell me what numbers mean I should call. Is 158/92 an emergency? Do I wait and see if it goes down? I'm keeping a log but honestly don't know what I'm looking for.*

Theme 3: Comorbid Care Management (n=136, 45.3%)

Overview

This theme reflected caregivers' understanding of hypertension as one condition among multiple comorbidities and their perception that ADRD complicated hypertension management. Three codes comprised this theme: Multiple Chronic Conditions Management (n=112), Caregivers' Theories About ADRD and Hypertension Relationships (n=13), and ADRD as a Barrier to Hypertension Management (n=11).

Multiple Chronic Conditions Management

Caregivers described managing hypertension alongside numerous other conditions, revealing the complexity of care for medically complex older adults:

> *My dad is very 'weep', just diagnosed two weeks ago after countless falls and cognitive issues. He has other issues, CPAP machine use, 30 pounds overweight, glaucoma and meds taking for high blood pressure. He also suffers from anxiety*

This quote exemplifies the clustering of chronic conditions (ADRD, hypertension, obesity, glaucoma, sleep apnea, and anxiety) requiring simultaneous management.

Caregivers' Theories About ADRD and Hypertension Relationships

Some caregivers speculated about causal relationships between hypertension and ADRD development, drawing connections between poor vascular health and cognitive decline:

> *Mom has had high blood pressure & cholesterol issues for most of her life so at 79 perhaps this is the result. Lack of blood oxygen to the brain, etc.*

ADRD as a Barrier to Hypertension Management

Caregivers identified cognitive impairment as directly interfering with hypertension self-management by the care recipient with ADRD, including tasks such as remembering medications, understanding dietary restrictions, and recognizing symptoms requiring medical attention:

> *The hardest part is the medication schedule. Mom takes her blood pressure pill in the morning and another one at night, but she forgets constantly. I'll find her pills still in the container at bedtime, or she'll ask me three times in one day if she already took them. Last week, she took her morning dose twice because she couldn't remember taking it the first time.*

Theme 4: Hypertension Management and Outcomes (n=71, 23.7%)

Overview

This theme captured acute hypertensive situations, consequences of poor control, and care partners' observations of BP fluctuations. Six codes comprised this theme: Hypertensive Episodes (n=23), Effects of Hypertension Noncompliance (n=12), Possible Hypertension (n=11), Hypertension as a Side Effect (n=10), Hypertension Exacerbation (n=8), and Hypertension to Normotension/Hypotension (n=7).

Acute Hypertensive Episodes and Stress Reactivity

Caregivers described observing acute BP elevations, often triggered by emotional distress or behavioral symptoms:

> *Decided to check blood pressure since she is so upset and it's off the charts. I try to calm her down and within five minutes or so blood pressure is back to normal. Afraid she is going to have a heart attack or stroke!!*

This quote reveals caregivers' recognition of stress-BP relationships and their fear of acute cardiovascular events.

Consequences of Noncompliance

Caregivers connected medication nonadherence to adverse outcomes, particularly cardiovascular events and dementia progression:

XSL•FO
RenderX

> *Don't want to scare you, but my mom's dementia got worse when she had a heart attack. Her heart attack was caused by high blood pressure because she was taking her meds in a sporadic manner.*

Theme 5: Managing Behavioral Symptoms (n=38, 12.7%)

Overview

This theme addressed BPSD that complicated hypertension management. Two codes comprised this theme: Resistance to Care (n=31) and Strategies for Compliance (n=7).

Resistance to Care

Caregivers frequently described care recipients' refusal of medical appointments, medications, or BP monitoring, often attributing resistance to anosognosia (lack of illness awareness):

> *My mom refuses medical care also. Like your dad she could care less what anyone else is doing or what their opinion is.*

This resistance created significant management challenges, forcing caregivers to balance respecting autonomy with ensuring necessary care.

Strategies for Compliance

Caregivers shared strategies for circumventing resistance, including "therapeutic fibbing" or "white lies":

> *Fib, tell her it's a new program that they are using to train young nurses and caregivers, and her insurance wants her to "volunteer..." worth a shot that worked for my mom. [I] told her the insurance company required these for training and a study, to see if blood pressure taken at home was lower than when taken at the doctor's office.*

Theme 6: Peer Medical Advice Exchange in Online Forum (n=9, 3%)

This small but notable theme captured instances of forum users providing medical advice or recommendations to other caregivers, often without identifying professional credentials or citing evidence sources. One code comprised this theme: Peer Medical Advice (n=9).

Caregivers offered opinions on medication efficacy, treatment decisions, and care strategies based on personal experiences:

> *When I read that your Mom "came back" after only 2 days of restarting the Aricept, it just solidified my opinion that this drug may help more than the Drs even know….Its only been on the market for 12 years, which is not a SUPER long time in the pharmaceutical industry.*

Theme 7: Caregiver Burden and Self-Care for Hypertension (n=25, 8.3%)

This theme revealed caregivers' own hypertension experiences, often attributed to caregiving stress. One code comprised this theme: Caregivers With Hypertension (n=25).

Caregivers disclosed developing hypertension or experiencing BP elevations they directly linked to caregiving demands:

> *I was so frustrated with them and the situation I was literally getting sick myself. I needed counselling myself, my blood pressure went up (now I'm on BP meds); my stomach was always in knots.*

This quote exemplifies the bidirectional perceived health impact of caregiving: care partners not only managed their care recipients' hypertension but also attributed the development of hypertension themselves as a manifestation of chronic stress. The psychosomatic framing ("literally getting sick myself") reflected caregivers' understanding of stress-health connections. Few care partners discussed self-care strategies, suggesting limited attention to their own health needs relative to care recipient needs.

Discussion

Study Overview

This qualitative analysis of 300 posts from the ALZConnected forum revealed that informal care partners of persons with ADRD managing comorbid hypertension navigate multifaceted challenges spanning clinical decision-making, behavioral symptom management, and health care system coordination—often developing hypertension themselves as a consequence of caregiving. These findings illuminate the complex intersection of dementia caregiving and chronic disease management, highlighting both the clinical challenges care partners face and the dual role of online forums as sources of peer support and potential misinformation. Importantly, these findings should be interpreted with attention to the population they most likely represent: caregivers who are digitally engaged, motivated to seek peer support, and comfortable sharing health-related experiences in online communities. This profile suggests a relatively activated, help-seeking subset of the broader caregiving population—one that may have higher health literacy, greater social connectedness, or more access to resources than caregivers who do not engage with online forums. Clinicians and researchers should consider how these characteristics may shape the themes identified and whether caregivers who are less digitally engaged or more socially isolated face additional or distinct challenges not captured in this dataset.

Principal Results

The findings from this analysis reflect the profound daily burdens described in existing caregiving literature while offering additional context about how those burdens manifest specifically in the management of hypertension. Posts in the forum frequently depicted care partners as primary decision-makers for complex, chronic disease management—roles for which many felt underprepared. The co-occurrence of ADRD and hypertension created particularly layered challenges, as cognitive and behavioral symptoms of dementia often complicated adherence to hypertension regimens. That some care partners also reported developing hypertension themselves further illustrates how the caregiving role can blur the boundary between caregiver and care recipient, with implications for how

XSL•FO
RenderX

clinicians conceptualize and address the health of the caregiving dyad.

Comparison With Prior Work

ADRD caregiving places heavy emotional and practical demands on family members. Care partners often spend more than 150 hours per month helping with difficult cognitive and behavioral changes [28]. Caregiving can cause long-term physical and psychological strain and create extra stress in multiple life domains, such as work and family relationships, which may lead to health issues, including hypertension [29]. Research demonstrates that caregiving stress increases the risk of developing hypertension and heart disease [25]. Caregivers experiencing high levels of stress frequently neglect their own health-promoting behaviors and self-care. Studies showed that caregiving duties, lack of energy, and breakdown of social networks lead caregivers to defer their own care [16,26]. The posts analyzed in this study suggest that some care partners may manage both their loved ones' BP and their own, resulting in ongoing health challenges for both.

Major interrelated clinical challenges emerged from the analysis.

The theme of *Pharmacotherapy* revealed patterns of autonomous decision-making, with care partners independently modifying or discontinuing medications without consulting health care providers. For example, care partners described making unilateral decisions about BP medications. Although care partners value their loved one's autonomy in medication management, transitions to care partners taking over medication responsibilities often occur after care partners observe medication errors [30]. Care partners in the online forum frequently reported making decisions about medication changes while facing challenges related to managing multiple medications and complex dosage regimens [31]. These accounts align with research showing that caregivers assisting with medication management are often untrained, underresourced, and unsupported, even though caregivers of persons living with ADRD are often required to manage multiple simultaneous medications and complex regimens [32]. Care partner forum comments suggest that the autonomous decision-making described likely results from several factors: distrust related to perceived overmedication, concerns about certain medications and possible drug interactions, limited understanding of drug effects, and a lack of caregiver-focused medication education during clinical encounters. These findings suggest that nurses may be well positioned to address these concerns by providing caregiver education about medication management, including how to communicate concerns about overmedication with health care providers to avoid risky medication changes.

The theme of *Patient Health Care Implementation Strategies* revealed that care partners in this sample described extensive and often uncompensated coordination work—arranging specialist appointments, navigating medicolegal decisions, managing home care transitions, and monitoring BP without clear clinical guidance on when readings warranted intervention. These findings align with a growing literature documenting that family care partners frequently assume quasi-clinical roles for which they receive little preparation or support [9,10]. The gap between what care partners were asked to do—conduct and interpret home BP monitoring, coordinate multidisciplinary care teams—and the guidance they received during clinical encounters was a recurring source of frustration and uncertainty. These accounts underscore the need for explicit caregiver education at clinical touchpoints, including clear parameters for when to act on home BP readings and how to navigate specialist referral pathways.

The theme of *Comorbid Care Management* revealed that care partners consistently described hypertension not as an isolated condition but as one element of a dense cluster of comorbidities, including obesity, sleep apnea, glaucoma, anxiety, and ADRD itself. This framing reflects the clinical reality of older adults with ADRD, who carry an average of 5 or more chronic conditions simultaneously [4]. Notably, care partners in this forum identified ADRD's cognitive and behavioral features—memory failure, anosognosia, and care resistance—as direct impediments to hypertension self-management by the care recipient. This is an important mechanistic observation: the disease that motivates hypertension management also undermines the patient's capacity to participate in it. Clinicians should anticipate this dynamic and explicitly shift responsibility and education toward the care partner rather than assuming the person with ADRD can self-manage.

The theme of *Managing Behavioral Symptoms* revealed that care resistance—rooted in anosognosia or disease progression—emerged as a substantial impediment to routine hypertension management, including medication administration, BP monitoring, and medical appointments. Care partners described developing workarounds such as therapeutic fibbing to achieve cooperation, reflecting the extent to which they absorbed the clinical burden of managing a patient who could not consent to or comprehend their own care. These strategies, while pragmatically effective, raise ethical questions about autonomy and dignity that care partners were navigating without professional guidance. This theme reinforces the value of behavioral symptom training for care partners and highlights a gap in current clinical practice: the management of ADRD-related care resistance is rarely addressed in the context of chronic disease management protocols.

The theme of *Caregiver Burden and Self-Care for Hypertension* revealed that some care partners shared that they developed high BP or saw their BP rise because of the stress of caregiving, even though this topic only came up in a small number of posts. One caregiver said, "I was literally getting sick myself...my BP went up (now I'm on BP meds)." This quote illustrates how caregiving can affect both the caregiver and the person receiving care. Caregivers often had to handle their loved ones' high BP while also dealing with their own. These accounts are consistent with research suggesting that caregiving can cause ongoing physical and mental strain, leading to adverse health outcomes, including hypertension [29]. Not many care partners discussed their own health or ways to care for themselves, which may be because they focus more on their loved ones' needs than on their own, even though many caregivers have high BP. This pattern means that nurses should consider asking caregivers about their own health during appointments with care recipients, including urging care partners to maintain their own medical appointments. Nurses and other health care providers can check

XSL•FO
RenderX

for caregiver stress and high BP, discuss health risks, and encourage caregivers to prioritize their own self-care. Health care systems might consider developing programs that assist care partners in managing both the daily responsibilities of caregiving and the associated health effects, such as offering stress management, health screenings, and practical support. The Guiding an Improved Dementia Experience (GUIDE) pilot program being offered through Medicare, as well as new billing codes to address the needs of care partners, are important advancements that allow nurses to implement person- and family-centered care to support caregivers [33].

Given these substantial burdens, online support groups have emerged as vital resources where care partners seek social support to buffer against the negative health impacts of caregiving [24]. These online platforms provide accessible spaces for informational, emotional, and peer support among individuals who share similar experiences. Notably, online forums offer care partners safe spaces to disclose experiences that might be stigmatized in offline settings, such as expressions of anger or negative emotions toward the care recipient, with such disclosures typically met with understanding and reciprocal narratives from other group members [34]. According to the stress process model of caregiving, access to social support, such as that from peers, has the potential to improve an individual's appraisal of a stressor and buffer the impact of caregiving demands [16,35]. The forum analyzed in this study functioned as precisely this type of space, where caregivers disclosed their daily struggles with others navigating similar circumstances.

Due to the heavy demands on caregivers, online support groups have become crucial places to find help and connection [24]. These forums provide caregivers with a way to support one another, especially for those who live far from others or have limited time. Research shows that caregivers with the highest needs are often the most active in these groups [34]. Peer support can help caregivers handle stress and feel less alone [35]. The forum in this study served as a venue where caregivers could discuss their daily challenges with others facing similar issues. By sharing their experiences, such as discussing ways to help loved ones take their medicine or finding comfort in each other's stories, caregivers received both helpful advice and emotional support that they might not have otherwise. At the same time, some of the information provided on online forums may not be accurate. Nurses could have a role as facilitators within these forums to provide accurate information and debunk potentially harmful advice.

Across all 7 themes, several cross-cutting patterns emerged that reflect broader systemic challenges in dementia caregiving. Care partners consistently reported unmet informational needs—particularly regarding medication management, behavioral symptom interpretation, and care coordination—that were not adequately addressed during clinical encounters. A recurring consequence was autonomous or underinformed decision-making, in which care partners acted independently due to insufficient guidance from health care providers. Fragmented care coordination further compounded these challenges, with care partners frequently navigating multiple specialists and systems without centralized support. Together, these patterns suggest that the needs of care partners managing comorbid hypertension in persons with ADRD extend well beyond the clinical encounter and underscore the value of proactive, family-centered approaches to chronic disease management in this population.

Limitations

These findings should be viewed in light of several limitations. First, the data come from a single online forum (ALZConnected), so they may not reflect the experiences of all caregivers, especially those who do not use online communities or who use different platforms. Caregivers who post online may differ from those who do not in terms of stress, support-seeking, or health care knowledge. Second, since the forum posts are cross-sectional, we cannot track how caregiving experiences and challenges change over time. Third, limited demographic information about forum users made it impossible to analyze differences by age, gender, relationship to the care recipient, or socioeconomic status. The forum data also lacked race and ethnicity information, limiting generalizability to minoritized communities that may face distinct caregiving burdens and health disparities; future studies should intentionally recruit diverse samples. Additionally, we were unable to verify the specific etiology of care recipients' dementia (eg, Alzheimer disease vs vascular or mixed pathology) or other clinical details, which may influence caregiving challenges and the relevance of hypertension management. Fourth, geographic data were unavailable, limiting interpretation of findings related to access to care. Fifth, the self-reported, unverified nature of online forum posts introduces self-report and potential reporting bias. Sixth, the search strategy relied exclusively on fully spelled-out terms ("hypertension" and "BP") and did not include common medical abbreviations such as "hypertension," "BP," "systolic blood pressure," or "diastolic blood pressure." Caregivers with greater medical familiarity or clinical backgrounds may preferentially use these abbreviations, and potentially relevant posts from more health-literate users may therefore have been excluded from the dataset. This represents an additional source of selection bias that could skew findings toward caregivers with less clinical knowledge, and future studies should test whether abbreviation-based search terms yield meaningfully different or additional posts. Finally, thematic analysis is inherently interpretive; future mixed methods studies linking qualitative themes to clinical or registry data would strengthen validation of these findings. Prospective studies with comprehensive demographic data and registry-based diagnosis verification are needed to address these limitations.

Conclusions

This analysis of online forum posts shows that care partners for people with ADRD and high BP face many challenges, such as handling medications, managing difficult behaviors, working with the health care system, and dealing with their own health. In the current sample, care partners frequently reported making medication-related decisions independently, struggled to differentiate between transient and chronic BP elevations, and sought ways to manage care resistance associated with anosognosia. These findings highlight significant gaps in the training and support available to caregivers. Health care

XSL•FO
RenderX

providers and systems should recognize that care partners are essential members of the care team, but they often lack the preparation and support they need.

Future research should investigate ways to address the challenges highlighted here, such as training programs for care partners on medication management, tools to aid in interpreting BP readings, and methods for managing care resistance in individuals with limited awareness of their illness. Long-term studies could help us understand how caregivers' needs change as dementia progresses. Research on how professional moderators influence online support communities could help identify the most effective ways to support caregivers and prevent the spread of incorrect information. Additional studies are also needed to examine caregivers' heart health and whether targeted programs can help prevent or manage high BP in caregivers.

By learning from the experiences shared in these online forums, health care providers and systems can better support care partners in their crucial role and attend to their own health needs. Equipping care partners with targeted education, care coordination support, and access to reliable health information represents a meaningful step toward addressing the dual burden documented in this study.

Acknowledgments

Generative AI tools were not used in the preparation of this manuscript.

Funding

This study was supported by the Center on Smart and Connected Health Technologies, UT Health San Antonio School of Nursing. The funding source had no role in study design, data collection, analysis, interpretation, or manuscript preparation.

Data Availability

The datasets analyzed during this study are publicly available on ALZConnected [23]. Specific extracted posts are not shared to protect user privacy despite their public availability.

Authors' Contributions

JK and JW contributed to study conceptualization, data collection, analysis, and manuscript drafting. JK, KM, and JW contributed to manuscript revision and approved the final version for publication. JK and JW take responsibility for the integrity of the data analysis.

Conflicts of Interest

None declared.

References

1. 2024 Alzheimer's disease facts and figures. Alzheimers Dement 2024 May;20(5):3708-3821. [doi: 10.1002/alz.13809] [Medline: 38689398]
2. Zissimopoulos JM, Tysinger BC, St.Clair PA, Crimmins EM. The impact of changes in population health and mortality on future prevalence of Alzheimer's disease and other dementias in the United States. J Gerontol B Psychol Sci Soc Sci 2018 Apr 16;73(suppl_1):S38-S47. [doi: 10.1093/geronb/gbx147] [Medline: 29669100]
3. Chen C, Zissimopoulos JM. Racial and ethnic differences in trends in dementia prevalence and risk factors in the United States. Alzheimers Dement (N Y) 2018 Jan;4(1):510-520. [doi: 10.1016/j.trci.2018.08.009] [Medline: 30364652]
4. Livingston G, Sommerlad A, Orgeta V, et al. Dementia prevention, intervention, and care. The Lancet 2017 Dec;390(10113):2673-2734. [doi: 10.1016/S0140-6736(17)31363-6] [Medline: 28735855]
5. Lourida I, Hannon E, Littlejohns TJ, et al. Association of lifestyle and genetic risk with incidence of dementia. JAMA 2019 Aug 6;322(5):430-437. [doi: 10.1001/jama.2019.9879] [Medline: 31302669]
6. Risk reduction of cognitive decline and dementia: WHO guidelines. : World Health Organization; 2019 URL: https://www.who.int/publications/i/item/risk-reduction-of-cognitive-decline-and-dementia [accessed 2026-07-09]
7. Eldholm RS, Persson K, Barca ML, et al. Association between vascular comorbidity and progression of Alzheimer's disease: a two-year observational study in Norwegian memory clinics. BMC Geriatr 2018 May 22;18(1):120. [doi: 10.1186/s12877-018-0813-4] [Medline: 29788900]
8. Mukadam N, Sommerlad A, Huntley J, Livingston G. Population attributable fractions for risk factors for dementia in low-income and middle-income countries: an analysis using cross-sectional survey data. Lancet Glob Health 2019 May;7(5):e596-e603. [doi: 10.1016/S2214-109X(19)30074-9] [Medline: 31000129]
9. Burgdorf J, Roth DL, Riffin C, Wolff JL. Factors associated with receipt of training among caregivers of older adults. JAMA Intern Med 2019 Jun 1;179(6):833-835. [doi: 10.1001/jamainternmed.2018.8694] [Medline: 30958503]
10. Reinhard S. Home alone revisited: family caregivers providing complex care. Innov Aging 2019 Nov 8;3(Supplement_1):S747-S748. [doi: 10.1093/geroni/igz038.2740]

XSL•FO
RenderX

11. van der Lee J, Bakker TJEM, Duivenvoorden HJ, Dröes RM. Multivariate models of subjective caregiver burden in dementia: a systematic review. Ageing Res Rev 2014 May;15:76-93. [doi: 10.1016/j.arr.2014.03.003] [Medline: 24675045]
12. van der Lee J, Bakker TJEM, Duivenvoorden HJ, Dröes RM. Do determinants of burden and emotional distress in dementia caregivers change over time? Aging Ment Health 2017 Mar;21(3):232-240. [doi: 10.1080/13607863.2015.1102196] [Medline: 26586266]
13. Riffin C, Van Ness PH, Wolff JL, Fried T. Family and other unpaid caregivers and older adults with and without dementia and disability. J Am Geriatr Soc 2017 Aug;65(8):1821-1828. [doi: 10.1111/jgs.14910] [Medline: 28426910]
14. Rodakowski J, Skidmore ER, Rogers JC, Schulz R. Role of social support in predicting caregiver burden. Arch Phys Med Rehabil 2012 Dec;93(12):2229-2236. [doi: 10.1016/j.apmr.2012.07.004] [Medline: 22824248]
15. Gallagher D, Ni Mhaolain A, Crosby L, et al. Self-efficacy for managing dementia may protect against burden and depression in Alzheimer's caregivers. Aging Ment Health 2011 Aug;15(6):663-670. [doi: 10.1080/13607863.2011.562179] [Medline: 21547745]
16. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct;30(5):583-594. [doi: 10.1093/geront/30.5.583] [Medline: 2276631]
17. Cheng ST. Dementia caregiver burden: a research update and critical analysis. Curr Psychiatry Rep 2017 Aug 10;19(9):64. [doi: 10.1007/s11920-017-0818-2] [Medline: 28795386]
18. Shiba K, Kondo N, Kondo K. Informal and formal social support and caregiver burden: the AGES Caregiver Survey. J Epidemiol 2016 Dec 5;26(12):622-628. [doi: 10.2188/jea.JE20150263] [Medline: 27180934]
19. Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. JAMA 2014 Mar 12;311(10):1052-1060. [doi: 10.1001/jama.2014.304] [Medline: 24618967]
20. Hamm MP, Chisholm A, Shulhan J, et al. Social media use among patients and caregivers: a scoping review. BMJ Open 2013 May 9;3(5):e002819. [doi: 10.1136/bmjopen-2013-002819] [Medline: 23667163]
21. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Serv Res 2016 Aug 26;16(1):442. [doi: 10.1186/s12913-016-1691-0] [Medline: 27562728]
22. Parker Oliver D, Patil S, Benson JJ, et al. The effect of internet group support for caregivers on social support, self-efficacy, and caregiver burden: a meta-analysis. Telemed J E Health 2017 Aug;23(8):621-629. [doi: 10.1089/tmj.2016.0183] [Medline: 28328392]
23. ALZConnected. URL: https://alzconnected.org/ [accessed 2026-07-17]
24. Daynes-Kearney R, Gallagher S. Online support groups for family caregivers: scoping review. J Med Internet Res 2023 Dec 13;25:e46858. [doi: 10.2196/46858] [Medline: 38090796]
25. Lambrias A, Ervin J, Taouk Y, King T. A systematic review comparing cardiovascular disease among informal carers and non-carers. Int J Cardiol Cardiovasc Risk Prev 2023 Mar;16:200174. [doi: 10.1016/j.ijcrp.2023.200174] [Medline: 36874047]
26. Vitaliano PP, Zhang J, Scanlan JM. Is caregiving hazardous to one's physical health? A meta-analysis. Psychol Bull 2003 Nov;129(6):946-972. [doi: 10.1037/0033-2909.129.6.946] [Medline: 14599289]
27. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
28. Ajrouch KJ, Webster NJ, Zahodne LB, Antonucci TC. Introduction to ADRD care in context. J Gerontol B Psychol Sci Soc Sci 2023 Mar 13;78(Suppl 1):S1-S3. [doi: 10.1093/geronb/gbac181] [Medline: 36477194]
29. Schulz R, Sherwood PR. Physical and mental health effects of family caregiving. Am J Nurs 2008 Sep;108(9 Suppl):23-27. [doi: 10.1097/01.NAJ.0000336406.45248.4c] [Medline: 18797217]
30. O'Conor R, Russell AM, Pack A, et al. Managing medications among individuals with mild cognitive impairment and dementia: patient-caregiver perspectives. J Am Geriatr Soc 2024 Oct;72(10):3011-3021. [doi: 10.1111/jgs.19065] [Medline: 39007450]
31. Gench M, Sawan MJ, Langford A, Gnjidic D. Tools to evaluate medication management for caregivers of people living with dementia: a systematic review. Health Expect 2021 Oct;24(5):1570-1581. [doi: 10.1111/hex.13318] [Medline: 34289214]
32. Werner NE, Campbell NL, Boustani M, Ganci A, Holden RJ. Helping the helpers - a research protocol for user-centered technology to aid caregiver management of medications for people with Alzheimer's disease and related dementias. Res Social Adm Pharm 2022 Sep;18(9):3680-3686. [doi: 10.1016/j.sapharm.2022.03.014] [Medline: 35396151]
33. Burgdorf JG, Freedman VA, Wolff JL. A look at caregivers for community-living persons with dementia: implications for the GUIDE model. Alzheimers Dement 2025 Mar;21(3):e70013. [doi: 10.1002/alz.70013] [Medline: 40156263]
34. Friedman EM, Trail TE, Vaughan CA, Tanielian T. Online peer support groups for family caregivers: are they reaching the caregivers with the greatest needs? J Am Med Inform Assoc 2018 Sep 1;25(9):1130-1136. [doi: 10.1093/jamia/ocy086] [Medline: 30016449]
35. Price ML, Surr C, Gough B, Howe D, Ashley L. The provision of social support in an online support forum for caregivers of people with comorbid dementia and cancer: content analysis study. JMIR Cancer 2025 Sep 29;11:e72217. [doi: 10.2196/72217] [Medline: 41022373]

XSL•FO
RenderX

Abbreviations

ADRD: Alzheimer disease and related dementias
BP: blood pressure
BPSD: behavioral and psychological symptoms of dementia
GUIDE: Guiding an Improved Dementia Experience

Edited by T Risling; submitted 22.Oct.2025; peer-reviewed by J Kramer, M Split; revised version received 04.May.2026; accepted 26.May.2026; published 28.Jul.2026.

Please cite as:
Ko J, Meyer K, Wang J
Care Partner Perspectives on Hypertension Management in Alzheimer Disease and Related Dementias in an Online Forum: Qualitative Thematic Analysis
JMIR Nursing 2026;9:e85818
URL: https://nursing.jmir.org/2026/1/e85818
doi:10.2196/85818

© Jisook Ko, Kylie Meyer, Jing Wang. Originally published in JMIR Nursing (https://nursing.jmir.org), 28.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Web-Based Mindfulness Intervention for Reducing Perceived Stress in Health Care Professionals: Secondary Analysis of Moderators in a Stepped-Wedge Cluster Randomized Trial

Gloria Guerrero-Pertiñez[1*], PhD; Vera Carbonell Aranda[1,2], Clin Psych; Jonathan Joseph Dawood-Histrova[1], MS; Gloria Pérez-Guerrero[3,4], MD; Adrián Pérez-Aranda[5,6], PhD; Selene Fernández-Martínez[7], PhD; Alicia Monreal-Bartolomé[8,9,10], PhD; Alberto Barceló-Soler[8,9,10], PhD; Javier García-Campayo[7,9], PhD; Yolanda López del Hoyo[8,9,10], PhD; Jesús Herrera-Imbroda[1,4], PhD; Jessica Marian Goodman-Casanova[1,4*], PhD; Jose Guzman-Parra[1,4*], PhD

[1]Grupo C-03: Investigación Básica, Clínica y Epidemiológica en Salud Mental, IBIMA Plataforma BIONAND, Instituto de Investigación Biomédica de Málaga, Plaza del Hospital Civil s/n, Málaga, Andalucía, Spain

[10]Research Network on Chronicity, Primary Care and Health Promotion (RICAPPS), Zaragoza, Aragón, Spain

[2]Hospital Universitario Virgen de la Victoria, Málaga, Andalucía, Spain

[3]Grupo C-06: Neuropsicofarmacología, IBIMA Plataforma BIONAND, Instituto de Investigación Biomédica de Málaga, Málaga, Andalucía, Spain

[4]Unidad de Gestión Clínica en Salud Mental, Hospital Regional Universitario de Málaga, Málaga, Andalucía, Spain

[5]Department of Clinical and Health Psychology, Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain

[6]Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública, Madrid, Madrid, Spain

[7]Faculty of Medicine, Universidad de Zaragoza, Zaragoza, Aragón, Spain

[8]Department of Psychology and Sociology, Universidad de Zaragoza, Zaragoza, Aragon, Spain

[9]Instituto de Investigación Sanitaria Aragón, Zaragoza, Aragon, Spain

[*]these authors contributed equally

Corresponding Author:
Gloria Guerrero-Pertiñez, PhD
Grupo C-03: Investigación Básica, Clínica y Epidemiológica en Salud Mental, IBIMA Plataforma BIONAND, Instituto de Investigación Biomédica de Málaga, Plaza del Hospital Civil s/n, Málaga, Andalucía, Spain

Abstract

Background: Health care professionals are exposed to high levels of occupational stress, emotional exhaustion, and anxiety-depressive symptoms. Digital interventions using mindfulness- and acceptance-based approaches have shown promising results in improving psychological well-being in this population. The MINDxYOU program is a self-guided digital intervention designed to reduce perceived stress and enhance emotional regulation, resilience, and psychological flexibility. Although its effectiveness has been demonstrated, treatment response may vary substantially between individuals.

Objective: The aims of this study are to explore the sociodemographic, occupational, and psychological moderators of the effectiveness of the digital MINDxYOU program on perceived stress among health care professionals and to identify differential response profiles.

Methods: This study is a secondary analysis of a stepped-wedge cluster randomized trial including 347 health care professionals from 6 clusters allocated to 3 intervention start sequences. The intervention (MINDxYOU) is a self-guided, web-based program designed to be completed over 8 weeks, grounded in third-wave psychological approaches and incorporating mindfulness-, compassion-, and acceptance-based components. Weekly support was provided via WhatsApp Messenger (Meta Platforms Inc), phone calls, or email to promote adherence. Assessments were conducted every 8 weeks across 5 measurement waves. Linear mixed-effects models were used to evaluate moderation effects by testing condition × moderator interactions, accounting for repeated measures within participants and clustering effects.

Results: Moderation analyses showed greater reductions in perceived stress among participants with higher baseline depressive symptom scores (9-item Patient Health Questionnaire [PHQ-9]; B=−0.22, SE 0.08; *P*=.004) and those with higher scores in the observing facet of mindfulness (15-item Five Facet Mindfulness Questionnaire [FFMQ-15]; B=−0.86, SE 0.41; *P*=.04). In addition, nursing staff exhibited greater reductions in perceived stress compared with physicians (B=1.92, SE 0.81; *P*=.02) and other health care professionals (B=1.80, SE 0.90; *P*=.04).

Conclusions: The results suggest that baseline psychological and occupational characteristics may contribute to differences in the effectiveness of the MINDxYOU program in reducing perceived stress. These findings provide preliminary evidence regarding

XSL•FO
RenderX

potential response profiles in digital mental health interventions and support the development of more personalized and targeted approaches to stress management in health care settings.

Trial Registration: ClinicalTrials.gov NCT05436717; https://clinicaltrials.gov/study/NCT05436717

International Registered Report Identifier (IRRID): RR2-10.1186/s12912-022-01089-5

(JMIR Nursing 2026;9:e98624) doi:10.2196/98624

KEYWORDS

perceived stress; health care professionals; digital intervention; mindfulness; depression; resilience; mental health

Introduction

The increasing levels of occupational stress among health care professionals represent a growing public health concern, with implications at both the individual and organizational levels. Prolonged exposure to high emotional demands, work overload, extended shifts, and ethical conflicts generates emotional exhaustion, compassion fatigue, and a high risk of burnout [1,2]. This situation has worsened in recent years due to the COVID-19 pandemic, which intensified health care pressure and increased levels of anxiety, depression, and psychological distress among health care workers [3,4].

The impact of chronic stress in this population affects not only professionals' well-being but also quality of care, patient safety, and the sustainability of health care systems. Recent studies have shown that workers experiencing high levels of burnout are more likely to commit medical errors, show reduced empathy, and report greater intentions to leave the profession [5]. In this context, promoting the mental health and psychological resilience of health care professionals has become a strategic priority at both the institutional and scientific levels.

Stress-reduction programs delivered face-to-face have demonstrated effectiveness; however, their implementation is often limited by time constraints, rotating shifts, and the emotional demands inherent to clinical practice. In this context, digital interventions based on third-wave cognitive behavioral therapies, such as acceptance and commitment therapy, mindfulness-based interventions, and compassion-focused approaches, have demonstrated effectiveness in reducing psychological distress and improving emotional coping across different occupational and clinical settings [6,7]. These digital interventions offer several advantages, including greater accessibility, lower costs, and scalability, which facilitate their implementation in high-demand health care environments [8]. Recent evidence has further highlighted the potential of digital mental health interventions to support workforce well-being and address unmet mental health needs in occupational settings [9-11]. These findings are particularly relevant in health care settings, where digital interventions have demonstrated promising effects on stress reduction, burnout prevention, and psychological well-being among health care professionals [9,10,12-14].

Within this framework, MINDxYOU was developed as a self-guided digital program designed to promote emotional self-regulation, acceptance, and resilience among health care professionals. Results from a stepped-wedge cluster randomized trial demonstrated that MINDxYOU significantly reduced levels of stress, anxiety, and depression in this population [12]. A subsequent mediation analysis revealed that improvements in resilience constituted a significant psychological mechanism underlying the observed intervention effects [15].

However, even when interventions show overall effectiveness, treatment response may vary substantially across individuals, with previous studies also showing substantial heterogeneity in the outcomes associated with this type of intervention [16]. Increasing attention has been devoted to identifying the moderators of treatment effects, defined as personal or contextual characteristics that determine for whom and under which conditions an intervention is the most effective [9,17]. This approach has gained increasing relevance in digital mental health research, where recent evidence highlights personalization and adaptive intervention strategies as key factors for improving intervention effectiveness, engagement, and implementation outcomes [17-19].

Although the MINDxYOU program has shown beneficial effects in reducing perceived stress among health care professionals, less is known about whether these effects may differ according to participants' baseline characteristics. Exploring potential sources of variability in response is relevant for advancing the personalization of digital mental health interventions and for improving their applicability in health care settings. Therefore, the present study aimed to examine whether individual, psychological, and occupational characteristics were associated with differential responses to the MINDxYOU program in terms of perceived stress reduction.

Methods

Design

This study followed a stepped-wedge cluster randomized trial design, in which 6 clusters of health care professionals were randomly assigned to 3 intervention start sequences. Each sequence began the intervention at different time points separated by 2-month intervals, ensuring that all groups eventually received the program. This design allowed comparisons between participants who had not yet accessed the intervention and those who had already completed it while maintaining temporal control of the intervention effects.

Participants

The sample included 347 health care professionals. Inclusion criteria were (1) employment as a health care professional (eg, physician, nurse, psychologist, or nursing assistant) or enrollment in training in any health-related field, (2) age between 18 and 70 years, (3) ability to understand Spanish, (4)

XSL•FO
RenderX

digital literacy and access to an internet-connected smartphone, tablet, or computer, and (5) expected continued employment in the same workplace for the following 6 months. Participants with incomplete data in the main study variables or with significant changes in their employment status during the follow-up period were excluded.

Participants were grouped into 6 clusters according to their workplace, and these clusters were randomly assigned to an intervention sequence (MINDxYOU) using a computerized simple randomization procedure. Assessments were conducted every 8 weeks across 5 time points throughout the study. At each measurement point, participants completed the same self-report questionnaires through the program's digital platform.

Intervention

The MINDxYOU program is a self-guided, web-based intervention consisting of 4 modules delivered over 8 weeks, integrating mindfulness-, compassion-, and acceptance-based strategies. Content was delivered through videos, audio-guided practices, and interactive exercises. Participants received weekly support via WhatsApp Messenger (Meta Platforms Inc), phone, or email to promote adherence. In the stepped-wedge design, participants in the control condition were placed on a waiting list and received the intervention at later time points. Further details on the intervention components, module contents, and delivery procedures are provided in the study protocol [20].

Ethical Considerations

The study protocol was approved by the Research Ethics Committee of the Autonomous Community of Aragon and the Ethics and Research Committee of Northeast Malaga in July 2022 (approval code PI22/341). All participants provided informed consent prior to participation, and the study was conducted in accordance with the Declaration of Helsinki and European data protection regulations [21].

Measures

In the MINDxYOU clinical trial, a comprehensive set of self-report instruments was administered to assess clinical and psychological variables. Detailed information regarding the structure, reliability, and validity of these measures is described in the study protocol [20].

The primary outcome was perceived stress, assessed using the 10-item Perceived Stress Scale (PSS-10; [22]; Spanish version [23]). Secondary outcome variables included (1) depressive symptoms, assessed with the 9-item Patient Health Questionnaire (PHQ-9; [24]; Spanish version [25]); (2) generalized anxiety, measured with the 7-item Generalized Anxiety Disorder ([26]; Spanish version [27]); and (3) general psychological symptoms, assessed using the 18-item Brief Symptom Inventory ([28]; Spanish version [29]).

Regarding psychological processes associated with change, mindfulness was assessed using the 15-item Five Facet Mindfulness Questionnaire (FFMQ-15; [30]; Spanish version [31]). Psychological resilience was measured using the Connor-Davidson Resilience Scale (CD-RISC; [32]; Spanish version [33]). Compassion and self-compassion were assessed using the Sussex-Oxford Compassion Scales ([34]; Spanish version [35]) and the Acceptance and Action Questionnaire-II ([36]; Spanish version [37]).

Variable Selection

The selection of moderator variables was conducted in several stages. First, a panel of researchers defined an initial set of potential moderators based on the previous literature on stress-reduction programs for health care professionals and on the theoretical plausibility of the programs' relationship with treatment response [38]. Subsequently, a correlation matrix was examined to assess the relationships between baseline variables and perceived stress symptoms. To avoid problems of multicollinearity and conceptual redundancy, priority was given to variables that showed moderate correlations with baseline perceived stress symptoms while representing distinct psychological or contextual constructs. Variables with correlations greater than r=0.65 were excluded.

Following this procedure, a final set of moderators was defined, including (1) sociodemographic variables, including age and gender; (2) work-related variables, including type of contract (temporary vs permanent), occupation (physician, nurse [nurse or nursing assistant], or other health care professional), management position (yes/no), and workplace setting (hospital, primary care center, nursing home, or other); (3) clinical variables, including depressive symptoms measured with the PHQ-9; and (4) process variables, including psychological resilience, measured with the CD-RISC, and the observing facet of mindfulness, assessed with the FFMQ-15.

Additionally, variables related to the intervention were analyzed separately in order to explore whether the level of participation and the timing of intervention initiation influenced treatment outcomes. These variables included the intervention sequence, corresponding to the different start times established by the stepped-wedge design. Adherence was also considered and was defined as completion of at least 3 (out of 4) modules among participants who initiated the program.

Thus, the main set of moderators included sociodemographic factors, occupational factors, clinical factors, and process variables, while intervention-related variables were analyzed in a separate block to provide complementary information regarding the role of program use and adherence in the effectiveness of the intervention, as well as the potential influence of delayed initiation of the intervention across sequences.

Statistical Analysis

Data were analyzed using linear mixed-effects regression models to account for the longitudinal and clustered structure of the data. Repeated assessments were nested within participants, and participants were clustered within health care centers. The primary outcome was the PSS-10 total score. For each potential moderator, a separate mixed-effects model was fitted. Each model included the condition (intervention vs control), the moderator, and the condition × moderator interaction as fixed effects. The interaction term was used to test whether the effect of the intervention on perceived stress differed according to the baseline value or category of the moderator. All models were adjusted for baseline PSS-10 score and assessment time.

XSL•FO
RenderX

Random intercepts were specified at the participant and cluster levels. Collinearity among baseline predictors was examined using a correlation matrix before fitting the models. All potential moderators were then entered into the same multivariable mixed-effects model. This model was used to examine whether each moderator was independently associated with a differential intervention response after adjusting for the remaining moderators. Intervention-related variables, including intervention sequence and adherence, were analyzed separately using mixed-effects models adjusted for the same baseline covariates. All analyses were conducted in R software (version 4.4.1; R Foundation for Statistical Computing) using the *nlme* and *lme4* packages [39]. A 2-tailed significance level of .05 was adopted for all statistical tests.

Results

Sample Characteristics

The final sample consisted of 347 health care professionals, including 297 (85.6%) women with a mean age of 45.01 (SD 11.2) years. Participants were classified as physicians (n=147, 42.4%), nursing staff including nurses and nursing assistants (n=93, 26.8%), and other health care professionals (n=107, 30.8%). Further details of the baseline characteristics of the overall sample and across the randomized intervention sequences are presented in Table 1.

XSL•FO
RenderX

Table . Baseline characteristics of the study sample.

Variables	Total sample (N=357)	Sequence 1 (n=138)	Sequence 2 (n=135)	Sequence 3 (n=74)
Sociodemographic characteristics				
Gender, women, n (%)	297 (85.6)	117 (84.2)	114 (84.4)	66 (89.2)
Age (years), mean (SD)	45 (11.2)	47 (11)	43.5 (10.1)	44.2 (12.8)
Work-related aspects, n (%)				
Type of contract				
Temporary contract	64 (18.4)	20 (13.9)	32 (23.2)	12 (16.0)
Permanent contract	283 (81.6)	124 (86.1)	106 (76.8)	63 (84.0)
Occupation				
Physician	147 (42.4)	63 (43.8)	50 (36.2)	34 (45.3)
Nursing staff	93 (26.8)	37 (25.7)	37 (26.8)	19 (25.3)
Other health care professional	107 (30.8)	44 (30.6)	51 (37.0)	22 (29.3)
Workplace setting				
Hospital	85 (23.8)	37 (25.7)	48 (34.8)	0 (0)
Primary care center	92 (25.8)	48 (33.3)	0 (0)	44 (58.7)
Nursing home	45 (12.6)	1 (0.7)	26 (18.8)	18 (24)
Other setting	135 (37.8)	58 (40.3)	64 (46.4)	13 (17.3)
Management role, yes	52 (15)	20 (14.5)	25 (18.5)	7 (9.5)
PSS-10[a], mean (SD), score range 0-40	16.9 (6.3)	17.4 (6.7)	15.8 (5.9)	18 (6.3)
Secondary outcome and process variables, mean (SD)				
PHQ-9[b], score range 0 - 27	6.3 (4.4)	6.6 (4.4)	5.7 (4.3)	6.6 (4.3)
CD-RISC[c], score range 0 - 40	27.3 (6.8)[d]	27 (6.9)	28.2 (6.7)	26.3 (6.7)
Observing facet of FFMQ-15[e], score range 1-5	2.8 (0.9)	2.9 (0.8)[d]	2.9 (0.9)	2.8 (0.9)[f]

[a]PSS-10: 10-item Perceived Stress Scale.

[b]PHQ-9: 9-item Patient Health Questionnaire.

[c]CD-RISC: Connor-Davidson Resilience Scale.

[d]One missing value.

[e]FFMQ-15: 15-item Five Facet Mindfulness Questionnaire.

[f]Two missing values.

Moderation Analysis

Professional occupation showed a significant moderation effect. Using nursing staff as the reference category, physicians (B=1.92, SE 0.81; *P*=.02) and other health care professionals (B=1.80, SE 0.90; *P*=.04) exhibited smaller reductions in perceived stress following the intervention. Further details of the univariate moderation analyses are presented in Table 2. Figure 1 graphically illustrates the changes in perceived stress levels during the preintervention and postintervention periods across professional groups.

XSL•FO
RenderX

Table . Univariate moderation analyses of the effect of the intervention on perceived stress.

Moderator	Estimated B[a] (SE)[b]	*t* test (*df*)	*P* value
Gender (reference: woman)	0.97 (0.82)	1.19 (497)	.24
Condition	−0.38 (1.22)	−0.31 (759)	.76
Condition × gender	−0.98 (0.95)	−1.03 (730)	.30
Age	0.02 (0.03)	0.90 (461)	.37
Condition	−1.24 (1.56)	−0.79 (775)	.43
Condition × age	−0.01 (0.03)	−0.19 (756)	.85
Employment type (temporary; reference: nontemporary)	0.07 (0.77)	0.09 (509)	.92
Condition	−1.443 (0.567)	−2.55 (814)	.01
Condition × employment	−0.30 (0.96)	−0.31 (776)	.76
Management position (reference: no)	−0.52 (0.80)	−0.65 (470)	.52
Condition	0.34 (1.22)	0.28 (767)	.78
Condition × management	−1.58 (0.95)	−1.66 (750)	.10
Workplace setting (reference: hospital)			
Primary care center	−1.22 (0.81)	−1.50 (491)	.13
Nursing home	−0.09 (0.97)	−0.09 (441)	.93
Other	−1.05 (0.75)	−1.392 (511)	.16
Condition	−1.35 (0.821)	−1.646 (766)	.10
Condition × primary care center	0.81 (0.98)	0.832 (755)	.41
Condition × nursing home	−0.76 (1.24)	−0.609 (725)	.54
Condition × other	−0.62 (0.87)	−0.717 (737)	.47
Occupation (reference: nurse)			
Physician	−0.63 (0.70)	−0.898 (481)	.37
Other	0.36 (0.77)	0.473 (483)	.64
Condition	−2.80 (0.76)	−3.674 (782)	<.001
Condition × physician	1.92 (0.81)	2.357 (741)	.02
Condition × other	1.80 (0.90)	2.014 (747)	.04
Baseline depression (PHQ-9[c])	0.27 (0.08)	3.257 (405)	<.001
Condition	−0.082 (0.728)	−0.113 (776)	.91
Condition × baseline depression	−0.22 (0.08)	−2.872 (757)	.004
Resilience (CD-RISC[d])	−0.142 (0.048)	−2.936 (414)	.003
Condition	−2.242 (1.41)	−1.596 (749)	.11
Condition × resilience	0.02 (0.05)	0.505 (732)	.61
Mindfulness—observing (FFMQ-15[e])	0.170 (0.342)	0.497 (458)	.62
Condition	0.92 (1.31)	0.697 (769)	.49
Condition × mindfulness—observing	−0.86 (0.41)	−2.073 (752)	.04

[a]B: unstandardized regression coefficient.

[b]Linear mixed-effects models include assessment time and baseline perceived stress as covariates.

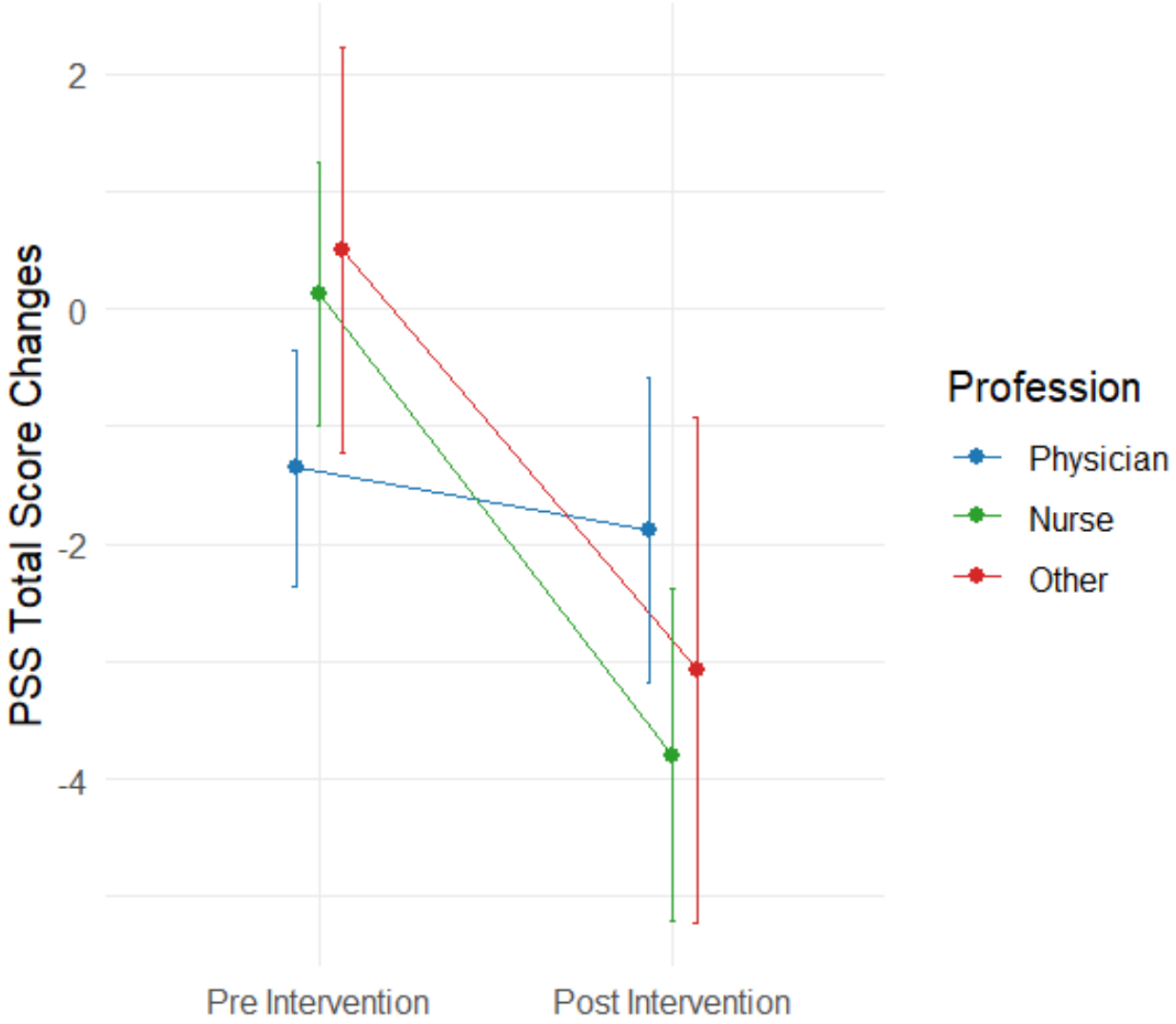
[c]PHQ-9: 9-item Patient Health Questionnaire.

[d]CD-RISC: Connor-Davidson Resilience Scale.

XSL•FO
RenderX

[e]FFMQ-15: 15-item Five Facet Mindfulness Questionnaire.

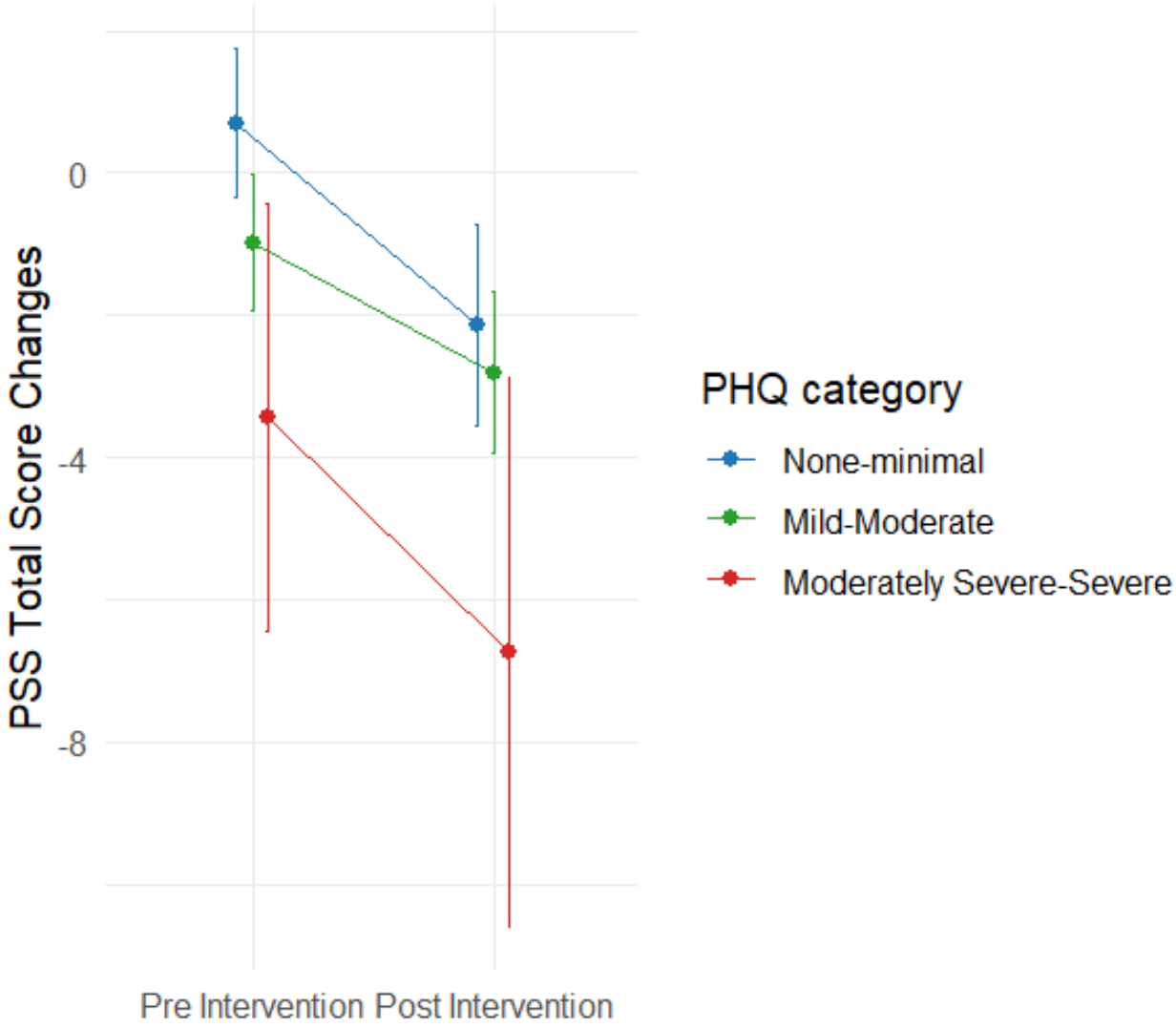
Figure 1. Changes in perceived stress using the 10-item Perceived Stress Scale (PSS-10) by professional category. Error bars indicate 95% CI.



Baseline depressive symptoms significantly moderated the intervention effect on perceived stress (B=−0.22, SE 0.08; *P*=.004). Higher baseline depression symptom scores were associated with greater reductions in perceived stress following the intervention. Further details of these analyses are presented in Table 2. Figure 2 illustrates the reductions in perceived stress during the preintervention and postintervention periods across 3 groups defined by baseline depressive symptom severity.

XSL•FO
RenderX

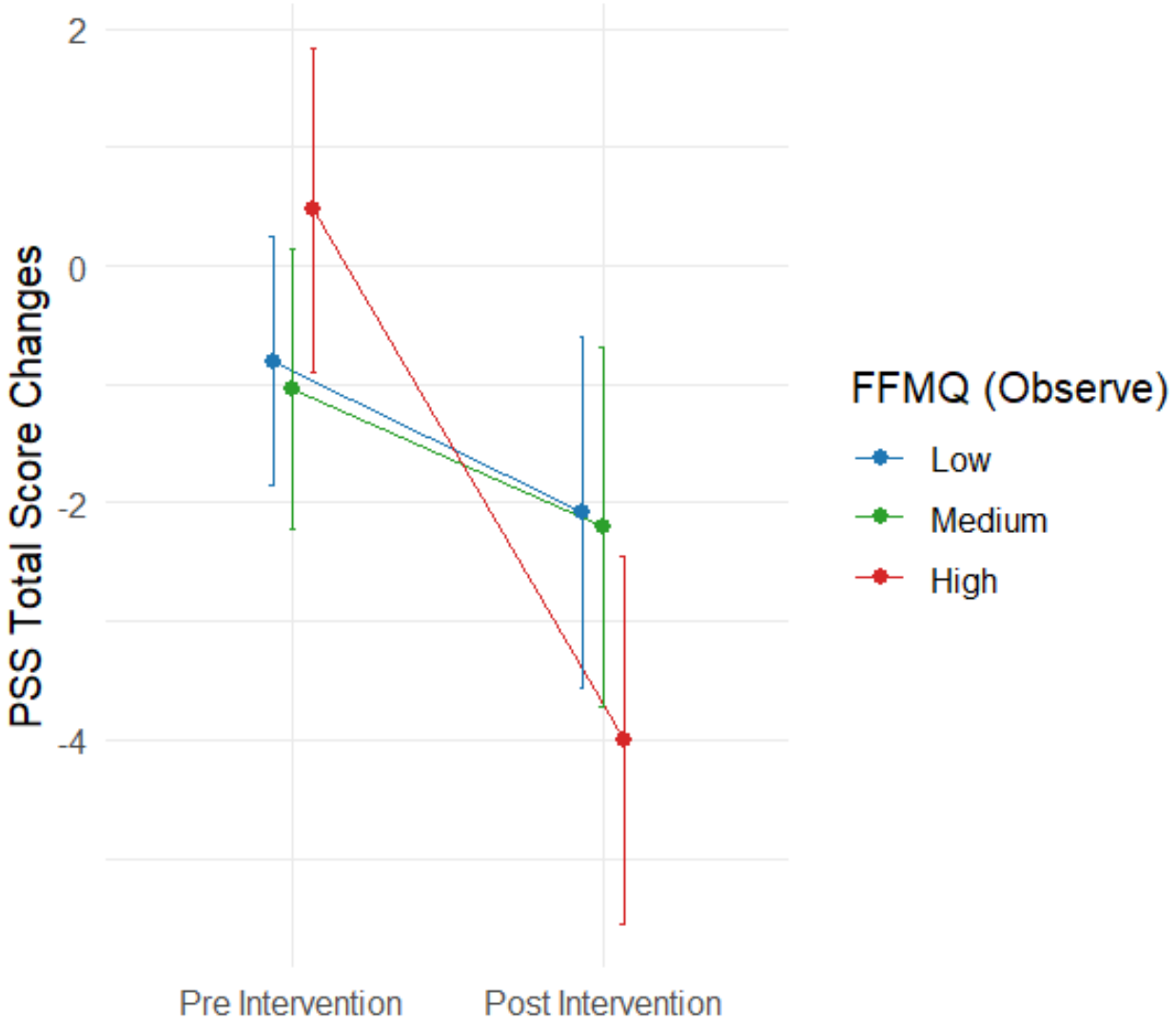
Figure 2. Changes in perceived stress using the 10-item Perceived Stress Scale (PSS-10) before and after the intervention by 9-item Patient Health Questionnaire (PHQ-9) depression severity category. None-minimal (0 - 4 score), mild-moderate (5 - 14 score), and moderately severe to severe depression (15 - 27 score). Error bars indicate 95% CI.



Additionally, the observing facet of mindfulness (FFMQ-15) also showed a significant moderation effect (B=−0.86, SE 0.41; *P*=.04). Health care professionals with higher baseline levels of mindful observation exhibited a more favorable response to the intervention. Further details of these analyses are presented in Table 2. Figure 3 illustrates the reductions in perceived stress during the preintervention and postintervention periods across 3 groups defined by baseline levels of the observing facet (low, medium, and high).

XSL•FO
RenderX

Figure 3. Changes in perceived stress using the 10-item Perceived Stress Scale (PSS-10) by levels of the 15-item Five Facet Mindfulness Questionnaire (FFMQ-15) observing facet. Low, medium, and high FFMQ—observing groups are defined using baseline tertiles. Error bars indicate 95% CI.



The results of the multivariate analysis are presented in Multimedia Appendix 1 and confirm the findings obtained in the univariate models. The moderating effects of baseline depressive symptoms (PHQ-9), nursing occupation, and the observing facet of mindfulness (FFMQ-15) on the effectiveness of the program remained statistically significant. In the multivariate model, these variables continued to play a significant moderating role in the reduction of perceived stress, whereas the remaining sociodemographic and occupational factors did not reach statistical significance.

In addition, the role of program adherence and the temporal sequence of the stepped-wedge design was explored. In the adjusted mixed-effects models, completing the program was not significantly associated with changes in perceived stress (B=−0.65, SE ##; *P*=.33), and no significant interaction was observed between the experimental condition and program completion (B=−0.99, SE 0.78; *P*=.20). Similarly, the temporal sequence of intervention implementation was not significantly associated with changes in perceived stress (B=−0.45, SE ##; *P*=.25) and did not moderate the intervention effect (B=0.61, SE ##; *P*=.19).

Discussion

Principal Findings

The present study explored individual, occupational, and psychological characteristics that may be associated with differential responses to the digital MINDxYOU program in reducing perceived stress among health care professionals. The findings suggest that baseline depressive symptoms, nursing occupation, and the observing facet of mindfulness may be potential moderators of intervention response. Although these results should be interpreted cautiously, they may provide preliminary insights into the characteristics of participants who could benefit most from this type of digital intervention. In this sense, the findings may contribute to ongoing efforts to develop more personalized and context-sensitive digital mental health programs for health care professionals.

First, the most robust finding of this study was the moderating role of baseline depressive symptoms. Participants with higher levels of depressive symptomatology at baseline benefited more from the intervention, showing larger reductions in perceived stress. This aligns with previous findings from an

XSL•FO
RenderX

internet-delivered mindfulness-based cognitive therapy study [40], in which more severe baseline depression predicted greater posttreatment improvement in anxiety among cancer survivors. Although further research is needed, this suggests that individuals with greater initial distress may benefit more from online mindfulness-based interventions. This finding is consistent with the literature describing a bidirectional relationship between stress and depression, in which chronic stress acts as both a precipitating and a maintaining factor of depression, while depressive symptoms increase threat perception and reduce coping resources [41]. From this perspective, it is plausible that an intervention based on acceptance and commitment principles and on mindfulness, both of which have demonstrated efficacy in reducing depressive symptoms, may be particularly beneficial in this subgroup [42,43]. These interventions promote acceptance of internal experiences, psychological flexibility, and self-compassion, processes that counteract the cognitive mechanisms underlying depression, such as rumination and experiential avoidance [44]. Thus, participants with more severe depressive symptoms may have benefited more from acquiring adaptive coping and self-compassion strategies targeting processes common to both stress and depression. These findings suggest the potential usefulness of MINDxYOU for health care professionals with greater emotional vulnerability.

Second, professional occupation showed a significant moderating effect. The results indicated that the magnitude of change in perceived stress differed across professional groups, with nursing staff showing greater reductions after the intervention than physicians and other health care professionals. These findings should be interpreted cautiously, as the present analyses do not establish the mechanisms underlying these differences. Nevertheless, professional roles in health care may involve distinct emotional, organizational, and caregiving demands, which could influence how participants engage with and respond to stress-reduction interventions [2,6]. These findings highlight the importance of adapting the content and support structure of digital interventions according to the specific emotional demands associated with different health care roles.

Finally, the observing facet of mindfulness appeared to moderate the intervention effect, suggesting that observing may be associated with differences in response to the program. Previous evidence has shown that mindfulness facets may have differential associations with stress and anxiety responses: whereas acceptance-related facets were associated with lower subjective and cortisol stress reactivity, monitoring-related facets such as observing were associated with higher cortisol reactivity and anxious arousal [45,46]. This result may also complement previous mediation analyses of the MINDxYOU trial [12] that identified resilience and the describing and nonreactivity facets of mindfulness as potential mechanisms of change. Existing evidence suggests that mindfulness-related processes may operate differently depending on the clinical outcome, context, and level of practice, with some studies identifying emotional regulation processes as potential mechanisms linking mindfulness to anxiety and depression, while others show that increases in mindfulness facets may mediate the association between formal practice and improvements in stress, symptoms, and well-being [47,48]. However, evidence regarding the moderating role of the observing facet remains limited, and the present findings should therefore be considered preliminary. Future studies are needed to clarify whether baseline mindful observation contributes to differential responses in digital mindfulness-based interventions.

The digital format of the intervention may facilitate the implementation of personalization strategies in real-world health care settings. These findings are consistent with the idea that mindfulness-based and third-wave digital interventions may not be uniformly effective across all users and that their impact may vary according to psychological and contextual factors. From an applied perspective, this knowledge may inform the development of more personalized digital mental health programs, in which the level of guidance, support, or sequencing of content could be adapted to the user's baseline profile [8]. Future studies should further investigate dynamic personalization mechanisms, integrating digital indicators of progress and hybrid clinical support to maximize the effectiveness of such interventions and enhance their real-world applicability in high-demand health care environments.

Limitations

This study has several limitations. Although the stepped-wedge design and the use of mixed-effects models allow partial control of time-related effects, the influence of unmeasured variables cannot be ruled out. In addition, the sample consisted predominantly of women health care professionals working in the public sector, which may limit the generalizability of the findings to other professional or health care contexts. Furthermore, the use of self-report measures may introduce potential biases, such as social desirability or recall bias. Future studies should incorporate physiological indicators and objective measures of program use to provide a more comprehensive assessment of intervention effects. Additionally, the relatively large number of moderators that were examined increases the possibility of a type I error. Therefore, the findings should be interpreted cautiously and considered exploratory and hypothesis-generating until replicated in future confirmatory studies specifically designed to evaluate moderation effects. Finally, interactions between moderators and patterns of digital adherence were not examined. These aspects could be explored in future research using adaptive and personalized intervention models, which may help optimize the personalization and effectiveness of digital mental health interventions for health care professionals.

Conclusions

The present study provides preliminary evidence that baseline psychological and occupational characteristics may influence responses to the MINDxYOU intervention among health care professionals. Higher depressive symptom scores, nursing occupation, and greater mindful observation were associated with larger reductions in perceived stress. These exploratory findings support the importance of considering individual characteristics when developing personalized digital mental health interventions.

XSL•FO
RenderX

Funding

This study was supported by the Instituto de Salud Carlos III (ISCIII), Spanish Ministry of Science and Innovation, through grants PI21/01328 and PI21/01338, cofunded by the European Regional Development Fund (ERDF; "A Way of Making Europe") and by the European Union. GG-P is supported by a postdoctoral junior researcher contract (Exp. RHJ-0043-2024) funded by the Consejería de Salud y Consumo of the Junta de Andalucía, cofunded by the European Social Fund Plus (ESF+). JG-P is supported by the Consejería de Salud y Consumo of the Junta de Andalucía through the Acción A–Intensification of Research Activity program. VCA, GP-G, and JMG-C are supported by Río Hortega contracts from the Instituto de Salud Carlos III (ISCIII). AP-A is supported by a Sara Borrell postdoctoral contract (CD20/00181) from the Instituto de Salud Carlos III (ISCIII), Spanish Ministry of Science and Innovation. AM-B is supported by a Margarita Salas postdoctoral fellowship at the University of Zaragoza, funded by the NextGenerationEU program, in collaboration with the Instituto de Investigación Sanitaria Aragón. The funders had no role in the design of the study; the collection, analysis, or interpretation of data; the writing of the manuscript; or the decision to publish the results.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Multivariate regression analyses examining the moderating effects of baseline depressive symptoms, nursing occupation, and the observing facet of mindfulness on intervention outcomes.
[DOCX File, 2941 KB - nursing_v9i1e98624_app1.docx]

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 1182 KB - nursing_v9i1e98624_app2.pdf]

References

1. Hall LH, Johnson J, Watt I, Tsipa A, O'Connor DB. Healthcare staff wellbeing, burnout, and patient safety: a systematic review. PLoS ONE 2016;11(7):e0159015. [doi: 10.1371/journal.pone.0159015] [Medline: 27391946]
2. West CP, Dyrbye LN, Shanafelt TD. Physician burnout: contributors, consequences and solutions. J Intern Med 2018 Jun;283(6):516-529. [doi: 10.1111/joim.12752] [Medline: 29505159]
3. Carmassi C, Foghi C, Dell'Oste V, et al. PTSD symptoms in healthcare workers facing the three coronavirus outbreaks: What can we expect after the COVID-19 pandemic. Psychiatry Res 2020 Oct;292:113312. [doi: 10.1016/j.psychres.2020.113312] [Medline: 32717711]
4. Shah MK, Gandrakota N, Cimiotti JP, Ghose N, Moore M, Ali MK. Prevalence of and factors associated with nurse burnout in the US. JAMA Netw Open 2021 Feb 1;4(2):e2036469. [doi: 10.1001/jamanetworkopen.2020.36469] [Medline: 33538823]
5. Ruotsalainen JH, Verbeek JH, Mariné A, Serra C. Preventing occupational stress in healthcare workers. Cochrane Database Syst Rev 2015 Apr 7;2015(4):CD002892. [doi: 10.1002/14651858.CD002892.pub5] [Medline: 25847433]
6. Kriakous SA, Elliott KA, Lamers C, Owen R. The effectiveness of mindfulness-based stress reduction on the psychological functioning of healthcare professionals: a systematic review. Mindfulness (N Y) 2021;12(1):1-28. [doi: 10.1007/s12671-020-01500-9] [Medline: 32989406]
7. Spijkerman MPJ, Pots WTM, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: A review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114. [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
8. Andersson G, Titov N, Dear BF, Rozental A, Carlbring P. Internet-delivered psychological treatments: from innovation to implementation. World Psychiatry 2019 Feb;18(1):20-28. [doi: 10.1002/wps.20610] [Medline: 30600624]
9. Cameron G, Mulvenna M, Ennis E, et al. Effectiveness of digital mental health interventions in the workplace: umbrella review of systematic reviews. JMIR Ment Health 2025 Jan 24;12:e67785. [doi: 10.2196/67785] [Medline: 39854722]
10. Aye LM, Tan MM, Schaefer A, et al. Self-help digital mental health intervention in improving burnout and mental health outcomes among healthcare workers: a narrative review. Digit Health 2024 Sep 9;10:20552076241278313. [doi: 10.1177/20552076241278313] [Medline: 39257871]
11. Larsson M, Ho DM, Kirschner M, Seifritz E, Manoliu A. Digital resilience interventions for healthcare workers: a systematic review. Front Psychiatry 2025 Sep 10;16:1519670. [doi: 10.3389/fpsyt.2025.1519670] [Medline: 41000340]
12. López-Del-Hoyo Y, Fernández-Martínez S, Perez-Aranda A, et al. Effectiveness of a web-based self-guided intervention (MINDxYOU) for reducing stress and promoting mental health among health professionals: results from a stepped-wedge cluster randomized trial. J Med Internet Res 2025 Feb 3;27:e59653. [doi: 10.2196/59653] [Medline: 39899345]

XSL•FO
RenderX

13. Taylor H, Cavanagh K, Field AP, Strauss C. Health care workers' need for Headspace: findings from a multisite definitive randomized controlled trial of an unguided digital mindfulness-based self-help app to reduce healthcare worker stress. JMIR Mhealth Uhealth 2022 Aug 25;10(8):e31744. [doi: 10.2196/31744] [Medline: 36006668]
14. Schwartz-Mette RA, Sapp M, Strout K. Online and in-person mindfulness-based stress reduction (MBSR) improves stress resilience in nursing education. Worldviews Evid Based Nurs 2025 Jun;22(3):e70031. [doi: 10.1111/wvn.70031] [Medline: 40350560]
15. Guerrero-Pertiñez G, Carbonell-Aranda V, Pérez-Guerrero G, et al. Resilience as a mediator in a web-based intervention (MINDxYOU) to reduce stress among health care professionals: stepped-wedge cluster randomized trial. JMIR Ment Health 2026 Feb 9;13:e82905. [doi: 10.2196/82905] [Medline: 41662682]
16. Yang JM, Ye H, Long Y, et al. Effects of web-based mindfulness-based interventions on anxiety, depression, and stress among frontline health care workers during the COVID-19 pandemic: systematic review and meta-analysis. J Med Internet Res 2023 Aug 29;25:e44000. [doi: 10.2196/44000] [Medline: 37527546]
17. Wanniarachchi VU, Greenhalgh C, Choi A, Warren JR. Personalization variables in digital mental health interventions for depression and anxiety in adolescents and youth: a scoping review. Front Digit Health 2025 May 15;7:1500220. [doi: 10.3389/fdgth.2025.1500220] [Medline: 40444184]
18. Hyland K, Romero D, Andreasson S, Hammarberg A, Hedman-Lagerlöf E, Johansson M. Individual differences in treatment effects of internet-based cognitive behavioral therapy in primary care: a moderation analysis of a randomized clinical trial. Addict Sci Clin Pract 2025 Feb 14;20(1):17. [doi: 10.1186/s13722-025-00546-1] [Medline: 39953626]
19. Graham AK, Lattie EG, Powell BJ, et al. Implementation strategies for digital mental health interventions in health care settings. Am Psychol 2020 Nov;75(8):1080-1092. [doi: 10.1037/amp0000686] [Medline: 33252946]
20. López-Del-Hoyo Y, Fernández-Martínez S, Pérez-Aranda A, et al. Effectiveness and implementation of an online intervention (MINDxYOU) for reducing stress and promote mental health among healthcare workers in Spain: a study protocol for a stepped-wedge cluster randomized trial. BMC Nurs 2022 Nov 10;21(1):308. [doi: 10.1186/s12912-022-01089-5] [Medline: 36357881]
21. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation). EUR-Lex. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj [accessed 2026-07-28]
22. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404] [Medline: 6668417]
23. Remor E. Psychometric properties of a European Spanish version of the Perceived Stress Scale (PSS). Span J Psychol 2006 May;9(1):86-93. [doi: 10.1017/s1138741600006004] [Medline: 16673626]
24. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
25. Diez-Quevedo C, Rangil T, Sanchez-Planell L, Kroenke K, Spitzer RL. Validation and utility of the patient health questionnaire in diagnosing mental disorders in 1003 general hospital Spanish inpatients. Psychosom Med 2001;63(4):679-686. [doi: 10.1097/00006842-200107000-00021] [Medline: 11485122]
26. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
27. García-Campayo J, Zamorano E, Ruiz MA, et al. Cultural adaptation into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. Health Qual Life Outcomes 2010 Jan 20;8:8. [doi: 10.1186/1477-7525-8-8] [Medline: 20089179]
28. Miaja Ávila M, De la Rubia JM. Validación del Inventario Breve de Síntomas (BSI-18) en mujeres mexicanas diagnosticadas con cáncer de mama [Article in Spanish]. Psicol 2017;14(2-3):307-324. [doi: 10.5209/PSIC.57088]
29. Derogatis L. Brief Symptom Inventory 18: Administration, Scoring and Procedures Manual: NCS Pearson Inc; 2001. URL: https://www.pearsonassessments.com/en-us/Store/Professional-Assessments/Personality-%26-Biopsychosocial/Brief-Symptom-Inventory-18/p/100000638 [accessed 2026-07-27]
30. Baer RA, Smith GT, Lykins E, et al. Construct validity of the five facet mindfulness questionnaire in meditating and nonmeditating samples. Assessment 2008 Sep;15(3):329-342. [doi: 10.1177/1073191107313003] [Medline: 18310597]
31. Cebolla A, García-Palacios A, Soler J, Guillen V, Baños R, Botella C. Psychometric properties of the Spanish validation of the Five Facets of Mindfulness Questionnaire (FFMQ). Eur J Psychiat 2012 Jun;26(2):118-126. [doi: 10.4321/S0213-61632012000200005]
32. Connor KM, Davidson JRT. Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). Depress Anxiety 2003;18(2):76-82. [doi: 10.1002/da.10113] [Medline: 12964174]
33. Soler Sánchez MI, Meseguer de Pedro M, García Izquierdo M. Propiedades psicométricas de la versión española de la escala de resiliencia de 10 ítems de Connor-Davidson (CD-RISC 10) en una muestra multiocupacional [Article in Spanish]. Rev Latinoam Psicol 2016 Sep;48(3):159-166. [doi: 10.1016/j.rlp.2015.09.002]
34. Gu J, Baer R, Cavanagh K, Kuyken W, Strauss C. Development and psychometric properties of the Sussex-Oxford Compassion Scales (SOCS). Assessment 2020 Jan;27(1):3-20. [doi: 10.1177/1073191119860911] [Medline: 31353931]

XSL•FO
RenderX

35. García-Campayo J, Barceló-Soler A, Martínez-Rubio D, et al. Exploring the relationship between self-compassion and compassion for others: the role of psychological distress and wellbeing. Assessment 2024 Jul;31(5):1038-1051. [doi: 10.1177/10731911231203966] [Medline: 37840255]
36. Bond FW, Hayes SC, Baer RA, et al. Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: a revised measure of psychological inflexibility and experiential avoidance. Behav Ther 2011 Dec;42(4):676-688. [doi: 10.1016/j.beth.2011.03.007] [Medline: 22035996]
37. Ruiz FJ, Langer Herrera AI, Luciano C, Cangas AJ, Beltrán I. Measuring experiential avoidance and psychological inflexibility: the Spanish version of the Acceptance and Action Questionnaire-II. Psicothema 2013 Feb;25(1):123-129. [doi: 10.7334/psicothema2011.239] [Medline: 23336554]
38. Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. Arch Gen Psychiatry 2002 Oct;59(10):877-883. [doi: 10.1001/archpsyc.59.10.877] [Medline: 12365874]
39. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw 2015 Oct 7;67(1):1-48 [FREE Full text] [doi: 10.18637/jss.v067.i01]
40. Nissen ER, Zachariae R, O'Connor M, et al. Internet-delivered mindfulness-based cognitive therapy for anxiety and depression in cancer survivors: predictors of treatment response. Internet Interv 2021 Mar;23:100365. [doi: 10.1016/j.invent.2021.100365] [Medline: 33552930]
41. Hammen C. Stress and depression. Annu Rev Clin Psychol 2005;1:293-319. [doi: 10.1146/annurev.clinpsy.1.102803.143938] [Medline: 17716090]
42. Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: model, processes and outcomes. Behav Res Ther 2006 Jan;44(1):1-25. [doi: 10.1016/j.brat.2005.06.006] [Medline: 16300724]
43. Kuyken W, Warren FC, Taylor RS, et al. Efficacy of mindfulness-based cognitive therapy in prevention of depressive relapse: an individual patient data meta-analysis from randomized trials. JAMA Psychiatry 2016 Jun 1;73(6):565-574. [doi: 10.1001/jamapsychiatry.2016.0076] [Medline: 27119968]
44. Gu J, Strauss C, Bond R, Cavanagh K. How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. Clin Psychol Rev 2015 Apr;37:1-12. [doi: 10.1016/j.cpr.2015.01.006] [Medline: 25689576]
45. Gallistl M, Linz R, Puhlmann LMC, Singer T, Engert V. Evidence for differential associations of distinct trait mindfulness facets with acute and chronic stress. Psychoneuroendocrinology 2024 Aug;166:107051. [doi: 10.1016/j.psyneuen.2024.107051] [Medline: 38678734]
46. Desrosiers A, Klemanski DH, Nolen-Hoeksema S. Mapping mindfulness facets onto dimensions of anxiety and depression. Behav Ther 2013 Sep;44(3):373-384. [doi: 10.1016/j.beth.2013.02.001] [Medline: 23768665]
47. Carmody J, Baer RA. Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. J Behav Med 2008 Feb;31(1):23-33. [doi: 10.1007/s10865-007-9130-7] [Medline: 17899351]
48. Desrosiers A, Vine V, Klemanski DH, Nolen-Hoeksema S. Mindfulness and emotion regulation in depression and anxiety: common and distinct mechanisms of action. Depress Anxiety 2013 Jul;30(7):654-661. [doi: 10.1002/da.22124] [Medline: 23592556]

Abbreviations

CD-RISC: Connor-Davidson Resilience Scale
FFMQ-15: 15-item Five Facet Mindfulness Questionnaire
PHQ-9: 9-item Patient Health Questionnaire
PSS-10: 10-item Perceived Stress Scale

Edited by UM Kinnunen; submitted 17.Apr.2026; peer-reviewed by M García-Toto, W Collinge; revised version received 11.Jun.2026; accepted 22.Jun.2026; published 14.Aug.2026.

Please cite as:
Guerrero-Pertiñez G, Carbonell Aranda V, Dawood-Histrova JJ, Pérez-Guerrero G, Pérez-Aranda A, Fernández-Martínez S, Monreal-Bartolomé A, Barceló-Soler A, García-Campayo J, López del Hoyo Y, Herrera-Imbroda J, Goodman-Casanova JM, Guzman-Parra J
Web-Based Mindfulness Intervention for Reducing Perceived Stress in Health Care Professionals: Secondary Analysis of Moderators in a Stepped-Wedge Cluster Randomized Trial
JMIR Nursing 2026;9:e98624
URL: https://nursing.jmir.org/2026/1/e98624
doi:10.2196/98624

XSL•FO
RenderX

© Gloria Guerrero-Pertiñez, Vera Carbonell Aranda, Jonathan Joseph Dawood-Histrova, Gloria Pérez-Guerrero, Adrián Pérez-Aranda, Selene Fernández-Martínez, Alicia Monreal-Bartolomé, Alberto Barceló-Soler, Javier García-Campayo, Yolanda López del Hoyo, Jesús Herrera-Imbroda, Jessica Marian Goodman-Casanova, Jose Guzman-Parra. Originally published in JMIR Nursing (https://nursing.jmir.org), 14.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Inpatient Nurses' Anticipatory Perspectives on Assistive Robots: Qualitative Study

Mavis Weiting Tan*, BSN (Hons); Timie Lai Teng Chan*, BSN (Hons); Charmaine Shi Min Taye*, BSN, MscEd, PhD

Department of Nursing, Woodlands Hospital, 17 Woodlands Drive,, Singapore, Singapore

*all authors contributed equally

Corresponding Author:
Mavis Weiting Tan, BSN (Hons)
Department of Nursing, Woodlands Hospital, 17 Woodlands Drive,, Singapore, Singapore

Abstract

Background: Assistive robots have been proposed to support nursing work by offloading routine, logistical, and physically demanding tasks. However, existing research has largely focused on postimplementation acceptance or usability, with fewer studies examining how nurses conceptualize assistive robots before routine deployment, particularly in settings where exposure remains limited.

Objective: This study aimed to explore inpatient nurses' anticipatory perspectives on assistive robots and to examine how perceived benefits, concerns, and implementation conditions are constructed at the preimplementation stage.

Methods: A qualitative descriptive study was conducted using semistructured, face-to-face interviews with 16 registered nurses aged 21 to 32 years from 4 general medical wards in a public tertiary hospital in Singapore. At the time of data collection, assistive robots had not been routinely implemented; participants' exposure was limited to pilots, demonstrations, and indirect sources. Interview transcripts were analyzed using reflexive thematic analysis.

Results: Participants' accounts reflected anticipatory acceptance characterized by cautious optimism regarding the potential of assistive robots to support nursing work, alongside concurrent concerns about safety, professional identity, relational care, and implementation feasibility. Four themes were developed: anticipatory optimism shaped by indirect exposure; negotiating uncertainty, safety, and professional boundaries; imagining assistive robots as workload and workflow support; and conditions for responsible implementation. Concerns centered on professional identity, patient safety, ethical boundaries, loss of human touch, technical reliability, cost-effectiveness, and organizational feasibility. Participants viewed assistive robots as acceptable only if they were integrated into nursing workflows, preserved clinical judgment and relational care, and were supported by training, technical infrastructure, leadership commitment, and clear governance.

Conclusions: Nurses' anticipatory perspectives suggest that acceptance of assistive robots depends not only on perceived usefulness, but also on alignment with professional values, patient safety, relational care, and organizational readiness. Early nurse involvement, human-centered implementation, structured training, and clear governance are needed to ensure that assistive robots augment rather than undermine nursing practice.

(JMIR Nursing 2026;9:e91106) doi:10.2196/91106

KEYWORDS

nurses; assistive robots; human-robot interaction; preimplementation; technology adoption

Introduction

Assistive Robots and Their Emergence in Health Care

Assistive robots are increasingly positioned within the broader category of service assistive robots. Service assistive robots are commonly defined as assistive robots for personal or professional use that perform useful tasks for humans or equipment, excluding industrial automation applications [1]. In health care, assistive robots have been applied in multiple areas, including rehabilitation, patient mobility support, logistics, medication or supply delivery, socially assistive companionship, monitoring, and patient-facing support functions [2-6].

Interest in assistive robots has grown in response to workforce shortages, aging populations, infection-control demands, and pressure to improve productivity in health care systems [7-9]. Developments in sensor technology, AI, autonomous navigation, and human-robot interaction have expanded the potential functions of assistive robots beyond fixed repetitive tasks toward more adaptive forms of clinical and social support [2-4]. These capabilities not only make assistive robots potentially useful in care environments but also raise questions about safety,

XSL•FO
RenderX

accountability, workflow integration, and the boundaries of human-centered care [4,10,11].

During the COVID-19 pandemic, socially assistive robots were deployed in some care settings to support symptom screening and exposure assessment, illustrating one way assistive robots were used to reduce infection-related contact while maintaining service delivery [12,13]. However, pandemic-related deployments also highlighted that assistive robots must be evaluated not only for technical feasibility but also for acceptability, workflow fit, and their effects on human interaction [12,13].

Within nursing contexts, the term "assistive robots" refers to robots designed to support care through routine, physical, logistical, automated, or interactive functions. These include assistive robots and automated devices used in nurses' work, physically assistive technologies, and socially assistive robots intended to support interaction, engagement, or care delivery [4-6,10]. By reducing physical strain and repetitive workloads, assistive robots may help support nursing work and preserve time for patient-centered care [4,5,10]. However, their introduction into nursing practice is not merely a technical intervention. It may reshape task ownership, professional responsibility, nurse-patient interaction, and perceptions of what constitutes appropriate care [4,10,11].

Technology Acceptance and Preimplementation Perspectives

Technology acceptance research provides a useful starting point for understanding how health professionals respond to emerging technologies. The Technology Acceptance Model (TAM) and its extensions, including the Unified Theory of Acceptance and Use of Technology (UTAUT), emphasize constructs such as perceived usefulness, perceived ease of use, social influence, facilitating conditions, and intention to use [14,15]. In robot acceptance research, the Almere Model further adapts technology acceptance concepts for assistive and socially assistive robots by incorporating constructs such as trust, anxiety, social presence, perceived enjoyment, and perceived sociability [16].

However, acceptance is not formed solely through direct use. In preimplementation contexts, health professionals may construct expectations through demonstrations, workplace discourse, media narratives, peer accounts, or limited observational exposure. These anticipatory judgments can influence whether a technology is perceived as useful, safe, legitimate, or compatible with professional values before it becomes part of routine practice [15-18].

In this study, these anticipatory judgments are conceptualized as *anticipatory acceptance*, referring to nurses' prospective evaluation of whether assistive robots are likely to be useful, safe, appropriate, and compatible with nursing practice before routine implementation occurs. Unlike postimplementation technology acceptance, which develops through direct interaction and accumulated user experience, anticipatory acceptance is informed by indirect exposure and expectations about future clinical practice. It is also distinct from organizational readiness, which reflects preparedness for implementation, and from general attitudes toward technology, which do not necessarily consider how an innovation may be integrated into everyday nursing work. By focusing on anticipatory acceptance, this study seeks to understand how nurses begin to negotiate the potential role of assistive robots in nursing practice before assistive robots become part of routine care.

In nursing, the preimplementation stage is especially important because assistive robots enter a practice environment where clinical judgment, embodied care, emotional labor, accountability, and workflow coordination are central [4,10,11]. Nurses' early interpretations may therefore shape not only their willingness to use assistive robots but also how they define appropriate task delegation, human oversight, and acceptable limits of automation.

Nurses' Acceptance of Assistive Robots

Existing studies suggest that nurses and care professionals often hold cautiously positive attitudes toward assistive robots [6,13,17-20]. Positive perceptions are commonly associated with the potential to reduce physical strain, support routine or repetitive tasks, improve workflow efficiency, and alleviate aspects of the nursing workload [4-6,10]. At the same time, concerns persist regarding technical reliability, patient safety, the loss of human touch, privacy, ethical accountability, costs, and potential effects on employment or professional identity [6,10,11,13,17-20].

Acceptance also appears to be shaped by task type and perceived fit with professional values. Assistive robots tend to be viewed as more acceptable for logistical, repetitive, or physically demanding tasks than for emotionally sensitive, relational, or judgment-intensive aspects of care [5,6,13,17,20]. This distinction is important in nursing because care work is not only defined by task completion but also by surveillance, interpretation, communication, and relational presence [10,11].

The work by Turja et al [17,18] is particularly relevant because it moves beyond general technology acceptance to examine care workers' readiness for robotization and their intention to use assistive robots in professional contexts. Their Robot Acceptance Model for Care (RAM-care) highlights that robot acceptance is shaped not only by perceived usefulness but also by social influence, attitudes, and whether robot use is perceived as consistent with personal and professional values [17]. This suggests that acceptance of assistive robots in nursing cannot be understood as an individual preference alone; it is also shaped by organizational expectations, task legitimacy, perceived value alignment, and the wider social meaning of robotized care [17,18].

Despite this growing literature, important gaps remain. Much of the existing evidence is quantitative, cross-sectional, or based on attitude and acceptance measures [17-20]. These studies are useful for identifying predictors of acceptance, but they provide limited insight into how nurses interpret, negotiate, and anticipate the implications of assistive robots before routine implementation. Existing qualitative work has begun to examine nurses' experiences and perceptions of socially assistive technologies, including during COVID-19-related deployments

XSL•FO
RenderX

[13], but fewer studies have explored anticipatory perspectives in acute inpatient settings where assistive robots remain emergent rather than embedded. This limits understanding of how nurses anticipate the ethical, professional, operational, and economic implications of robotic systems before they become part of everyday clinical workflows.

Conceptual and Ethical Considerations in Nursing Practice

The integration of assistive robots into nursing practice raises conceptual and ethical questions about professional identity, task ownership, accountability, and the boundaries of relational care [4,10,11]. Although these technologies are often framed as supportive tools, their introduction may reshape how nursing work is organized, experienced, and valued.

Existing literature reflects these tensions. Some nurses and care professionals view assistive robots positively, emphasizing their potential to alleviate workload and support care delivery, whereas others express concerns about technical reliability, reduced human interactions, ethical trade-offs, and accountability when care is mediated by machines [10,11,13,16,19]. Practical barriers, including high acquisition and maintenance costs, technical complexity, maintenance requirements, and limited institutional support infrastructure, may further complicate adoption [8,10]. Hesitancy may also reflect anxieties about job displacement or perceived misalignment between robotic functions and core nursing values [6,10,11,17,18].

Local Context and Study Aim

Nurses play a central role in inpatient care delivery and are directly affected by how assistive robots may reshape workflows, task boundaries, and clinical responsibility. At the same time, the nursing workforce faces sustained pressure from staff turnover, aging populations with complex needs, and increasing care demands [7,8]. Assistive robots and other automated technologies have been proposed as potential tools to support nursing work and care delivery [4,5,10,21], yet successful integration depends heavily on nurses' engagement, trust, and sense-making [6,13,17-19].

As primary users and workflow mediators, nurses' involvement in the design, implementation, and evaluation of assistive robots is critical, as they hold firsthand knowledge of clinical demands, patient needs, workflow constraints, and contextual risks [4,5]. Prior research suggests that attitudes toward assistive robots are shaped by familiarity, perceived usefulness, perceived ease of use, and prior or indirect exposure [6,16-20]. However, much of the existing evidence relies on quantitative instruments such as robot attitude or acceptance scales [17-20]. These approaches are useful for measuring acceptance-related constructs but provide limited insight into how nurses interpret, negotiate, and anticipate the implications of assistive robots in everyday inpatient practice.

In the study hospital, assistive robots had been piloted on a limited basis, but were not routinely implemented in inpatient nursing care. Local exposure was primarily limited to demonstrations, institutional briefings, and small-scale pilots involving nonclinical robotic functions, such as logistics support (eg, supply delivery) and task-support activities, such as providing on-demand audio-visual support for patient orientation and educational activities. This created an opportunity to examine how nurses formed anticipatory perspectives in the absence of sustained hands-on experience with assistive robots. This study, therefore, aimed to explore how inpatient nurses conceptualized, interpreted, and anticipated the use of assistive robots in clinical practice, including their perceived benefits, concerns, ethical boundaries, and conditions influencing future adoption.

Methods

Study Design

This study used a qualitative descriptive design to explore inpatient nurses' anticipatory perspectives on assistive robots in clinical care. This design was appropriate because the study sought to generate a practice-near understanding of their expectations, concerns, and perceived implementation conditions in a context where routine exposure to assistive robots was limited. Interview data were analyzed using reflexive thematic analysis.

Participants and Setting

A convenience sampling strategy was used to recruit registered nurses from 4 general medical inpatient wards in a public tertiary hospital in Singapore. These wards were selected because they were the clinical areas where small-scale pilots, demonstrations, and informal exposure to assistive robots had occurred. At the time of data collection, assistive robots had not been routinely implemented in inpatient nursing care at the study site. Nurses' exposure was therefore largely indirect and was limited primarily to small-scale pilots and demonstrations involving nonclinical robotic functions, particularly logistics and task-support activities, together with institutional briefings, educational videos, and informal discussions among colleagues. None of these robotic systems were routinely integrated into bedside nursing practice, and participants did not report sustained hands-on experience using assistive robots during their daily clinical work. Consequently, the study explored anticipatory perspectives informed by limited local exposure rather than reflections based on routine clinical use.

Eligible participants were registered nurses working in the selected wards who had direct experience with inpatient nursing workflows. Potential participants were identified through ward rosters and approached by members of the research team during nonclinical hours. Convenience sampling was used because the pool of nurses with contextual exposure to assistive-robot pilots was limited.

The possibility of sampling bias cannot be excluded. Nurses who participated in the interviews could have held stronger preconceptions about assistive robots compared to those who declined. In addition, the sample comprised mainly early-career nurses, which may have shaped the concerns and expectations, and professional identity negotiations reflected in the findings. Nonetheless, the findings may be transferable to similar acute inpatient settings where assistive robots are at an early or anticipatory stage of implementation.

XSL•FO
RenderX

Researcher Positionality and Reflexivity

The research team comprised registered nurses with at least 5 years of clinical experience in acute inpatient care and prior experience conducting qualitative research. None of the team members had direct hands-on responsibility for implementing assistive robots in clinical practice. This positioned the team as clinically informed about inpatient nursing work but not directly invested in any specific implementation outcome.

Interviews were conducted by 2 members of the research team, TLTC and MWT, both registered nurses who did not hold supervisory, managerial, or appraisal responsibilities for the participants. This helped reduce potential power imbalances during recruitment and interviews.

The researchers recognized that their nursing backgrounds could sensitize them to issues such as workload, patient safety, professional accountability, and relational care, while also potentially shaping how participants' accounts were interpreted. At the same time, the shared professional background between interviewers and participants was considered likely to facilitate rapport and enable discussion of workplace realities using familiar clinical language. The researchers remained mindful; however, that this insider position could also encourage implicit assumptions about nursing practice or shared understandings that might otherwise have been explored more explicitly during interviews. Interviewers therefore used follow-up and clarifying questions to encourage participants to elaborate on their perspectives rather than assuming shared meanings.

To support reflexivity, the team held regular analytic discussions throughout data collection and analysis. These discussions examined assumptions, considered alternative interpretations, and avoided framing participants' views as either simple resistance to innovation or uncritical acceptance of technology.

Procedures

Potential participants were approached in-person by members of the research team during nonclinical hours and provided with an information sheet detailing the study aims and procedures. They were given time to consider their participation. Of the 20 nurses approached, 16 consented to participate, while 4 declined due to scheduling constraints or lack of interest. Written informed consent was obtained from all participants.

A pilot interview was conducted with one registered nurse from a general medical inpatient ward within the same hospital to assess the clarity, flow, and relevance of the interview questions. The pilot also provided the interviewers with an opportunity to familiarize themselves with the interview process and refine their interviewing approach. No substantive revisions to the interview guide were required following the pilot interview, and the same interview guide was subsequently used for all participants. Consistent with qualitative research practice, the pilot interview was not included in the final analysis.

Semistructured, face-to-face interviews were conducted with the 16 participants between October 2023 and June 2024. Interviews took place in a private room within the hospital to ensure confidentiality and minimize interruptions. Both interviewers received the same training to ensure consistency in interview conduct and reduce variation in interview style.

At the start of each interview, participants were provided with a brief description of assistive robots to establish a common frame of reference. Examples included nursing-care assistive robots (eg, medication delivery and patient transport), physically assistive robots (eg, patient lifting and mobility support), and socially assistive robots (eg, companionship and cognitive support). These examples were presented solely to ensure a consistent understanding of the terminology and did not imply that such assistive robots were routinely available at the study hospital. Participants were subsequently invited to describe what they understood by the term "assistive robots" and to reflect on how such technologies might influence nursing practice based on their own observations, indirect exposure, or personal interpretations.

The interview guide explored participants' perceptions of assistive robots, including perceived benefits, concerns, ethical considerations, and expectations for implementation. Open-ended questions were used, and participants were encouraged to provide concrete examples. Interviews lasted between 60 and 90 minutes. All interviews were audio-recorded with participants' permission and transcribed verbatim. Identifying information was removed from transcripts to protect participant anonymity.

Data Adequacy

Recruitment and data collection occurred concurrently with preliminary analysis. Data adequacy was assessed iteratively during data collection, with attention to code saturation and thematic sufficiency rather than to exhaustive meaning saturation. After approximately 14 interviews, no substantially new codes were identified across the main topic areas, and participants' accounts showed repeated patterns regarding perceived benefits, safety concerns, professional boundaries, and implementation conditions. Two additional interviews were conducted to assess the stability of the developing thematic structure. The final sample of 16 participants was judged sufficient for the study aim, design, and focused clinical context, while recognizing that findings remain contextually situated rather than statistically generalizable.

Data Analysis

Interview transcripts were analyzed using reflexive thematic analysis as described by Braun and Clarke [22]. The analysis involved iterative phases of familiarization, initial coding, theme development, reviewing, and defining and naming themes.

Consistent with the qualitative descriptive design, the analysis aimed to provide a clinically grounded account of nurses' perspectives while recognizing that theme development involved active interpretation by the researchers. Themes, therefore, moved beyond simple topic summaries to capture patterns of shared meaning across participants' accounts, while remaining closely anchored in their descriptions of anticipated experiences with assistive robots.

Initial coding was conducted by 1 author (TCLT) to generate codes that captured patterns of meaning relevant to the study

XSL•FO
RenderX

aims. Coding was conducted primarily inductively, while remaining sensitized to concepts identified in prior literature, including perceived usefulness, workflow fit, patient safety, ethical concerns, and professional boundaries. Codes were revised as the analysis progressed and as new interpretations emerged from repeated engagement with the data.

Theme development was interpretive and reflexive rather than based on consensus coding. Provisional themes were generated through repeated engagement with coded data extracts, the full dataset, and emerging analytic interpretations. Regular analytic discussions were held among the research team to interrogate assumptions, explore alternative explanations, and refine the coherence and distinctiveness of themes. These discussions were used to support reflexive interpretation rather than to achieve intercoder agreement.

Themes were finalized based on coherence, distinctiveness, and explanatory relevance to the research questions. An audit trail documenting coding decisions, theme development, and analytic refinements was maintained throughout the process. Data were managed using Microsoft Excel to support systematic coding and documentation.

The study is reported in accordance with the COREQ (Consolidated Criteria for Reporting Qualitative Research) Checklist 1 [23].

Ethical Considerations

The study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. Ethical approval was obtained from the National Healthcare Group Domain Specific Review Board (reference number: 2023/00679). Written informed consent was obtained from all participants prior to the interviews.

Participants were informed that participation was voluntary and that refusal or withdrawal would not result in any adverse consequences, including effects on employment status or performance appraisals. Participants were advised that they had the right to decline to answer any interview questions and to withdraw from the study at any time prior to data analysis.

No reimbursement or incentives were provided for study participation. Participant confidentiality and anonymity were ensured throughout the study. Audio recordings and interview transcripts were stored securely, with access limited to members of the research team. All personally identifiable information was removed during transcription and data analysis. Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans for this research.

Results

Participant Characteristics and Overview of Themes

Demographic characteristics of the 16 participants are presented in Table 1. No participants withdrew after providing informed consent. Participants ranged in age from 21 to 32 years, with a mean age of 26 (SD 2.94) years. The majority of participants were Singaporean citizens or permanent residents (n=11). Five participants were foreign nurses: 4 were from Malaysia and 1 was from Myanmar. The mean duration of professional nursing experience was 34 (SD 2.05) months.

XSL•FO
RenderX

Table . Characteristics of study participants (N=16).

Characteristic	Value, n (%)
Sex	
Female	16 (100)
Male	0 (0)
Age range (y)	
20 - 25	7 (43.8)
26 - 30	7 (43.8)
31 - 35	2 (12.5)
Designation	
Staff nurse	10 (62.5)
Senior staff nurse	6 (37.5)
Professional nursing experience (y)	
<1	4 (25)
1 - 5	11 (68.8)
6 - 10	1 (6.3)
Highest nursing education level	
Polytechnic diploma	8 (50)
University degree	8 (50)
Nationality or residency status	
Singapore citizen or permanent resident	11 (68.8)
Foreign nurse: Malaysia	4 (25)
Foreign nurse: Myanmar	1 (6.3)

Four themes were developed from the analysis: (1) anticipatory optimism shaped by indirect exposure; (2) negotiating uncertainty, safety, and professional boundaries; (3) imagining assistive robots as workload and workflow support; and (4) conditions for responsible implementation. Table 2 summarizes the relationship between initial codes, subthemes, and overarching themes. The thematic structure illustrates that participants' perspectives were not formed through routine use of assistive robots, but through anticipatory sense-making shaped by limited exposure, imagined use, perceived usefulness, safety concerns, professional boundaries, and implementation conditions.

Table . Development of themes from initial codes to subthemes and overarching themes.

Initial codes or analytic ideas	Subthemes	Overarching themes
Curiosity, demonstrations, media exposure, imagined use, and cautious optimism	Indirect exposure, imagined usefulness, and conditional openness	Anticipatory optimism shaped by indirect exposure
Malfunction, overreliance, patient safety, job displacement, human touch, counseling discomfort, and cost concerns	Safety and reliability, professional role boundaries, relational care limits, and organizational feasibility	Negotiating uncertainty, safety, and professional boundaries
Lifting support, patient turning, transport, basic requests, information retrieval, workload reduction, and staff retention	Physical workload support, routine task support, workflow responsiveness, and organizational benefits	Imagining assistive robots as workload and workflow support
Training, hands-on practice, protected time, phased pilots, technical support, leadership, and champions	Competency development, gradual implementation, support infrastructure, and leadership and buy-in	Conditions for responsible implementation

Table 2 illustrates how participants' anticipatory perspectives on assistive robots were shaped by perceived usefulness, safety concerns, professional boundaries, workload expectations, and implementation conditions.

Theme 1: Anticipatory Optimism Shaped by Indirect Exposure

Participants expressed curiosity and cautious optimism toward assistive robots as potential tools to support nursing work and alleviate routine workload. Although most nurses had limited

XSL•FO
RenderX

direct exposure to assistive robots in health care, they articulated generally positive expectations about their future role in nursing practice, particularly in relation to physically demanding, repetitive, or time-consuming tasks.

Participants' perspectives were shaped largely by anticipatory and indirect exposure rather than sustained hands-on use. They drew on demonstrations, informal observations, media portrayals, peer discussions, and imagined future scenarios to consider how assistive robots might function in clinical practice. These indirect forms of exposure contributed to both interest and caution, suggesting that nurses began forming views about assistive robots before routine implementation. One participant stated, "But hopefully in the future I have a robot to help us to turn the patient, take orders from the patient… and engage the patient as well" (P13).

Some participants envisioned assistive robots as intelligent support tools that could enhance efficiency at the point of care by facilitating rapid access to information:

> *If the robot… can act like a computer, like Siri… then it can show me immediately. Don't need to search.* [P10]

Despite this openness, participants' optimism was tempered by uncertainty. Nurses emphasized the need for evidence of effectiveness, reliability, and sustainability before fully trusting assistive robots in clinical care. Concerns were raised about the overreliance on technology and the potential consequences of system failure in patient care settings:

> *It's 50–50… it would be really helpful if you can rely on it. But sometimes I'm worried that we might over rely on it.* [P14]

Others questioned whether assistive robots would genuinely reduce workload or instead introduce additional responsibilities for nurses, particularly if the technology requires ongoing supervision or troubleshooting. One participant stated, "Will it be a helping hand or extra workload for us?" (P13).

Overall, nurses' accounts reflected openness to innovation alongside a cautious appraisal. Their optimism was not unconditional but grounded in practical concerns about whether assistive robots would be reliable, useful, and compatible with everyday nursing work.

Theme 2: Negotiating Uncertainty, Safety, and Professional Boundaries

Participants articulated several uncertainties about the integration of assistive robots into nursing practice. A prominent concern was related to job security and role displacement. Some nurses worried that assistive robots might encroach on aspects of nursing work, while simultaneously emphasizing that core nursing functions involving human judgment, accountability, and emotional engagement could not be replicated by machines. One participant stated, "I hope we can use it to relieve the workload, but not totally relying on it… otherwise they would take over our role" (P14).

Concerns about technical reliability were also prominent. Participants highlighted the risk of system malfunction and stressed the importance of maintaining human oversight, particularly in situations where patient safety could be affected. One participant stated, "Assistive robots can malfunction… we as humans should know how to override it and take over" (P07).

Several participants cautioned against excessive reliance on assistive robots, emphasizing that patient safety ultimately remains dependent on nurses' situational awareness, monitoring, and ability to intervene when needed. One participant stated, "You still need to check the robot because it is about patient safety" (P01).

Participants consistently raised concerns about the potential loss of human touch in nursing care. Emotional support, physical reassurance, and empathetic communication were described as fundamental aspects of nursing that should not be fully delegated to assistive robots. Importantly, nurses did not frame these concerns as a rejection of assistive robots, but rather as boundary-setting around the relational and emotional dimensions of care. One participant stated, "Touch therapy is one aspect that can't be replaced by assistive robots" (P14).

Some participants expressed discomfort with the idea of assistive robots engaging in emotionally sensitive interactions, such as counseling or responding to patient distress, suggesting limits to the appropriateness of robotic involvement in relational care. One participant stated, "If you have a robot coming to give you counseling, that's a bit weird… they probably need a human" (P06).

Beyond clinical and ethical considerations, participants framed assistive robots in explicitly economic terms. Bedside nurses-raised concerns about cost-effectiveness, maintenance burden, and return on investment, reflecting a pragmatic awareness of organizational resource constraints and sustainability. One participant stated, "If it's very expensive and doing very minimal work, then definitely not sustainable" (P06).

Together, these uncertainties show that participants' concerns were multidimensional. They involved not only whether assistive robots would work technically but also whether they would preserve safety, professional identity, ethical boundaries, organizational feasibility, and relational care.

Theme 3: Imagining Assistive Robots for Workload and Workflow Support

Despite these uncertainties, participants identified several ways in which assistive robots could positively contribute to nursing practice. Many viewed assistive robots as potential tools for increasing nursing capacity by taking over selected routine or physically demanding tasks. This was perceived as a way to reduce physical strain, conserve cognitive resources, and allow nurses to focus on more complex clinical responsibilities. One participant stated, "It can help reserve our brain power for more critical things" (P08).

Participants anticipated benefits for physically demanding activities, such as patient lifting and transport, which they believed would reduce occupational strain and help prevent injury. One participant stated, "If there is a robot that can help lift a heavy patient, it helps prevent back pain" (P09).

XSL•FO
RenderX

Some participants also linked assistive robots to improved responsiveness to patients' basic needs. They suggested that assistive robots could help reduce waiting times for simple requests, such as obtaining water or retrieving items, thereby improving the patient experience while allowing nurses to attend to more urgent clinical priorities. One participant stated, "If the robot can help take simple things like water immediately, then the patient will be happy" (P10).

At the organizational level, participants perceived that the use of assistive robots could reap potential operational benefits, including improved productivity, staff satisfaction, and retention. One participant stated, "The staff workload can be a bit lesser… then you can retain your staff" (P02).

Some nurses also noted that visible adoption of assistive robots could enhance the hospital's public image as innovative and technologically progressive. Importantly, participants consistently framed the benefits of assistive robots as contingent on their integration into nursing practice, emphasizing alignment with nursing workflows, the reduction rather than the addition of workload, preservation of clinical judgment, and the protection of relational care.

Theme 4: Conditions for Responsible Implementation

Participants emphasized that successful integration of assistive robots would depend on several implementation conditions. Structured and ongoing training was viewed as essential for building confidence and competence, with nurses highlighting the importance of hands-on learning, clear guidance, and opportunities to practice using the technology before being expected to incorporate it into routine care. One participant stated, "If I want to use this robot, I would really want to go for a training session first" (P13).

Protected training time during working hours was considered critical to prevent cognitive overload and ensure meaningful engagement. Participants were concerned that implementation would be less effective if training were added on top of the existing workload or delivered without sufficient opportunity for practice.

Participants also emphasized the importance of phased implementation through small-scale pilots. Gradual introduction was viewed as a way for nurses to assess reliability, understand the robot's functions, and build trust over time. One participant stated, "Give the staff some time to get used to how reliable the robot is" (P14).

Adequate technical support and rapid troubleshooting were identified as essential to prevent assistive robots from becoming additional burdens rather than supportive tools. One participant stated, "It's supposed to offload us, not add on" (P08).

Stakeholder buy-in was another key implementation condition. Participants highlighted the role of leadership support, clear communication, feedback mechanisms, and staff champions in encouraging adoption. One participant stated, "You just need a champion… then everybody will hop on to the idea" (P05).

At the same time, nurses acknowledged variability in receptivity to assistive robots, noting differences in comfort with technology among individuals. Participants suggested that tailored training, leadership engagement, and clear expectations may be required to support successful integration. Overall, participants viewed responsible implementation as requiring more than access to the technology itself. It depended on preparation, training, technical infrastructure, leadership support, and governance processes that ensured assistive robots were introduced as supportive tools rather than additional burdens.

Discussion

Principal Findings

This study explored inpatient nurses' anticipatory perspectives on assistive robots in a clinical context where such technologies had been piloted on a limited basis but were not routinely implemented. Participants' accounts reflected anticipatory acceptance characterized by cautious optimism regarding the potential of assistive robots to support nursing work, alongside concurrent concerns about safety, professional identity, relational care, and implementation feasibility.

The findings extend existing literature by demonstrating that nurses' views on assistive robots are formed before routine use through anticipatory sense-making. Participants' perspectives were shaped largely by indirect exposure, including demonstrations, peer discussions, and imagined future use. This preimplementation perspective is important because early interpretations may subsequently influence trust, acceptance, resistance, and implementation trajectories.

Importantly, nurses did not view assistive robots merely as technological devices. Instead, they assessed these technologies in relation to their potential impact on patient safety, professional responsibility, therapeutic relationships, clinical workflows, and the long-term sustainability of health care organizations. Their acceptance of assistive robots was, therefore, contingent upon the extent to which these technologies were perceived as supporting and enhancing, rather than substituting, professional nursing judgment and the relational aspects of care.

Anticipatory Acceptance Before Routine Implementation

Participants' cautious optimism is consistent with previous studies showing that nurses and other care professionals often view assistive robots positively when they are perceived to reduce physical workload, support routine tasks, and improve workflow efficiency [4-6,10]. Concerns relating to technical reliability, reduced human interaction, cost, and role displacement align with prior research on nurses' and care professionals' attitudes toward assistive robots in health and social care settings [6,13,17-20].

This study adds nuance by focusing on the preimplementation stage. TAMs such as the TAM and the UTAUT emphasize perceived usefulness, perceived ease of use, social influence, facilitating conditions, and intention to use as key determinants of acceptance [14,15]. The Almere model further highlights constructs such as trust, anxiety, perceived sociability, and enjoyment [16]. Building on these concepts, our findings suggest that anticipatory acceptance extends beyond conventional

XSL•FO
RenderX

technology assessments of usefulness and usability to include concerns about safety, professional boundaries, relational care, and potential consequences of implementation for everyday work (ie, whether the technology will add to or reduce work).

This finding is consistent with the work by Turja et al [17] which demonstrated that assistive robot acceptance is influenced by factors beyond technical utility. RAM-Care highlights that the intention to use assistive robots is shaped by attitudes, social influence, perceived usefulness, and value alignment, while research on care workers' readiness for robotization has identified psychological and sociodemographic influences on readiness and acceptance [18]. Our present study extends this literature by showing how nurses reason through these issues before routine implementation, particularly in an acute inpatient environment where assistive robots remain an emerging rather than an embedded technology.

Task Boundaries, Professional Identity, and Relational Care

Rather than rejecting assistive robots outright, participants appeared to define moral and professional boundaries around their use. Their accounts suggest that acceptance depended less on the technology itself than on whether its intended role aligned with nurses' expectations regarding appropriate task delegation, clinical judgment, accountability, and relational care. Concerns about "human touch" therefore reflect efforts to preserve aspects of nursing work regarded as fundamentally human, rather than resistance to technology. Viewed in this way, participants' concerns functioned as a form of boundary-setting through which nurses distinguished tasks that could appropriately be supported by robotic tools from those that should remain primarily human.

This interpretation aligns with the literature suggesting that assistive robots may support selected nursing tasks while also raising questions about the fundamentals of care, professional identity, and relational practice [4-6,10]. Archibald and Barnard [10] argue that assistive robots should be considered in relation to fundamental nursing care goals rather than being treated only as technological innovations. Ethical analyses of assistive robots in health care have also emphasized how robotic systems may alter responsibility, agency, and human interaction in care contexts [11].

Participants' view that assistive robots should augment rather than replace nursing work is also consistent with local policy directions that position technology as an enabler of care rather than a substitute for professional expertise [24]. Our study findings contribute a frontline perspective by illustrating how nurses define the acceptable limits of robotic assistance in relation to patient safety, accountability, and relational care.

Safety, Trust, and Human Oversight

Trust emerged as a governance issue rather than merely being an individual attitude. Participants' accounts suggest that confidence in assistive robots depends on organizational arrangements that ensure reliability, transparency, accountability, and appropriate human oversight. These concerns are consistent with the broader human factors literature on automation bias and complacency, where users may place excessive trust in automated systems or fail to detect errors despite their fallibility [25]. Recent health care–focused work similarly highlights automation bias as a risk associated with the increasing introduction of AI applications into clinical practice when failure modes, accountability, and user oversight are not adequately addressed [26].

For nursing practice, these concerns are especially important because patient safety depends on continuous surveillance, contextual judgment, and timely escalation. Participants not only framed assistive robots as autonomous replacements for nurses but also as tools that should operate under human supervision. This supports a governance approach in which assistive robots are introduced with clear task boundaries, escalation criteria, downtime procedures, incident-reporting mechanisms, maintenance responsibilities, and competency requirements.

The implication is that trust should not be treated as a matter of users' attitudes alone. Trust must be designed and governed. Nurses are more likely to view assistive robots as acceptable when systems are reliable, transparent in function, supported by technical infrastructure, and embedded within clear accountability arrangements.

Economic Reasoning and Organizational Readiness

A noteworthy finding was that frontline nurses incorporated organizational and economic considerations into their anticipatory evaluations of assistive robots. Participants questioned whether assistive robots would deliver sufficient benefits to justify investment, whether they could perform meaningful functions within existing workflows, and whether maintenance or troubleshooting would create additional workload. Although economic evaluation and return on investment are often assumed to be managerial concerns, our study findings suggest that frontline nurses also consider organizational sustainability when judging the value of emerging technologies.

This finding extends the literature on robot acceptance by demonstrating that nurses evaluate assistive robots not only through individual acceptance constructs but also through broader considerations such as perceived organizational value. Practical implementation barriers such as cost, technical complexity, and institutional support have similarly been identified in studies of robotic surgery and socially assistive robot implementation [27,28]. Our study's findings indicate that these considerations are also evident in nurses' own reasoning even before routine implementation.

The findings also align with implementation science perspectives that view technology adoption as a complex systems process rather than a simple user acceptance issue. The Non-Adoption, Abandonment, Scale-up, Spread and Sustainability (NASSS) framework highlights that adoption and sustainability of health technologies are influenced by the technology, value proposition, adopters, organizational context, wider systems, and adaptation over time [9]. Applying this lens, successful implementation of assistive robots in nursing would depend not only on whether nurses perceive assistive robots as useful but also on whether the organization can provide infrastructure, technical support,

XSL•FO
RenderX

workflow redesign, training, maintenance, and credible evidence of value.

Practical Implications

Our findings suggest that the implementation of assistive robots in inpatient nursing should follow a staged and governance-oriented approach.

First, organizations should conduct workflow mapping before procurement or implementation. This will allow organizations to identify tasks that are repetitive, physically demanding, time-consuming, or logistical, and distinguish them from tasks requiring clinical judgment, emotional support, complex communication, or relational presence. This approach may help ensure that assistive robots are introduced in ways that are most likely to complement, rather than disrupt, nursing work.

Second, implementation should begin with small-scale pilots before wider deployment. Pilot evaluation should assess not only technical performance but also the impact on nursing workload, patient safety, patient experience, staff trust, workflow integration and/or disruption, maintenance burden, and unintended consequences. This is especially important given participants' concerns that assistive robots could increase rather than reduce workload if poorly integrated.

Third, training should be structured, hands-on, and competency-based. Training should include technical operation, troubleshooting, escalation procedures, downtime management, infection prevention, patient communication, and the limits of robotic assistance. Protected training time may be necessary, as training delivered on top of the routine workload may reinforce perceptions that assistive robots add to the burden rather than reduce it.

Fourth, governance structures should clearly define accountability and support nurses' involvement throughout the implementation process. Organizations should define who is responsible for monitoring robot performance, responding to malfunctions, reporting incidents, maintaining equipment, reviewing safety data, and deciding when human intervention is required. These arrangements are essential to address participants' concerns about overreliance, malfunctions, and patient safety.

Fifth, there should be organizational support for nurses' involvement throughout the implementation stage. Nurses should be engaged early in design, procurement, pilot testing, and evaluation to help identify workflow mismatches, clarify appropriate task boundaries, and address concerns about professional identity and relational care. Nurse champions may support engagement, but champion-led adoption should be accompanied by leadership support, technical infrastructure, and feedback mechanisms.

Overall, implementation strategies should frame assistive robots as supervised tools that augment nursing work, rather than as substitutes for professional judgment or human interaction. Organizations should also establish predefined evaluation metrics to determine whether assistive robots achieve their intended clinical and organizational benefits. These measures may include nursing workload (eg, time spent on routine tasks), operational efficiency (eg, response times and task completion rates), patient safety (eg, robot-related incidents and downtime events), patient experience (eg, patient satisfaction and acceptability), workforce outcomes (eg, nurse satisfaction, perceived workload, and confidence in using assistive robots), and organizational outcomes (eg, robot use, maintenance burden, and service interruptions). Routine monitoring of these indicators would enable organizations to evaluate implementation effectiveness, identify unintended consequences, and refine deployment strategies over time.

Theoretical Implications

This study contributes to the technology acceptance and robot acceptance literature by foregrounding the anticipatory stage of acceptance. Existing models such as TAM, UTAUT, the Almere Model, and RAM-Care provide useful constructs for understanding perceived usefulness, ease of use, social influence, trust, anxiety, and intention to use [14-17]. The findings suggest that in inpatient nursing, anticipatory acceptance also involves professional and ethical reasoning. Nurses evaluated assistive robots in relation to patient safety, relational care, task legitimacy, accountability, workload implications, and organizational sustainability.

The study also contributes to the sociotechnical understanding of health care technology adoption. Participants evaluated assistive robots within broader clinical, organizational, and professional contexts rather than as isolated technological artifacts. Acceptance was shaped by anticipated interactions between technology and existing staffing pressures, workflow demands, patient needs, professional values, maintenance systems, and leadership support. Taken together, our findings support the view that the implementation of assistive robots represents a sociotechnical change in nursing work rather than simply the introduction of a new device.

Limitations

This study has several limitations. First, participants were recruited from 4 general medical inpatient wards in one public tertiary hospital in Singapore where assistive robots had been piloted on a limited basis but were not routinely embedded in practice. The findings therefore reflect anticipatory perspectives rather than experiences of sustained use. Nurses' views may change after direct and repeated interaction with assistive robots in routine care.

Second, convenience sampling may have introduced selection bias. Nurses with stronger views about assistive robots may have been more willing to participate than those who declined participation.

Third, the sample consisted mainly of early-career nurses, which may have shaped the emphasis on professional identity, role security, and confidence in managing emerging technologies. More experienced nurses, advanced practice nurses, and nurse leaders may place greater emphasis on clinical governance, service redesign, workflow integration, and organizational implementation considerations. Future studies should therefore include participants across different career stages and nursing roles.

XSL•FO
RenderX

Fourth, the study was conducted in a digitally enabled hospital context, which may limit transferability to settings with different staffing models, technological maturity, infrastructure, or organizational readiness. Although the findings may be relevant to similar acute inpatient settings where assistive robots are at an early stage of implementation, they should not be assumed to represent all nursing contexts.

Future research should examine how nurses' perceptions evolve after sustained exposure to assistive robots and should incorporate the perspectives of multiple stakeholders. Given participants' concerns about human touch, counseling, and patient experience, future studies should also explore how patients and family caregivers perceive acceptable uses for assistive robots in inpatient care, particularly for activities involving communication, emotional support, and relational aspects of care. Longitudinal and comparative studies are also needed to evaluate the effects of assistive robots on workload, patient safety, patient experience, staff trust, cost-effectiveness, and long-term sustainability.

Conclusions

This study demonstrates that inpatient nurses form nuanced anticipatory judgments about assistive robots before routine implementation. Participants expressed cautious optimism about the potential for assistive robots to reduce workload and support selected nursing tasks, but their acceptance was conditional on safety, reliability, workflow fit, human oversight, relational care, cost-effectiveness, and organizational readiness.

The findings suggest that assistive robot implementation should not be viewed as a purely technical undertaking. Responsible and successful integration will depend on early nurse engagement, workflow-sensitive design, phased piloting, structured training, clear governance, and safeguards that ensure assistive robots augment rather than undermine nursing practice. Nurses' acceptance depends on whether robotic systems are perceived as compatible with professional judgment, patient-centered care, ethical accountability, and practical realities of inpatient work. By foregrounding the anticipatory stage of acceptance, this study highlights how nurses evaluate assistive robots not merely as technical tools, but as potential change agents that may shape professional work, relational care, safety, accountability, and organizational resource use.

Acknowledgments

The authors acknowledge the participants in this study for sharing their time and insights and the Woodlands Hospital nursing leaders for their guidance and support throughout the study. The authors also thank Yong Keng Kwang for his input on the initial study design and Dong Lijuan, Chia Rui Lin, and Sheena Nishanti D/O Ramasamy for their constructive comments and suggestions, which helped improve the quality and clarity of this manuscript. During manuscript revision, the authors used ChatGPT (OpenAI) to assist with language editing, manuscript organization, reviewer-response drafting, and identification of areas requiring clearer explanation. The authors reviewed, edited, and verified all AI-assisted output, including references and substantive claims, and take full responsibility for the accuracy, integrity, and final content of the manuscript. ChatGPT was not used to generate or analyze study data, conduct coding, develop themes, or make final analytic interpretations.

Funding

This work was supported by the Woodlands Health Small Projects Utilising Teams (SPROUTS) Grant (SPR5-24-CI-02).

Data Availability

Deidentified data excerpts relevant to the findings are included in the article. Full interview transcripts are not publicly available due to participant confidentiality and ethical restrictions.

Authors' Contributions

All authors contributed to study conceptualization, methodology, investigation, data curation, formal analysis, funding acquisition, project administration, and manuscript writing, review, and editing. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Checklist 1
COREQ Checklist
[PDF File, 112 KB - nursing_v9i1e91106_app1.pdf]

References

1. World Robotics 2025 report – SERVICE ROBOTS – released by IFR. International Federation of Robotics (IFR). URL: https://ifr.org/ifr-press-releases/news/service-robots-see-global-growth-boom [accessed 2026-07-08]

XSL•FO
RenderX

2. Alboul L, Dimitrova M, Lekova A, Kaburlasos VG, Mitrouchev P. Editorial: emerging technologies for assistive robotics: current challenges and perspectives. Front Robot AI 2023;10:1288360. [doi: 10.3389/frobt.2023.1288360] [Medline: 37881772]
3. Javaid M, Haleem A, Pratap Singh R, Rab S, Suman R, Kumar L. Utilization of robotics for healthcare: a scoping review. J Ind Integr Mgmt 2025;10(1):43-65. [doi: 10.1142/S2424862222500154]
4. Christoforou EG, Avgousti S, Ramdani N, Novales C, Panayides AS. The upcoming role for nursing and assistive robotics: opportunities and challenges ahead. Front Digit Health 2020;2:585656. [doi: 10.3389/fdgth.2020.585656] [Medline: 34713058]
5. Kangasniemi M, Karki S, Colley N, Voutilainen A. The use of robots and other automated devices in nurses' work: an integrative review. Int J Nurs Pract 2019 Aug;25(4):e12739. [doi: 10.1111/ijn.12739] [Medline: 31069892]
6. Papadopoulos I, Koulouglioti C, Ali S. Views of nurses and other health and social care workers on the use of assistive humanoid and animal-like robots in health and social care: a scoping review. Contemp Nurse 2018;54(4-5):425-442. [doi: 10.1080/10376178.2018.1519374] [Medline: 30200824]
7. Poon YSR, Lin YP, Griffiths P, Yong KK, Seah B, Liaw SY. A global overview of healthcare workers' turnover intention amid COVID-19 pandemic: a systematic review with future directions. Hum Resour Health 2022 Sep 24;20(1):70. [doi: 10.1186/s12960-022-00764-7] [Medline: 36153534]
8. Mitchell E, Walker R. Global ageing: successes, challenges and opportunities. Br J Hosp Med (Lond) 2020 Feb 2;81(2):1-9. [doi: 10.12968/hmed.2019.0377] [Medline: 32097073]
9. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
10. Archibald MM, Barnard A. Futurism in nursing: technology, robotics and the fundamentals of care. J Clin Nurs 2018 Jun;27(11-12):2473-2480. [doi: 10.1111/jocn.14081] [Medline: 28940657]
11. Stahl BC, Coeckelbergh M. Ethics of healthcare robotics: towards responsible research and innovation. Rob Auton Syst 2016 Dec;86:152-161. [doi: 10.1016/j.robot.2016.08.018]
12. Mucchiani C, Cacchione P, Johnson M, Mead R, Yim M. Deployment of a socially assistive robot for assessment of COVID-19 symptoms and exposure at an elder care setting. Proc IEEE Int Conf Robot Hum Interact Commun (RO-MAN) 2021:1189-1195. [doi: 10.1109/RO-MAN50785.2021.9515551]
13. Kang HS, Koh IS, Makimoto K, Yamakawa M. Nurses' perception towards care robots and their work experience with socially assistive technology during COVID-19: a qualitative study. Geriatr Nurs 2023;50:234-239. [doi: 10.1016/j.gerinurse.2023.01.025] [Medline: 36805955]
14. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
15. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep;27(3):425-478. [doi: 10.2307/30036540]
16. Heerink M, Kröse B, Evers V, Wielinga B. Assessing acceptance of assistive social agent technology by older adults: the Almere model. Int J Soc Robotics 2010 Dec;2(4):361-375. [doi: 10.1007/s12369-010-0068-5]
17. Turja T, Aaltonen I, Taipale S, Oksanen A. Robot acceptance model for care (RAM-care): a principled approach to the intention to use care robots. Inf Manage 2020 Jul;57(5):103220. [doi: 10.1016/j.im.2019.103220]
18. Turja T, Taipale S, Kaakinen M, Oksanen A. Care workers' readiness for robotization: identifying psychological and socio-demographic determinants. Int J Soc Robotics 2020 Jan;12(1):79-90. [doi: 10.1007/s12369-019-00544-9]
19. Andtfolk M, Nyholm L, Eide H, Rauhala A, Fagerström L. Attitudes toward the use of humanoid robots in healthcare—a cross-sectional study. AI Soc 2022 Dec;37(4):1739-1748. [doi: 10.1007/s00146-021-01271-4]
20. Felding SA, Koh WQ, Teupen S, Budak KB, Laporte Uribe F, Roes M. A scoping review using the Almere model to understand factors facilitating and hindering the acceptance of social robots in nursing homes. Int J Soc Robotics 2023 Jul;15(7):1115-1153. [doi: 10.1007/s12369-023-01012-1]
21. National Academies of Sciences, Engineering, and Medicine. The Future of Nursing 2020-2030: Charting a Path to Achieve Health Equity: National Academies Press; 2021. [doi: 10.17226/25982]
22. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
23. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
24. Keynote Address by Mr Ong Ye Kung, Minister for Health, at the Singapore Health & Biomedical Congress 2024, on Thursday 10 October 2024, Singapore Expo. Ministry of Health. URL: https://www.moh.gov.sg/newsroom/keynote-address-by-mr-ong-ye-kung--minister-for-health--at-the-singapore-health---biomedical-congress-2024--on-thursday-10-october-2024--singapore-expo [accessed 2026-07-08]
25. Wickens CD, Clegg BA, Vieane AZ, Sebok AL. Complacency and automation bias in the use of imperfect automation. Hum Factors 2015 Aug;57(5):728-739. [doi: 10.1177/0018720815581940] [Medline: 25886768]

XSL•FO
RenderX

26. Abdelwanis M, Alarafati HK, Tammam MMS, Simsekler MCE. Exploring the risks of automation bias in healthcare artificial intelligence applications: a Bowtie analysis. J Saf Sci Resilience 2024 Dec;5(4):460-469. [doi: 10.1016/j.jnlssr.2024.06.001]
27. Lawrie L, Gillies K, Duncan E, Davies L, Beard D, Campbell MK. Barriers and enablers to the effective implementation of robotic assisted surgery. PLoS One 2022;17(8):e0273696. [doi: 10.1371/journal.pone.0273696] [Medline: 36037179]
28. Zafrani O, Nimrod G, Krakovski M, Kumar S, Bar-Haim S, Edan Y. Assimilation of socially assistive robots by older adults: an interplay of uses, constraints and outcomes. Front Robot AI 2024;11:1337380. [doi: 10.3389/frobt.2024.1337380] [Medline: 38646472]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
NASSS: Non-Adoption, Abandonment, Scale-up, Spread and Sustainability
RAM-Care: Robot Acceptance Model for Care
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by E Borycki; submitted 11.Jan.2026; peer-reviewed by FA Dhabbari, J Leoste; revised version received 09.Jul.2026; accepted 16.Jul.2026; published 10.Aug.2026.

Please cite as:
Tan MW, Chan TLT, Taye CSM
Inpatient Nurses' Anticipatory Perspectives on Assistive Robots: Qualitative Study
JMIR Nursing 2026;9:e91106
URL: https://nursing.jmir.org/2026/1/e91106
doi:10.2196/91106

© Mavis Weiting Tan, Timie Lai Teng Chan, Charmaine Shi Min Taye. Originally published in JMIR Nursing (https://nursing.jmir.org), 10.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Insights Into Factors Affecting Nurses' Knowledge of and Attitudes Toward AI and Implications for Successful AI Integration in Critical Care: Cross-Sectional Study

Habib Alrashedi[1], PhD; Saad M Alderaan[1,2], MSN; Nader Alnomasy[1], PhD; Hamdi Lamine[3], PhD; Khalil A Saleh[1], PhD; Sameer A Alkubati[1], PhD

[1]Medical-Surgical Nursing Department, College of Nursing, University of Ha'il, Baqaa street, Hail, Saudi Arabia

[2]Hail Health Cluster, King Khaled Hospital, Hail, Saudi Arabia

[3]Community Health Nursing Department, College of Nursing, University of Ha'il, Hail, Saudi Arabia

Corresponding Author:
Habib Alrashedi, PhD
Medical-Surgical Nursing Department, College of Nursing, University of Ha'il, Baqaa street, Hail, Saudi Arabia

Abstract

Background: Assessing the current landscape of nurses' knowledge and attitudes is a critical first step in facilitating a smooth and effective transition toward artificial intelligence (AI)–enhanced critical care.

Objective: This study aimed to assess the levels of and factors affecting the knowledge of and general attitudes toward AI in critical care among nurses.

Methods: A cross-sectional correlational design was used with 203 critical care nurses in Hail, Saudi Arabia, using the Nurses' AI Knowledge Questionnaire and the 20-item General Attitudes Toward Artificial Intelligence Scale from May 2025 to July 2025. Data were analyzed using 2-tailed *t* tests, ANOVA, Pearson correlation, and multivariable linear regression. Statistical significance was set at $P<.05$.

Results: Critical care nurses demonstrated moderate knowledge of (mean score 4.93, SD 1.78) and positive attitudes toward AI (mean score 64.39, SD 8.26). A moderate positive correlation was found between knowledge of and attitudes toward AI ($r=0.45$; $P<.001$). In multivariable analyses, older age was associated with lower knowledge (≥40 years: β=−1.29, 95% CI −2.12 to −0.45; $P=.003$) and less positive attitudes (β=−8.97, 95% CI −12.49 to −5.44; $P<.001$). Female nurses reported lower knowledge (β=−0.69, 95% CI −1.20 to −0.19; $P=.007$) and less positive attitudes (β=−2.65, 95% CI −4.78 to −0.52; $P=.02$) than male nurses. Greater experience (>5 years) was positively associated with knowledge (β=1.20, 95% CI 0.65 - 1.75; $P<.001$) and attitudes (β=8.08, 95% CI 5.76 - 10.41; $P<.001$).

Conclusions: Critical care nurses in Hail demonstrated moderate knowledge of and positive attitudes toward AI, which varied based on their demographic and professional characteristics. These findings highlight the need to strengthen AI literacy and provide targeted support to groups with lower scores, which may enhance readiness for AI integration in critical care settings.

(JMIR Nursing 2026;9:e85649) doi:10.2196/85649

KEYWORDS

artificial intelligence; AI; knowledge; attitudes; factors; critical care nurses; Saudi Arabia

Introduction

The global health care model is undergoing a great revolution directed by the rapid incorporation of artificial intelligence (AI) [1,2]. AI is defined as a system designed to perceive the environment and take action to achieve specific goals [3]. AI encompasses machine learning, natural language processing, and robotics [2,4]. In medicine, AI has vast and promising potential for enhancing diagnostic precision; personalizing treatment plans; optimizing operational efficiency; and, ultimately, improving patient outcomes [5,6]. This technological shift aligns with global initiatives such as Vision 2030, which actively encourages innovation and digital transformation within its health sector to build a robust, data-driven health care system [7,8].

Critical care units, including intensive care units and emergency departments, represent environments with exceptionally high pressure [9]. Clinicians in these settings are required to process large amounts of complex, real-time patient data to make swift decisions. AI applications can be leveraged for the early detection of patient deterioration, prediction of sepsis, forecasting intensive care unit length of stay, and managing ventilatory support. These capabilities support clinical decision-making and potentially reduce human burden [10]. By

XSL•FO
RenderX

automating routine tasks and administrative burdens, AI can free critical care nurses (CCNs) and physicians to focus on more complex clinical reasoning and direct patient care. This shift enhances both the efficiency and humanistic aspects of treatment [11].

As frontline clinicians, nurses are pivotal to the successful adoption of new technologies in clinical practice [12]. Their role involves continuous patient monitoring, assessment, and execution of complex care plans, making them key end users of AI-driven tools. Therefore, the effective integration of AI into critical care is linked to nurses' acceptance, which is shaped by their knowledge, attitudes, and willingness to incorporate these technologies into their workflow [13,14]. However, the introduction of AI in the nursing domain has sparked debate. While some view AI as a tool to augment nursing practice and mitigate workload, others perceive it as a threat to the essential human-to-human interactions that form the bedrock of compassionate care. This difference in perspective raises concerns about dehumanization and ethical implications [15-17].

A significant barrier to AI's integration is the current underrepresentation of nurses in the development, implementation, and evaluation of AI systems for health care [18,19]. This gap can lead to a misalignment among technological solutions, actual clinical needs, and workflow. Furthermore, studies have indicated that nurses' perceptions of AI are mixed and vary widely based on their understanding of its capabilities, reliability, and potential to replace human judgment [20]. Therefore, assessing the current landscape of nurses' knowledge and attitudes is a critical first step in facilitating a smooth and effective transition toward AI-enhanced critical care.

Previous research has begun to explore health care professionals' perspectives on AI, but studies focused on CCNs within the Middle Eastern context, particularly in Saudi Arabia, remain limited. Understanding the demographic, educational, and experiential factors that influence these perceptions is crucial for developing targeted educational and training programs. As critical care environments become increasingly technologically advanced, ensuring that the nursing workforce is not only proficient but also confident and ethically grounded in using AI is paramount.

This study aimed to bridge this knowledge gap by assessing CCNs' level of knowledge of and general attitudes toward AI in Hail, Saudi Arabia. The findings provide valuable insights for hospital administrators, educators, and policymakers in designing strategies that foster AI literacy and address concerns. The ultimate goal is to harness the full potential of AI to support rather than replace the critical role of nurses in delivering high-acuity patient care.

Methods

Design, Setting, Population, and Sample

A cross-sectional correlational design was used in this study. The target population of this study comprised CCNs employed in public hospitals located in the Hail region. After obtaining institutional review board approval, meetings were conducted with the heads of critical care units across participating hospitals. In addition, formal communication was established with continuing nursing education offices within these institutions. The survey was designed using Google Forms. The link to the questionnaire and informed consent form was disseminated to the CCNs. Participants were first given an electronic information sheet outlining the study's goals, procedures, risks, benefits, confidentiality measures, and voluntary nature of participation. Data collection was conducted over 3 months, from May 2025 to July 2025. Eligibility criteria required participants to have at least 1 year of continuous experience working in critical care departments, ensuring that only nurses with sufficient exposure to clinical practice in critical situations were included. Nurses serving primarily in administrative roles, as well as those with less than 1 year of critical care department experience, were excluded to maintain a focus on direct patient care providers with adequate professional backgrounds. The required sample size was determined using OpenEpi (version 3.01). On the basis of an estimated total population of approximately 420 CCNs in the region, the minimum sample size necessary to achieve adequate statistical power was 201, with a 95% confidence level and 5% margin of error. To enhance representativeness and mitigate the potential impact of nonresponses, 220 self-administered questionnaires were distributed. Of these 220 questionnaires, 203 (92.3%) were returned. The electronic survey required responses to all the items before submission; therefore, there were no missing data.

Instruments

The questionnaire consisted of 3 parts. In the first part, the characteristics of the nurses, including their age, sex, marital status, educational level, years of experience, type of shift work, unit type, and prior experience with AI in health care, were examined. We used a previously validated tool by Swed et al [21] (the Nurses' AI Knowledge Questionnaire) in the second segment of the questionnaire to gauge nurses' awareness of AI. It consists of 7 yes-or-no questions regarding common AI terms used in health care designed to gauge nurses' familiarity with this key vocabulary. The scoring system is as follows: "yes" answers are scored as 1 point, and "no" answers are scored as 0 points. The total score ranges from 0 to 7, with a higher score indicating a higher level of knowledge regarding AI terminology [21]. As Swed et al [21] noted, the Cronbach α value of 0.795 demonstrated the tool's internal consistency among the subscales. In this study, the reliability of this instrument was confirmed with a Cronbach α of 0.765.

The 20-item General Attitudes Toward Artificial Intelligence Scale (GAAIS) created by Schepman and Rodway [22] constituted the third section of the questionnaire. It gauged nurses' opinions on the use of AI in medical environments. The items are divided into positive (12 items) and negative (8 items). Positive items are scored on a 5-point Likert-type scale, with 1 denoting "strongly disagree" and 5 denoting "strongly agree." Negative items are reverse scored, with 1 denoting "strongly agree" and 5 denoting "strongly disagree." Thus, the scores range from 20 to 100, with higher scores on each subscale reflecting more positive attitudes [22]. According to Schepman and Rodway [22], the GAAIS has demonstrated a high degree of internal consistency, with Cronbach α values for the 12

XSL•FO
RenderX

positive items and 8 negative items being 0.88 and 0.82, respectively. In this study, reliability was confirmed with a Cronbach α of 0.969 for positive items and 0.952 for negative items.

Ethical Considerations

Institutional review board approval was obtained from the University of Hail (H-2025-718) on March 10, 2025, and from the Ministry of Health (2025-37) on March 18, 2025. In compliance with institutional review board approval, informed consent was obtained electronically: participants were required to study the information page before completing and submitting the survey, which constituted informed consent. The participants' anonymity and confidentiality were maintained throughout the study. As the survey platform required all items to be completed before submission, there were no partial responses or missing data. No compensation or incentives were provided for participation in the study.

Data Analysis

SPSS Statistics (version 27; IBM Corp) was used to analyze the data. The Shapiro-Wilk test was used to test for normality of the data ($P>.05$). Independent-sample *t* tests and one-way ANOVA were used to investigate the relationship between the dependent and independent variables. Using multivariable linear regression analysis, significant factors affecting CCNs' knowledge and attitudes were identified. Correlations between variables were assessed using the Pearson correlation coefficient. The *P* value was set at less than .05.

Results

Table 1 shows that most participants were aged 20 to 29 years (121/203, 59.6%), male (129/203, 63.5%), and single (140/203, 69%). Most nurses held a bachelor's degree (141/203, 69.5%), worked rotating shifts (131/203, 64.5%), and had 5 years or less of nursing experience (117/203, 57.6%). Younger nurses showed significantly higher knowledge of ($P=.01$) and more positive attitudes toward AI ($P=.002$) than older nurses. Male nurses reported higher knowledge and more positive attitudes than female nurses ($P<.001$ in both cases). Single nurses scored higher on knowledge ($P=.03$) and attitudes ($P=.046$) than married nurses. Nurses with a master's degree had higher knowledge ($P=.02$) and more positive attitudes ($P<.001$) than those with a bachelor's degree. Additionally, nurses with more than 5 years of experience exhibited higher knowledge ($P=.02$) and more positive attitudes ($P<.001$) than less experienced nurses.

XSL•FO
RenderX

Table . Relationship between critical care nurses' (CCNs) sociodemographic characteristics and knowledge of and attitudes toward artificial intelligence (N=203).

Variable and categories	CCNs, n (%)	Knowledge			Attitudes		
		Score (0-7), mean (SD)	*t* test (*df*) or *F* test (*df*)	*P* value	Score (20-100), mean (SD)	*t* test (*df*) or *F* test (*df*)	*P* value
Age (years)			4.61 (2)[a]	.01		6.55 (2)[a]	.002
20-29	121 (59.6)	5.20 (1.68)			65.60 (8.16)		
30-39	60 (29.6)	4.68 (1.94)			63.96 (8.44)		
≥40	22 (10.8)	4.09 (1.60)			58.90 (5.99)		
Sex			3.78 (201)[b]	<.001		3.54 (201)[b]	<.001
Male	129 (63.5)	5.27 (1.65)			65.90 (8.32)		
Female	74 (36.5)	4.32 (1.85)			61.75 (7.53)		
Marital status			2.20 (201)[b]	.03		1.88 (201)[b]	.046
Single	140 (69.0)	5.11 (1.71)			65.12 (8.62)		
Married	63 (31.0)	4.52 (1.88)			62.77 (7.22)		
Educational level			−2.35 (201)[b]	.02		−4.02 (201)[b]	<.001
Bachelor's	141 (69.5)	4.73 (1.80)			62.90 (7.17)		
Master's	62 (30.5)	5.37 (1.68)			67.79 (9.56)		
Shift type			−0.41 (201)[b]	.68		−1.09 (201)[b]	.28
Day	72 (35.5)	4.86 (1.78)			63.54 (7.79)		
Rotating	131 (64.5)	4.96 (1.79)			64.86 (8.51)		
Experience (years)			−2.41 (201)[b]	.02		−4.38 (201)[b]	<.001
≤5	117 (57.6)	4.67 (1.78)			62.30 (6.91)		
>5	86 (42.4)	5.27 (1.73)			67.23 (9.12)		

[a]*F* test.

[b]2-tailed *t* test.

The mean score for CCNs' knowledge of AI was 4.93 (SD 1.78; range 2 - 7), indicating a moderate level of knowledge. The mean score for attitudes was 64.39 (SD 8.26; range 47 - 95), reflecting a generally positive attitude toward AI (Table 2).

Table 3 shows that there was a moderate positive correlation between CCNs' knowledge of and attitudes toward AI (r=0.45; P<.001). This finding indicates that higher knowledge levels are associated with more positive attitudes.

Table . Means of critical care nurses' knowledge of and attitudes toward artificial intelligence (N=203).

Variable	Score, mean (SD; range)
Knowledge (0-7)	4.93 (1.78; 2-7)
Attitudes (20-100)	64.39 (8.26; 47-95)

Table . Correlation between study variables (N=203).

	Knowledge	Attitudes
Knowledge	1	0.45
Attitudes	0.45[a]	1

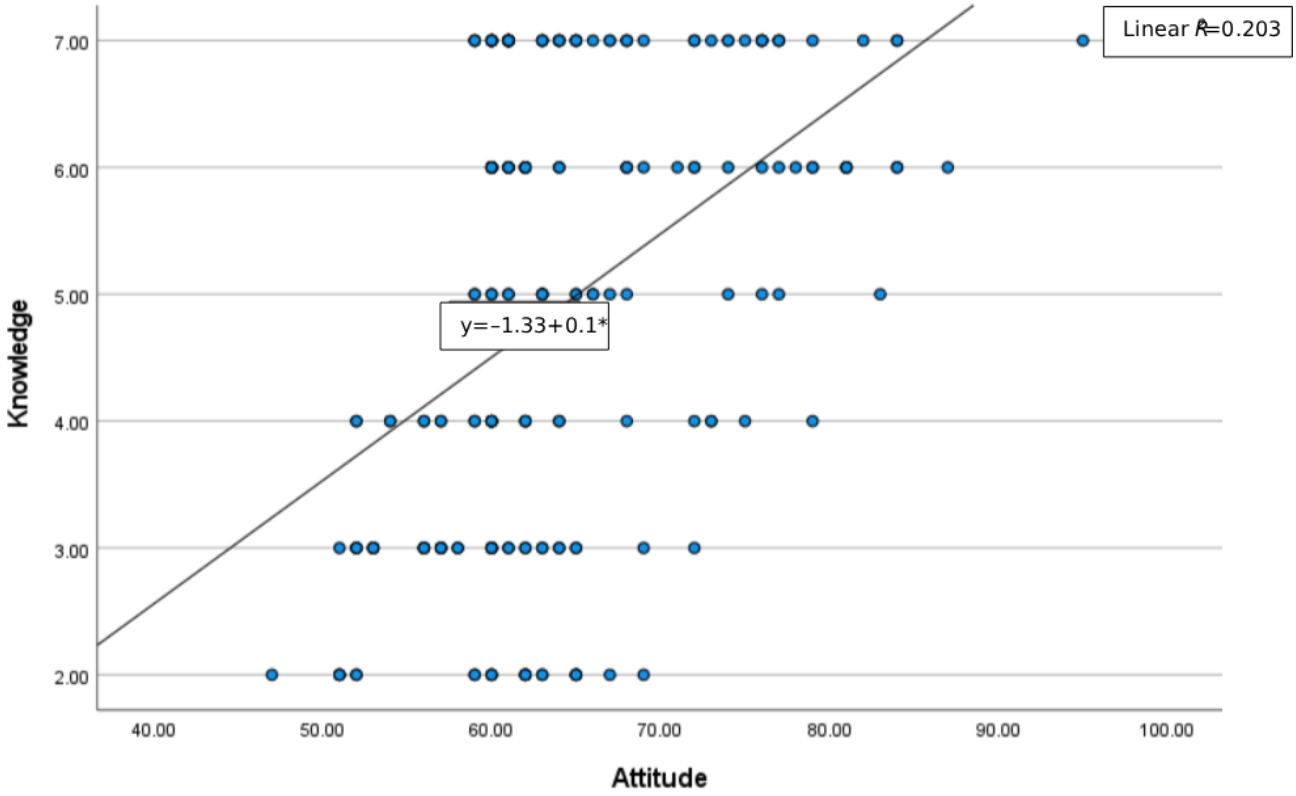
[a]Correlation is significant at the .01 level (2 tailed).

Figure 1 illustrates the moderate positive correlation between CCNs' knowledge of and attitudes toward AI. The scatterplot shows that higher attitude scores were associated with higher knowledge levels. The regression line indicates a significant positive linear relationship (R^2=0.20, P<.001), suggesting that

XSL•FO
RenderX

attitudes were associated with approximately 20% of the variance in knowledge.

Figure 1. Scatterplot of the correlation between critical care nurses' knowledge of and attitudes toward artificial intelligence.



The regression models identified several sociodemographic predictors of CCNs' knowledge of and attitudes toward AI. Age and sex were found to be significant predictors of knowledge. Nurses aged 30 to 39 years (β=–0.80; *P*=.02) and those aged ≥40 years (β=–1.29; *P*=.003) had lower knowledge scores than those aged 20 to 29 years. Similarly, female nurses reported significantly lower knowledge scores than their male counterparts (β=−0.69; *P*=.007). In contrast, nurses with more than 5 years of experience had significantly higher knowledge levels (β=1.20; *P*<.001). The model's R^2 was 0.19 (adjusted R^2=0.17; *P*<.001), indicating that the included predictors were associated with approximately 19.4% of the variability in knowledge scores.

Age, sex, educational level, and experience were significant predictors of attitudes. Nurses aged 30 to 39 years (β=–4.81; *P*=.001) and those aged ≥40 years (β=–8.97; *P*<.001) reported less positive attitudes than those aged 20 to 29 years. Female nurses had significantly less positive attitudes than male nurses (β=–2.65; *P*=.02). Conversely, nurses with a master's degree (β=3.38; *P*=.002) and those with more than 5 years of experience (β=8.08; *P*<.001) demonstrated more positive attitudes. The model's R^2 was 0.33 (adjusted R^2=0.31; *P*<.001), indicating that the included predictors were associated with approximately 33.2% of the variability in attitude scores (Table 4).

XSL•FO
RenderX

Table . Multiple linear regression for factors affecting critical care nurses' knowledge of and attitudes toward artificial intelligence.

Factor	Knowledge[a]		Attitudes[b]	
	β (95% CI)	*P* value	β (95% CI)	*P* value
Age (years)				
20-29	Reference	Reference	Reference	Reference
30-39	−0.80 (−1.47 to –0.14)	.02	−4.81 (−7.59 to –2.02)	.001
≥40	−1.29 (−2.12 to –0.45)	.003	−8.97 (−12.49 to –5.44)	<.001
Gender				
Male	Reference	Reference	Reference	Reference
Female	−0.69 (−1.20 to –0.19)	.007	−2.65 (−4.78 to –0.52)	.02
Marital status				
Single	Reference	Reference	Reference	Reference
Married	−0.44 (−0.98 to 0.10)	.11	−1.73 (−4.01 to 0.55)	.14
Educational level				
Bachelor's degree	Reference	Reference	Reference	Reference
Master's degree	0.40 (−0.10 to 0.90)	.12	3.38 (1.270 to 5.492)	.002
Experience (years)				
≤5	Reference	Reference	Reference	Reference
>5	1.20 (0.65 to 1.75)	<.001	8.08 (5.76 to 10.41)	<.001

[a]R^2=0.19; adjusted R^2=0.17; P<.001.

[b]R^2=0.33; adjusted R^2=0.31; P<.001.

Discussion

CCNs' Knowledge of and Attitudes Toward AI

This study provides a timely investigation of CCNs' knowledge of and attitudes toward AI in the Hail region of Saudi Arabia, a context undergoing rapid digital transformation as part of Vision 2030. The findings revealed a moderate level of AI knowledge (mean score 4.93, SD 1.78) and a generally positive attitude (mean score 64.39, SD 8.26) among CCNs. Crucially, a significant positive correlation was established, indicating that higher levels of AI knowledge were associated with more favorable attitudes. This aligns with the technology acceptance model (TAM), which posits that perceived usefulness and ease of use are key determinants of technology adoption and that these perceptions are inherently linked to an individual's understanding of the technology [23-25]. Our results suggest that educational interventions aimed at improving AI literacy could be a powerful lever for enhancing acceptance among the nursing workforce.

Predictors of CCNs' Knowledge of and Attitudes Toward AI

Sociodemographic analyses yielded insightful results. Younger nurses (aged 20-29 years) exhibited significantly higher knowledge and more positive attitudes than their older counterparts. This generational divide is consistent with the broader literature on technology adoption, where younger individuals, often "digital natives," tend to be more comfortable and familiar with emerging technologies [26,27]. This finding underscores the need for age-tailored training programs that support more experienced nurses in developing comparable levels of perceived ease of use and usefulness, thereby reducing TAM-related barriers among older staff while leveraging their clinical expertise.

Gender emerged as a significant predictor, with male nurses reporting higher knowledge and more positive attitudes than female nurses. This disparity may reflect broader societal and educational trends in the science, technology, engineering, and mathematics fields, where gender gaps in confidence and participation persist [28]. In the nursing context, which is predominantly female in many countries but has a different demographic profile in regions such as Saudi Arabia, this finding underscores the need for equitable access to AI training and leadership opportunities. Ensuring that these opportunities are encouraging and available to all genders is essential to preventing a new form of digital gender divide within the profession. From a TAM perspective, such differences may translate into unequal perceptions of ease of use and self-efficacy with AI systems, highlighting the importance of designing AI training and leadership opportunities that actively foster confidence and perceived control among women to prevent a digital gender divide in nursing.

Educational attainment has a strong positive influence. Nurses with a master's degree had significantly more positive attitudes and higher knowledge scores than those with a bachelor's degree. This finding reinforces the pivotal role of advanced education in fostering a forward-looking, evidence-based, and innovative mindset. This suggests that integrating AI concepts and applications into postgraduate nursing curricula is essential

XSL•FO
RenderX

for preparing future nurse leaders [6,29]. Furthermore, contrary to what might be assumed, nurses with more than 5 years of experience reported higher knowledge and more positive attitudes. This compelling finding challenges the notion that experienced clinicians are resistant to change. Instead, it implies that experienced nurses, with their developed clinical expertise, may better appreciate AI's potential to alleviate cognitive burdens, reduce errors, and enhance patient safety [11].

The regression models indicated that 19.4% of the variance in knowledge and 33.2% of the variance in attitudes were related to demographic factors. However, other variables not measured in this study also played a significant role. These could include organizational culture, quality of previous technology implementation experiences, perceived organizational support for training, and the level of trust in the institution's data governance and ethical frameworks [30,31]. Future research should explore these organizational and psychological determinants to provide a more holistic understanding of the factors that influence AI integration in nursing.

Correlation Between CCNs' Knowledge of and Attitudes Toward AI

The moderate positive correlation between knowledge and attitude (r=0.45; P<.001) strongly implies that resistance or skepticism toward AI is not unchangeable but can be mitigated through education and exposure. Within the TAM framework, this implies that structured education and meaningful hands-on experience can reshape nurses' beliefs about AI's usefulness and ease of use, moving them from passive compliance to active, informed adoption of AI tools in clinical workflows. This finding aligns with the work by Dornan [32], who suggested that a basic understanding of AI is essential for its acceptance and use in clinical practice. When nurses understand how AI works, what it can offer, and what its limits are, the technology becomes less intimidating [32]. This awareness helps them move from being passive recipients of change to being active and informed participants. Such a shift encourages genuine engagement rather than simple compliance [32]. Therefore, the observed link highlights the need for targeted education programs designed to build nurses' confidence and skills in working with AI, ensuring that technology, rather than distance, enhances nursing care. Thus, the primary barrier is not an inherent opposition to technology but a lack of structured and accessible education on what AI truly entails in nursing practice.

Implications for Nursing Education and Practice

The findings underscore the need for specific measures to improve AI preparedness among nurses through education. Older, female, and less experienced nurses had less knowledge of and a negative attitude toward AI, suggesting possible gaps in confidence and exposure that must be addressed through systematic training activities. Incorporating fundamental AI principles and practical applications into undergraduate and postgraduate nursing courses is critical for ensuring that all future nurses are prepared to work with developing technology. In nursing practice, continuing professional development programs that include practical training, simulation-based learning, and case-based scenarios can help improve comprehension and minimize anxiety. Furthermore, nursing professionals with a more optimistic mindset can act as clinical representatives to assist in collaborative learning and ensure the smooth integration of AI in critical care settings. Improving AI literacy across the nursing profession will eventually lead to safer and more efficient clinical practice and successful incorporation of AI-driven strategies in patient care.

These findings are particularly important in the context of Saudi Arabia's ambitious health sector reforms. For AI to be successfully leveraged to build a robust, data-driven health care system as envisioned in Vision 2030, the readiness of the nursing workforce is fundamental. Aligning educational strategies, professional development, and organizational policies with TAM principles by explicitly targeting perceived usefulness, ease of use, and supportive conditions can help translate the current foundational willingness among nurses into sustained, confident use of AI in everyday practice. Our study confirms that foundational willingness is present but must be actively cultivated through targeted, demographically sensitive, and continuous educational strategies.

Strengths and Limitations

This study has several strengths. First, it addresses a significant gap in the literature by focusing specifically on CCNs in the underresearched Middle Eastern context, thus providing valuable insights for regional policymaking and educational planning. The use of validated instruments such as the Nurses' AI Knowledge Questionnaire and the GAAIS enhances the reliability and validity of the findings. Furthermore, the high response rate (203/220, 92.3% returned questionnaires) and rigorous sample size calculation strengthened the statistical power and representativeness of the results for the target population in Hail.

Despite these strengths, this study has several limitations that must be acknowledged. The cross-sectional correlational design captures a snapshot in time and cannot establish causality between the variables. This study was conducted in a single region of Saudi Arabia (Hail), which may limit the generalizability of the findings to other regions or countries with different cultural and health care infrastructures. The reliance on self-reported data for knowledge and attitudes is susceptible to social desirability bias, in which participants may have provided answers that they believed were expected rather than their true beliefs. Finally, the knowledge assessment was based on a 7-item yes-or-no questionnaire, which, while reliable, may not capture the full depth and nuance of a nurse's understanding of AI concepts and applications.

Recommendations

On the basis of this study's findings, the following recommendations are proposed.

Recommendations for Practice

Health care institutions should implement structured, ongoing AI education programs that build basic literacy for all CCNs and provide advanced modules for those in higher-responsibility roles. These programs should use practical, critical care examples to enhance perceived usefulness and ease of use. Targeted support is also needed for older and female nurses

XSL•FO
RenderX

who showed lower knowledge and more negative attitudes through tailored workshops, mentorship, and simulation-based training to ensure equitable AI readiness. Additionally, nurses with master's degrees and more than 5 years of experience should be leveraged as AI champions to mentor colleagues and support effective implementation across critical care teams.

Recommendations for Education

Nursing schools and universities should integrate core AI content—concepts, clinical applications, ethics, and limitations—into undergraduate curricula while expanding advanced, practice-oriented AI training in postgraduate programs, reflecting the strong association between higher education and more positive perceptions. Alignment between academic preparation and workplace training will create a continuous pipeline of nurses who are both knowledgeable about and favorably disposed toward AI in critical care.

Recommendations for Policy

Hospital leadership and AI developers should systematically involve CCNs from diverse age groups, genders, and educational backgrounds in the design, piloting, and evaluation of AI tools because the regression results suggest that attitudes and knowledge are shaped partly by contextual and experiential factors beyond demographics alone. Participatory implementation can improve the perceived relevance and usability of AI systems, thereby reinforcing the positive relationship between knowledge and attitudes and supporting Saudi Arabia's Vision 2030 for a technology-driven, nurse-ready health care system.

Conclusions

This study demonstrated that CCNs in Hail, Saudi Arabia, possess a moderate level of knowledge and a generally positive attitude toward AI, with a clear correlation between the 2. Key sociodemographic factors, including age, sex, educational level, and clinical experience, significantly influenced these perceptions. The findings underscore that the successful integration of AI into critical care is not a technological challenge but a human-centric one. Readiness of the nursing workforce is a critical determinant of success. By investing in comprehensive, inclusive, and continuous education and actively involving nurses in the development process, health care leaders can harness the full potential of AI. This will ensure that these powerful technologies act as supportive tools that augment the clinical judgment and compassionate care provided by nurses, ultimately leading to enhanced patient outcomes and the realization of a technologically advanced, efficient, and resilient health care system as envisioned in the Saudi Vision 2030.

Acknowledgments

The authors express their gratitude to all critical care nurses who participated in their research.

Funding

This research was funded by the Deanship of Scientific Research at the University of Hail, Saudi Arabia, under project RG-23 048.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: HA
Data curation: HA, SMA, NA, KAS, HL, SAA
Formal analysis: SAA
Investigation: HA, SMA, NA, KAS, HL
Methodology: HA, SAA
Resources: HA, SAA
Validation: HA
Visualization: HA
Writing – original draft: HA
Writing – review & editing: HA, SMA, NA, KAS, HL, SAA
All authors read, revised, and accepted the final version of the manuscript for publication.

Conflicts of Interest

None declared.

References

1. Alnomasy N, Pangket P, Alreshidi B, et al. Artificial intelligence in health care: assessing impact on professional roles and preparedness among hospital nurse leaders. Digit Health 2025;11:20552076251356362. [doi: 10.1177/20552076251356362] [Medline: 40631348]

XSL•FO
RenderX

2. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021 Jul;8(2):e188-e194. [doi: 10.7861/fhj.2021-0095] [Medline: 34286183]
3. Collins C, Dennehy D, Conboy K, Mikalef P. Artificial intelligence in information systems research: a systematic literature review and research agenda. Int J Inf Manage 2021 Oct;60:102383. [doi: 10.1016/j.ijinfomgt.2021.102383]
4. Hendy A, Ibrahim RK, Abdelaliem SM, et al. Supervised machine learning for classification and prediction of stunting among under-five Egyptian children. BMC Pediatr 2025 Sep 18;25(1):681. [doi: 10.1186/s12887-025-06138-x] [Medline: 40963124]
5. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Dec;2(4):230-243. [doi: 10.1136/svn-2017-000101] [Medline: 29507784]
6. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019 Jan;25(1):44-56. [doi: 10.1038/s41591-018-0300-7]
7. Saudi Vision 2030. URL: https://www.vision2030.gov.sa/en [accessed 2025-10-05]
8. AlWatban N, Othman F, Almosnid N, AlKadi K, Alajaji M, Aldeghaither D. The emergence and growth of digital health in Saudi Arabia: a success story. In: Kozlakidis Z, Muradyan A, Sargsyan K, editors. Digitalization of Medicine in Low- and Middle-Income Countries: Paradigm Changes in Healthcare and Biomedical Research: Springer International Publishing; 2024. [doi: 10.1007/978-3-031-62332-5_3]
9. Alrashedi H, Alnomasy N, Saleh KA, Lamine H, Alkubati SA. The mediating effect of moral distress on the relationship between work environment and depression among critical care nurses. BMC Nurs 2025 Sep 26;24(1):1201. [doi: 10.1186/s12912-025-03873-5] [Medline: 41013489]
10. Shimabukuro DW, Barton CW, Feldman MD, Mataraso SJ, Das R. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. BMJ Open Respir Res 2017;4(1):e000234. [doi: 10.1136/bmjresp-2017-000234] [Medline: 29435343]
11. Robert N. How artificial intelligence is changing nursing. Nurs Manage 2019;50(9):30-39. [doi: 10.1097/01.NUMA.0000578988.56622.21]
12. Barchielli C, Marullo C, Bonciani M, Vainieri M. Nurses and the acceptance of innovations in technology-intensive contexts: the need for tailored management strategies. BMC Health Serv Res 2021 Jul 3;21(1):639. [doi: 10.1186/s12913-021-06628-5] [Medline: 34215228]
13. Ronquillo CE, Peltonen LM, Pruinelli L, et al. Artificial intelligence in nursing: priorities and opportunities from an international invitational think-tank of the nursing and artificial intelligence leadership collaborative. J Adv Nurs 2021 Sep;77(9):3707-3717. [doi: 10.1111/jan.14855] [Medline: 34003504]
14. Atalla AD, El-Gawad Mousa MA, Hashish EA, Elseesy NA, Abd El Kader Mohamed AI, Sobhi Mohamed SM. Embracing artificial intelligence in nursing: exploring the relationship between artificial intelligence-related attitudes, creative self-efficacy, and clinical reasoning competency among nurses. BMC Nurs 2025 Jun 20;24(1):661. [doi: 10.1186/s12912-025-03306-3] [Medline: 40542340]
15. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on the domains of nursing: scoping review. JMIR Nurs 2020 Dec 17;3(1):e23939. [doi: 10.2196/23939] [Medline: 34406963]
16. Ibrahim AM, Zoromba MA, Abousoliman AD, et al. Ethical implications of artificial intelligence integration in nursing practice in Arab countries: literature review. BMC Nurs 2025 Feb 11;24(1):159. [doi: 10.1186/s12912-025-02798-3] [Medline: 39934834]
17. Alnawafleh KA, Almagharbeh WT, Alfanash HA, et al. Exploring the ethical dimensions of AI integration in nursing practice: a systematic review. J Nurs Regul 2025 Oct;16(3):228-237. [doi: 10.1016/j.jnr.2025.08.001]
18. O'Connor S, Yan Y, Thilo FJ, Felzmann H, Dowding D, Lee JJ. Artificial intelligence in nursing and midwifery: a systematic review. J Clin Nurs 2023 Jul;32(13-14):2951-2968. [doi: 10.1111/jocn.16478] [Medline: 35908207]
19. Ramadan OME, Alruwaili MM, Alruwaili AN, Elsehrawy MG, Alanazi S. Facilitators and barriers to AI adoption in nursing practice: a qualitative study of registered nurses' perspectives. BMC Nurs 2024;23(1). [doi: 10.1186/s12912-024-02571-y]
20. Seibert K, Domhoff D, Bruch D, et al. Application scenarios for artificial intelligence in nursing care: rapid review. J Med Internet Res 2021 Nov 29;23(11):e26522. [doi: 10.2196/26522] [Medline: 34847057]
21. Swed S, Alibrahim H, Elkalagi NK, et al. Knowledge, attitude, and practice of artificial intelligence among doctors and medical students in Syria: a cross-sectional online survey. Front Artif Intell 2022;5:1011524. [doi: 10.3389/frai.2022.1011524] [Medline: 36248622]
22. Schepman A, Rodway P. The General Attitudes towards Artificial Intelligence Scale (GAAIS): confirmatory validation and associations with personality, corporate distrust, and general trust. Int J Hum Comput Interact 2023 Aug 9;39(13):2724-2741. [doi: 10.1080/10447318.2022.2085400]
23. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
24. El-Ashry AM, Al Saleh NS, AlOtaibi NG, et al. The impact of artificial intelligence attitudes and acceptance on critical thinking motivation among nursing students in Saudi Arabia. SAGE Open Nurs 2025;11:23779608251369564. [doi: 10.1177/23779608251369564] [Medline: 40809301]

XSL•FO
RenderX

25. El-Gazar HE, Zoromba MA, Fayed SM, et al. Nurturing success: e-learning readiness and academic self-efficacy in nursing students. BMC Nurs 2024 Jul 19;23(1):495. [doi: 10.1186/s12912-024-02125-2] [Medline: 39030566]
26. Prensky M. Digital natives, digital immigrants part 1. On Horizon 2001 Sep;9(5):1-6. [doi: 10.1108/10748120110424816]
27. Cecconi C, Adams R, Cardone A, et al. Generational differences in healthcare: the role of technology in the path forward. Front Public Health 2025;13:1546317. [doi: 10.3389/fpubh.2025.1546317] [Medline: 40078753]
28. Shim J, Park DI. The influence of gender equity in nursing education programs on nurse job satisfaction. Healthcare (Basel) 2023 May 4;11(9):1318. [doi: 10.3390/healthcare11091318] [Medline: 37174860]
29. Zhang X, Yang L, Bai Y, Zhang L. Postgraduate nursing students' knowledge, attitudes, and practices regarding artificial intelligence: a qualitative study. BMC Med Educ 2025;25(1):1269. [doi: 10.1186/s12909-025-07757-9]
30. Salameh B, Eddy LL, Batran A, Hijaz A, Jaser S. Nurses' attitudes toward the use of an electronic health information system in a developing country. SAGE Open Nurs 2019;5:2377960819843711. [doi: 10.1177/2377960819843711] [Medline: 33415233]
31. Ibrahim S, Donelle L, Regan S, Sidani S. Predicting registered nurses' behavioural intention to use electronic documentation system in home care: application of an adapted unified theory of acceptance and use of technology model. Nurs Leadersh (Tor Ont) 2019 Jun;32(2):48-70. [doi: 10.12927/cjnl.2019.25961] [Medline: 31613214]
32. Dornan M. Every nurse an AI nurse: a framework for integrating artificial intelligence across nursing practice, education, research and policy. Digit Health 2025;11:20552076251377939. [doi: 10.1177/20552076251377939] [Medline: 41036436]

Abbreviations

AI: artificial intelligence
CCN: critical care nurse
GAAIS: General Attitudes Toward Artificial Intelligence Scale
TAM: technology acceptance model

Edited by D Dordunoo, E Borycki; submitted 10.Oct.2025; peer-reviewed by B Siam, CP Tuppal, J Waterson; accepted 19.Dec.2025; published 16.Jan.2026.

Please cite as:
Alrashedi H, Alderaan SM, Alnomasy N, Lamine H, Saleh KA, Alkubati SA
Insights Into Factors Affecting Nurses' Knowledge of and Attitudes Toward AI and Implications for Successful AI Integration in Critical Care: Cross-Sectional Study
JMIR Nursing 2026;9:e85649
URL: https://nursing.jmir.org/2026/1/e85649
doi:10.2196/85649

© Habib Alrashedi, Saad M Alderaan, Nader Alnomasy, Hamdi Lamine, Khalil A Saleh, Sameer A Alkubati. Originally published in JMIR Nursing (https://nursing.jmir.org), 16.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Readiness Assessment for AI in Nursing Care Projects: Multimethods Study

Kathrin Seibert[1], RN, Dr; Dominik Domhoff[1], RN, MA; Janissa Altona[1], MA; Sebastian Jäger[2], MA; Felix Bießmann[2], Prof Dr; Alessia Nowak[3], MA; Rahel Gubser[3], MA; Matthias Schulte-Althoff[3], MA; Daniel Fürstenau[3], Prof Dr; Jörg Pohle[4], Dr; Lea Bergmann[5], MA; David Walter[6], MA; Kathi Beier[7], Dr; Dagmar Borchers[7], Prof Dr; Karin Wolf-Ostermann, Prof Dr

[1]Institute of Public Health and Nursing Research, University of Bremen, Universitätsallee 1B-C, Bremen, Germany

[2]Faculty VI Information Science and Media, Berliner Hochschule für Technik (BHT), Berlin, Germany

[3]Institute of Medical Informatics, Charité Universitätsmedizin, Berlin, Germany

[4]Alexander von Humboldt Institute for Internet and Society, Berlin, Germany

[5]Association for Digitalisation in the Social Sector e.V., Halle (Saale), Germany

[6]Institute of Business Administration & Information Systems, University of Hildesheim, Hildesheim, Germany

[7]Institute for Philosophy, University of Bremen, Bremen, Germany

Corresponding Author:
Kathrin Seibert, RN, Dr
Institute of Public Health and Nursing Research, University of Bremen, Universitätsallee 1B-C, Bremen, Germany

Abstract

Background: Integrating artificial intelligence (AI) systems into nursing care often encounters obstacles stemming from unmet requirements and insufficient engagement with well-documented sociotechnical pitfalls. Readiness models offer a systematic way to evaluate project preparedness and to build the capabilities needed for successful artificial intelligence in nursing care (AINC) research, development, and implementation. As of yet, an evidence-based AI readiness assessment prioritizing AINC projects and accounting for their diversity in care settings is missing.

Objective: This study aimed to develop a comprehensive artificial intelligence nursing care readiness assessment (AINCRA) to support planning, execution, and evaluation of AINC projects.

Methods: In a sequential exploratory multimethods bottom-up approach to maturity model development, key AI readiness dimensions and attributes were identified to develop a pilot readiness assessment. The pilot version was grounded on insights from an expert workshop (n=21) and expert interviews (n=14), an online survey (n=53), a rapid review (n=292), and a nominal group consensus process. A systematic literature review (n=7) further triangulated AI readiness attributes. Finally, a think-aloud interview study and focus group discussions involving experts (n=18) from nursing practice, nursing science, and AI research and development who had conducted AINC projects prior to data collection validated the attributes.

Results: The resulting AINCRA encompasses 5 core dimensions: regulatory, processual, technical, social, ethical, and community building requirements and aspects. Including 69 attributes and capabilities of AI nursing care readiness, the core dimensions reflect key areas of action where AINC project stakeholders can influence project outcomes. Clinical partners can assess their organization's maturity level in relation to the implementation of AI. An assessment of each dimension and its attributes across 5 maturity levels allows reflecting on and proactively shaping individual project approaches. Overall, experts regarded AINCRA as a useful instrument for the development, management, and evaluation of AINC projects while emphasizing that established principles of good practice in project and data management should not be neglected when using AINCRA as a project management tool.

Conclusions: AINCRA enables practitioners from AI research and development, clinical partners, and nursing and health scientists to plan, evaluate, and enhance AI projects across their lifecycle, thereby supporting effective AI integration in nursing care. While AINCRA was developed within the European and German legal framework for AI in health care settings, respective attributes can be adapted to international requirements.

(JMIR Nursing 2026;9:e84148) doi:10.2196/84148

KEYWORDS

artificial intelligence; nursing; care; readiness; maturity model

XSL•FO
RenderX

Introduction

Artificial intelligence (AI) systems promise to contribute to process optimization, workload reduction, patient safety, and quality assurance in nursing care worldwide [1-3]. Following the European Commission's guidelines on AI system definition as entailed in the European Union's Artificial Intelligence (EU AI) Act, AI systems in nursing care refer to machine-based systems developed to function with different degrees of autonomy and potentially adapt over time after being deployed [4]. AI systems process input data to produce outputs such as predictions, content, recommendations, or decisions based on explicit or implicit goals, and these outputs can affect physical or digital environments [4]. Main AI techniques encompass machine learning (ML) approaches including methods such as supervised learning, unsupervised learning, self-supervised learning, and reinforcement learning, or deep learning [4]. Additionally, AI systems include logic- and knowledge-based approaches that infer from encoded human knowledge or symbolic representation of the task to be solved [4]. Logic- and knowledge-based methods encompass areas such as knowledge representation, inductive logic programming, inference engines, expert systems, symbolic reasoning, and search and optimization techniques [4], highlighting diverse methodological possibilities for the development of AI systems in nursing care. Currently, opportunities for AI system development arising from applying generative AI models specifically trained on electronic medical record (or nursing record) data [5] further expand the scope of methods and use cases.

While AI is already supporting clinical decision-making of nurses in wound assessment [6] or diabetes care [7], monitoring and detection of risks and clinical deterioration throughout the care process [8-10], or nursing documentation via hybrid speech assistants [11,12], many unexplored use cases in acute and long-term care persist [13]. Furthermore, studies aiming to assess the effectiveness of AI systems in nursing care under real-world conditions using designs capable of determining cause-and-effect relationships are still scarce [14,15]. The potential benefits and expectations of artificial intelligence in nursing care (AINC) are high and are emphasized by professional associations and the World Health Organization, in light of growing global challenges to ensure high-quality care despite a declining number of health care professionals and increasingly complex disease and care trajectories of patients [16,17]. Funding bodies are investing in AI research and development (R&D), and the number of AINC projects is growing. In Germany, for example, the Federal Ministry of Health and the Federal Ministry of Research, Technology and Space support AINC projects in the areas of new care models, health care research, assistance of nurses and family caregivers, and enhancing autonomy and quality of life of people in need of care.

AINC projects face not only technical and regulatory requirements (eg, aspects of interoperability, compliance with national or international data laws, such as privacy and data protection laws, medical device regulations, or the EU AI Act) but also procedural, ethical, and social challenges of AI development and deployment (eg, the micro-, meso-, and macro-level impact of AI on nursing care, ethical-normative values of nursing and care, and translational factors such as acceptance of AI or added practical value) [14,18,19]. This applies to the planning, implementation, and evaluation phases of AINC projects. Project leaders encounter different care settings with many possible use cases for AI (such as AI systems for care in hospitals, nursing homes, outpatient care, or in initial, further, and continuing education) [2], with each care setting influencing the project process with its own unique organizational logic and culture. AI systems may be more complex to implement and use than many other digital technologies [20]. Organizations that aim to develop, deploy, and use AI systems face technical and human-centered challenges that can be overcome by building AI maturity and ensuring they are well-prepared for AI within their organizational context [20-22]. We propose that the same holds true for AINC projects, for which a systematic, comprehensive reflection of AI readiness attributes across various dimensions can help align, justify, and document actions and strategies across all project stages, and may serve as a motivation to foster communication and collaboration within interprofessional project consortia.

Maturity or, often synonymously used, readiness models provide structured guidance to enhance capabilities for effective (AI) technology integration [20,23,24]. Transferred from insights on organizations, readiness models provide features and guidance that help AINC projects to consider context-specific factors while not overlooking humanistic objectives and aspects of sociotechnical interplay next to focusing on technical and regulatory requirements [20,21]. There is a growing body of both theoretical and empirical research on readiness and maturity models [24]. Among these, stage growth models are particularly prevalent in information systems research [24]. These models typically define a progression of maturity levels, applied to an organization or process, ranging from minimal capability to full maturity, outlining an expected or ideal development path, accompanied by a measurement instrument [23,24]. As an example, in hospitals, the maturity model for ML systems [25] is structured around 3 dimensions: organization, adopter system, and patient data. These are broken down into 12 attributes and assessed across 5 maturity levels [25]. The model addresses aspects such as ML strategy, technical infrastructure, ML expertise, user acceptance, and the quality and standardization of patient data [25]. However, so far, an evidence-based AI readiness assessment prioritizing AINC projects and accounting for their diversity beyond R&D on ML in diverse nursing care settings is missing. Hence, the objective for our research is to develop a comprehensive artificial intelligence nursing care readiness assessment (AINCRA) tool for planning, conducting, and evaluating AINC projects responsibly. Therefore, in a sequential exploratory multimethods study, we ask the following questions:

1. What are dimensions, attributes, and levels of a readiness assessment intended to support those responsible for AINC projects as a reflection tool during the planning, implementation, and evaluation phases of their projects?
2. How do experts with practical experience in planning, implementing, and evaluating AINC projects understand

XSL•FO
RenderX

and use the AINCRA, and how do they assess it with regard to selected evaluation criteria?

We present AINCRA as a tool for decision-makers in R&D and clinical partners in AINC projects to reflect on their handling of prerequisites and aspects of these projects and for designing AINC projects in a promising and successful way. Further, we compare with recommendations for the development of maturity models and common criticisms of maturity models [24] to highlight methodological strengths and weaknesses of AINCRA.

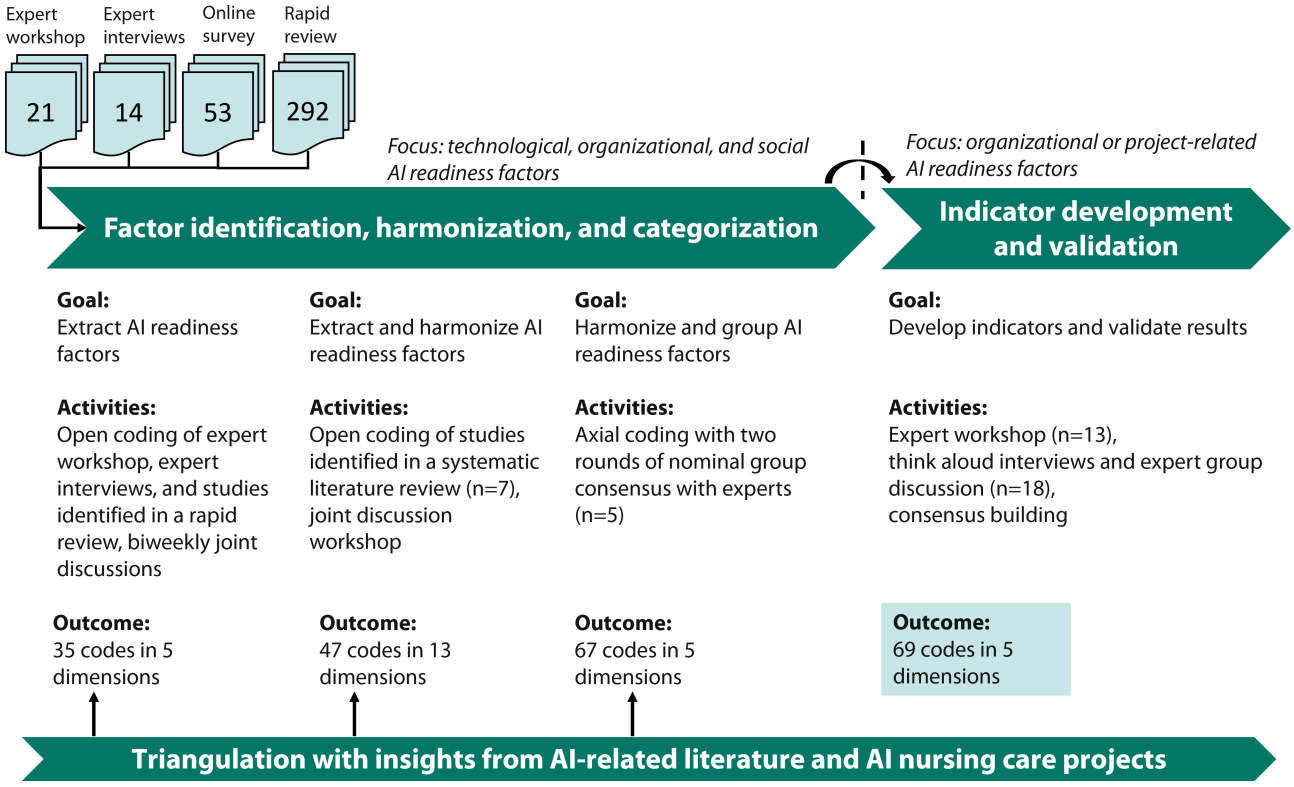
Methods

Overview

We conducted a sequential exploratory multimethods study following established design recommendations and structures of maturity models [20,23,24] to develop AINCRA. Figure 1 presents the research approach, highlighting goals, activities of data collection, analysis, and outcomes for the study period lasting from December 2021 to March 2025. In a stepwise bottom-up approach, we categorized AI readiness factors into capabilities, applying an iterative validation process. Building on preliminary qualitative and quantitative findings from our research group [3,13], we extracted AI readiness factors and conducted a complementary systematic literature review to inform the development of an initial AINCRA version as well as guiding questions for subsequent think-aloud interviews and discussions with experts. Additionally, we triangulated results and findings from the different steps of data collection with insights from AI-related literature and from 8 AINC projects we had been involved in or which we had been accompanying during the last 5 years by repeatedly comparing preliminary drafts of AINCRA with documented project plans to assess completeness.

We follow the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 statement for reporting systematic reviews (Checklist 1) [26]; furthermore, the COREQ (Consolidated Criteria for Reporting Qualitative Research; Checklist 2) [27] and criteria on mixed methods reporting [28] guide this research where applicable. A short description of the rationale, basic methods of development, and a short generalized overview of the AINCRA dimensions has been published elsewhere [29].

Figure 1. Overview of the research approach. AI: artificial intelligence.



Systematic Literature Review: Information Sources, Eligibility Criteria, and Review Protocol

Leading health and information sciences databases Scopus, PubMed, ACM Digital Library, AIS Electronic Library, and EconLit were searched in February 2023 for English- or German-language articles published from 2012 onward which either refer to a specific AI readiness assessment tool (defined by containing at least one specific question or evaluation criterion of AI readiness or maturity with a qualitative assessment option) or refer to frameworks or single studies that theoretically or empirically define AI readiness. Further, articles had to refer to AI readiness of institutions, organizations, projects, teams, or higher-level units (sectors and disciplines). Health care and nursing settings differ from other business sectors and societal areas, among other things, in that they often process highly sensitive personal data and aim to make it usable for AI systems. In addition, they are based on different sociocultural and organizational logics compared to, for example, the agricultural or logistics sector or the timelines and work culture of data science [30]. Therefore, following the assumption that AI readiness models established for other

XSL•FO
RenderX

sectors cannot be transferred directly to the nursing care context, publications had to refer to nursing or health care settings. Eligibility criteria were established before conducting the initial searches, and a review protocol was developed but not registered.

Search Strategy and Selection Process

The search strategy followed a block building approach [31] applied to title, abstract, and keyword searches using the Boolean operator OR within a single block and the Boolean operator AND across blocks to combine the search terms and synonyms for the blocks encompassing "artificial intelligence," "readiness," and "assessment" are depicted in Multimedia Appendix 1. Two independent reviewers (KS and DW) screened titles, abstracts, and full texts according to predefined inclusion and exclusion criteria using the online resource Rayyan [32]. Conflict between reviewers was resolved by including a third reviewer (DD).

Data Collection Process and Data Items

Data extraction was carried out by one reviewer (KS) in a data extraction form created in Microsoft Excel for this purpose. Data items extracted from each included study entail the following: reference (author, year of publication, and title), country in which the study data were collected, objective of study, study design as stated, setting as stated; population: type and sample size (N); definition of AI: presented (yes or no), definition as stated; AI readiness: addressed (yes or no), definition as stated, explicit readiness assessment presented (yes or no), level of application (organization, project, team, and setting), health or nursing care addressed (yes or no), reporting of attribute origin (yes or no), origin of attributes as stated, number of attributes, description or labels of attributes, and labels of dimensions; readiness framework: presented (yes or no), name as stated, empirically tested (yes or no); success factors and challenges: reported (yes or no), success factors and challenges as stated; limitations as stated; ethics vote obtained: yes or no; funding; conflicts of interest. Missing data were coded as not reported.

Study Risk of Bias Assessment, Effect Measures, and Synthesis Methods

As a conclusion on the effectiveness of interventions was not an intended result of the systematic literature review, we did not conduct a risk of bias assessment and other assessments associated with good practice of methodology for meta-analysis. Results were summarized narratively after inductively developing structuring categories for readiness descriptions and dimensions, as well as for readiness attributes. Key findings are described narratively in the results section and supplemented with tables and figures where appropriate.

Think-Aloud Interviews and Group Discussions

Think-aloud interviews and group discussions were conducted online and digitally recorded with OBS Studio (version 29.1.3; OBS Project) in December 2024 and January 2025. Due to individual availability, 3 experts did not participate in a group discussion but answered group discussion guiding questions (Multimedia Appendix 1) in an interview following the think-aloud task. Details about the involved researcher's background and relationship with participants are provided in Multimedia Appendix 1.

Selection and Recruitment of Expert Participants

We applied a purposive sampling approach [33] by sending personalized, electronic invitations to 42 experts who had either been actively involved in AINC bar camps held by our study team in the past or had conducted and published AINC-related R&D in Germany and were known to the study team. A total of 19 of 42 experts replied (response rate of 45.24%), and 17 experts consented to participation. The reason for nonparticipation was unavailability during the data collection phase. Additionally, all project leads of AINC projects in the German funding program "Making Repositories and AI Systems Usable in Everyday Care" were contacted via email and invited to participate or invite members of their project teams to participate. The possibility for participation was also advertised in the newsletter for said funding program, resulting in recruitment of one participant. We also asked the funding body to propose additional experts but did not receive a response.

Ethical Considerations

The ethics committee of the German Society of Nursing Science granted ethical clearance and approved this study (application 22 - 030). Participation was voluntary, and participants received no monetary or immaterial incentives or compensation. After receiving written information including statements on data protection, privacy, and confidentiality as well as concordance with the European General Data Protection Regulation, all participants provided a written informed consent to participation as well as to the pseudonymized analysis of the data obtained and to the publication of anonymized results. The study was not of an interventional nature. Participants were considered a nonvulnerable population, and participation was not classified as of particular risk.

Data Collection

An interview guideline structured and semistandardized the data collection process, which started with a 5-minute introduction of the interviewers, including their academic degree, field of research, and personal involvement in the research project, followed by an overview of the data collection process, an explanation of data protection, consent, and voluntary participation. After that, participants were asked to introduce themselves (name, work setting, and focus) and provide a short statement on their experiences with AINC projects and implementing AINC settings. A 10-minute introduction to the objectives of the study and AINCRA, during which participants were able to ask questions for clarification, made way for the 30-minute think-aloud interview. Participants were assigned to online breakout rooms with 1 or 2 members of the study team who also took field notes. The AINCRA dimension to be discussed was accessible on a Miro Board on which participants could navigate freely. Using the prompt outlined in the "Results" section, participants reflected on an AINC project they either had conducted, were conducting at the time of data collection, or were planning. They were asked to imagine that they are alone in the room and to talk to themselves. All thoughts were allowed to be expressed. If they

XSL•FO
RenderX

were silent for a longer period, the interviewer asked them to continue to speak. Concurrent verbal probing [34] completed the think-aloud interview strategy. After the think-aloud interview, participants were guided into an online meeting room, where one researcher (KS) moderated the 35-minute group discussion structured by the guiding questions (Multimedia Appendix 1). Other researchers took field notes and asked follow-up questions at the end of the discussion but did not intervene. Guiding questions were based on prior national research on maturity model development [35], and refined in consultation with project members. A 5-minute exit phase concluded data collection, during which participants were informed about the availability of results and invited to the final study event. Afterward, researchers held a 15-minute debriefing to discuss key impressions and themes from their field notes.

Qualitative Content Analysis

Applying a deductive qualitative content analysis approach [36], coding was done by 5 members of the study team (KS, JA, LB, AN, and JP) using a coding form developed in Microsoft Excel for this purpose. Deductive content analysis allows for retesting existing data in a new context, involving testing of models [36]. Units of analysis are the participants' verbal contributions during the think-aloud interviews and focus group discussion. Audio files recorded during the think-aloud interview were randomly assigned to one coder, who listened to the recording, transcribed anchor statements, and paraphrased them. A second coder (KS) reviewed all transcripts and codes to validate the first coder's analysis. The structured categorization matrix [36] consisted of all dimensions and attributes of the AINCRA pilot version. For each interviewed person and AINCRA dimension, attribute, and level description, statements were documented as anchor statements. The same was done for statements on the relevance of attributes and suggestions made for revision or changes to AINCRA. The coding framework for the focus group discussion followed the guiding questions, with the overall coding process being the same as for the think-aloud interviews, using an unconstrained categorization matrix, where various categories were formed within the deductive framework based on the principles of inductive content analysis [36].

Mixing of Results and Consensus-Building on Final AINCRA Version, Translation Strategy

Overall, our integration strategy consisted of the following steps: deriving AI readiness attributes and dimensions from prior work and the results of the systematic review. Building on this, an initial AINCRA version was developed, including the definition of attributes for discussion with experts. The results of these discussions were taken into account in the development of the final AINCRA version: in a half-day consensus-building workshop, results from the think-aloud interviews and focus group discussions were reviewed by 6 members of the study team and discussed one last time against overlapping or contrasting themes derived from the systematic literature search. Further, consensus on which changes suggested by the experts should be incorporated into the final AINCRA version was established by discussion, and the final AINCRA version and a user manual were developed in German and English language. As the qualitative data collection was conducted in the German language, results were first compiled in German and then translated into English for publication. Primary translation was performed by the first author (KS) who was the main researcher involved in data collection and data analysis. All members of the study team had access to the German-language results and validated and revised the translated results, with multiple researchers having extensive experience in English-language knowledge transfer and dissemination of research results (DF, FB, KWO, MS, JP).

Results

Overview

Multimedia Appendix 2 presents the final AINCRA and Multimedia Appendix 3 contains the AINCRA user manual. An open-access online version of AINCRA is available [37]. In the following, we outline the iterative research process to contextualize the results contributing to the final AINCRA, before presenting an overview of AINCRA and its recommended application.

Preliminary Database

First, we built on preliminary work from our working group which has been published elsewhere: data from a rapid review [3] including 292 publications on AI systems in nursing care, an expert workshop with 21 experts (including nurses, nursing directors, health care managers from hospitals, nursing homes, and home care services, digitalization officers and professionals with experience in digital health or routine nursing data analysis, informal caregivers, nursing scientists, and researchers in nursing education, computer science, AI, and ethics) as well as interviews with 14 experts and an online survey (n=53) were used to identify success factors and prerequisites of AINC projects [13]. Five initial key dimensions (namely, regulatory requirements, processual and translational requirements and aspects, technical requirements, social and ethical aspects, and community building) entailing 35 codes emerged as a result of qualitative structuring content analysis involving deductive as well as inductive category and code building.

Systematic Literature Review on AI Readiness Models in Health Care and Nursing

Second, we conducted a systematic literature review to identify existing AI readiness models focusing on nursing or health care settings and harmonize AI readiness capabilities, attributes, and dimensions with results from the first step. Screening of 4748 records identified from 5 leading health and information sciences databases resulted in the inclusion of 7 single publications (Figure 2) which reported either at least one question or evaluation criterion of AI readiness with a qualitative assessment option or a framework for theoretically or empirically defining AI readiness in health care or nursing care.

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of identification of studies via databases. AI: artificial intelligence.

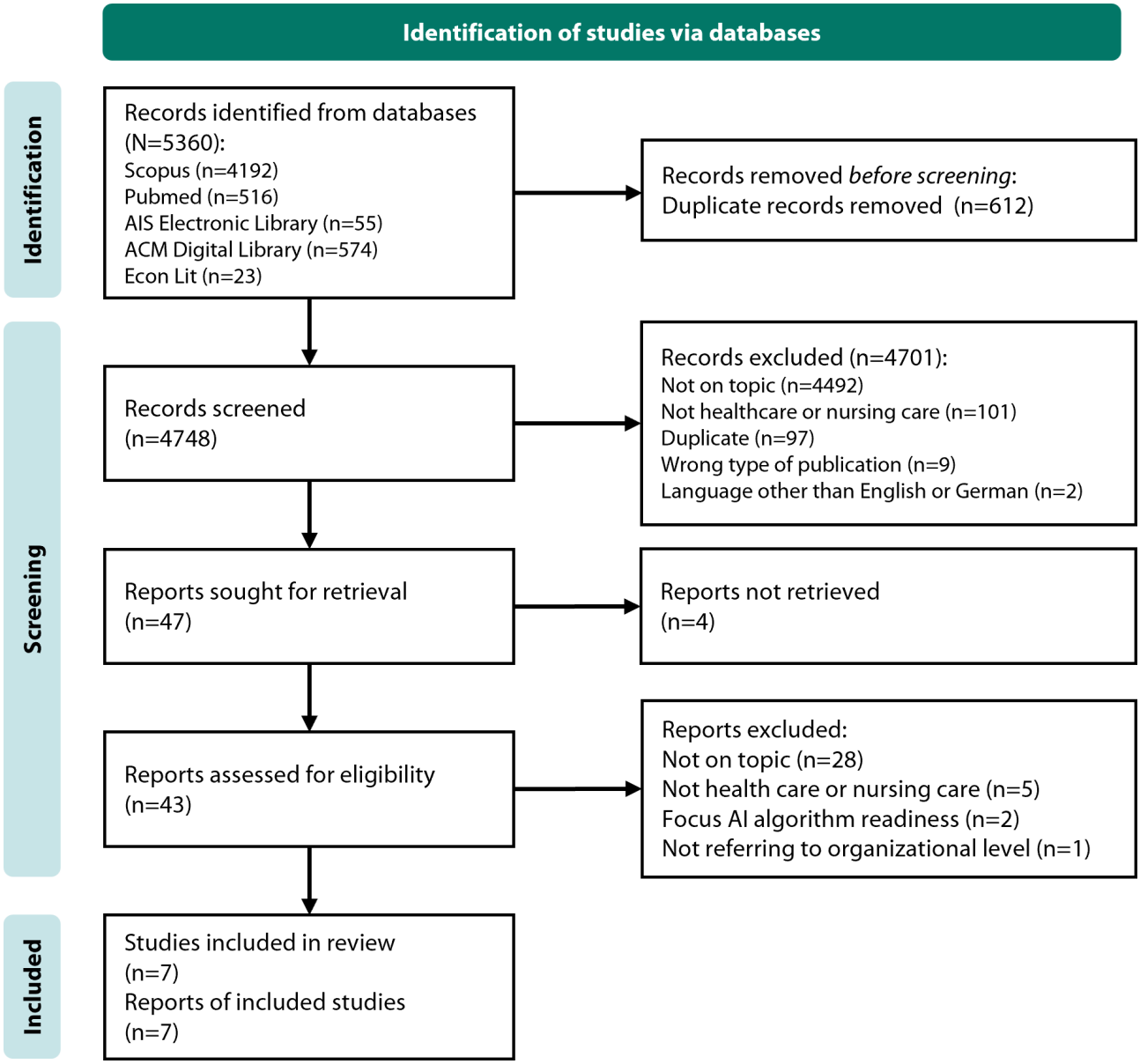


Table 1 provides an overview of the included studies and reports, 3 of which used observational designs with cross-sectional surveys among health care and IT professionals [38-40]. One study applied a qualitative design [25], 1 paper was classified as an expert reflection [41], and 2 papers did not report specific study designs to derive their results from [30,42]. Health care settings focused on either hospitals or selected professional groups or remained unspecified, without referencing specific nursing care settings in acute or long-term care. Overall, we extracted 74 single AI readiness attributes, which we inductively grouped into 47 codes in 13 AI readiness dimensions depicted in Table 2.

XSL•FO
RenderX

Table . Overview of included studies (N=7).

Reference; country of origin	Design	Aim	Population (N)	Reporting of				Number of readiness attributes
				Definition of AI[a]	Definition of AI readiness, explicit perspective	Readiness assessment description	Origin of readiness attributes	
Abuzaid et al 2022 [38]; United Arab Emirates	Observational, cross-sectional survey	Investigate knowledge, perception, readiness, and challenges regarding AI integration into radiology practice	Radiology professionals (n=153)	Y[b]	N[c] Organizational readiness	N	N	3
Alami et al 2020 [41]; International	Expert reflection	Bring forward the importance of studying organizational readiness to integrate AI applications	Not applicable, experiences with AI systems to support clinical decision-making reported	Y	Y Organizational readiness	Y Adapted framework from Jennett et al [43]	Y	Not applicable, readiness dimensions reported
Andersson et al 2021 [39]; Sweden	Observational, cross-sectional survey	Establish a foundation for a Swedish perspective on the potential effect of AI on the medical physics profession	Medical physicists (n=163)	N	N Workplace preparedness	N	Y	2
Chang 2020 [30]; country not reported	Unclear	Not reported	Not applicable	N	N	Y	N	10
Pumplun et al 2021 [25]; Germany and Switzerland	Qualitative	Explore factors that influence the adoption process of ML[d] systems for medical diagnostics in clinics; demonstrate how factors can be used to determine the ML maturity score of clinics	Medical experts and suppliers with profound knowledge in the field of ML (n=22)	Y	N	Y Maturity Model for ML Systems in Clinics	Y	12

XSL•FO
RenderX

Reference; country of origin	Design	Aim	Population (N)	Reporting of				Number of readiness attributes
				Definition of AI[a]	Definition of AI readiness, explicit perspective	Readiness assessment description	Origin of readiness attributes	
Weinert et al 2022 [40]; Germany	Observational, cross-sectional survey	Investigate factors influencing AI readiness as well as possible barriers to AI adoption and implementation in hospitals; assessed the status quo regarding the dissemination of AI tools in hospitals	Hospital chief information officers (n=40)	Y	Y Organizational readiness	Y Model by Jöhnk et al [21]	Y	47
Wiljer and Hakim 2019 [42]; Canada	Unclear	Not reported	Not applicable	Y	N	Y AI-enabled organization	N	Not applicable, readiness dimensions reported

[a]AI: artificial intelligence.

[b]Y: yes.

[c]N: no.

[d]ML: machine learning.

Table . Dimensions of AI[a] readiness derived from the literature (thematic, nonhierarchical grouping).

AI readiness dimension	Number of codes	Derived from
Personnel resources and competencies	12	[25,30,38-42]
Strategic planning	7	[25,30,38,40-42]
Data quality	8	[25,30,40,41]
Financial resources and investments	4	[25,30,40,41]
Acceptance and stakeholder participation	3	[30,39,41]
Technical infrastructure	3	[25,30,40]
Data protection and data safety	2	[30,41]
Intangible assets	2	[30,40]
Leadership culture	2	[25,30]
Needs and problems (in clinical practice and of patients)	1	[41]
Practical benefit and added value	1	[41]
Time resources	1	[40]
Organizational culture	1	[30]

[a]AI: artificial intelligence.

Axial coding and contrasting of these 74 readiness attributes with the 35 codes extracted in step 1 provided the database for 2 rounds of nominal group consensus involving 5 experts (2 nursing scientists, 1 ML expert, 1 information systems specialist, and 1 expert representing the German association for digitalization in the social economy with experience in consulting of clinical partners) to harmonize and group AI readiness factors. By merging semantically equivalent attributes and removing those not directly influenced by AINC stakeholders, the AI readiness attributes were reduced to 67 codes in 5 dimensions, serving as the baseline for the development of a pilot AINCRA version.

XSL•FO
RenderX

Indicator Development and Expert Validation

The pilot AINCRA version included feedback from 13 AINC project experts to establish a common AI readiness vocabulary for level distinctions. Building on vocabularies of existing maturity or readiness models [23,44], expert feedback was obtained from researchers and practitioners working in ongoing AINC projects at the time of data collection during a 45-minute bar camp session in September 2024 which was open for participation for all people currently used in 1 of 9 AINC projects in the German Ministry for Research, Technology and Space's funding program "Making Repositories and AI Systems Usable in Everyday Care." Further, level labels, frequencies, and numbering reported in the literature [23,25] were discussed with these experts, resulting in the decision to adopt 5 levels (1: initial, 2: assessing, 3: determined, 4: managed, and 5: optimized) with the possibility to rate an attribute as not applicable depending on the type of AINC project, AI system, or organizational features of clinical partners. The complete AI readiness vocabulary for level distinction is included in Multimedia Appendix 3. The use of the vocabulary and the 67 attributes derived from the third step of the research process as input for GPT-4o (OpenAI)–generated 335 level descriptions, which, after expert review and revision by all members of the study team, were incorporated into the AINCRA pilot version.

Finally, to evaluate the AINCRA pilot version, we applied a think-aloud interview approach [45,46] including a concurrent think-aloud session and reflexive group discussion with 18 experts with a proven track record in carrying out AINC projects to refine AINCRA dimensions, attributes, level distinctions, and descriptions. The think-aloud approach can be useful to "better understand thought processes during assessments as a strategy to describe what [an] assessment is measuring" [46]. Further, the think-aloud approach provided information about the cognitive thought processes of AINC project stakeholders using AINCRA "pertaining to usability problems they would encounter" [47] to reflect on a specific AINC project. Considering possible ways to evaluate maturity models, we followed good-practice recommendations by involving experts "on the type of process that is intended to be improved by the maturity model, but who have not been involved in the actual development of the maturity model" [48].

Characteristics of Think-Aloud Study Participants

At the time of data collection, 10 experts are involved in AINC projects outside of the German funding program "Making Repositories and AI Systems Usable in Everyday Care," while 8 experts are working in ongoing AINC projects in said funding program (characteristics of participants are depicted in Multimedia Appendix 1). Eight experts report less than 5 years of experience with AINC projects, while 5 experts indicate experience of 5 years or longer. These 5, as well as 5 experts who did not report conclusively on the duration of their experience, had also been involved in projects focusing on developing and implementing other digital technologies in nursing care, digitalization, and advancement of IT infrastructures in hospitals and long-term care for several years. Considering the main AINC project stakeholder groups regarding the composition of project consortia, 9 experts represent a nursing or health science perspective, often accompanied by responsibilities and roles connected to project coordination, while in some cases also holding topic-related qualifications or work experience such as nursing informatics expertise. Five experts represent clinical partners in AINC projects, and 5 experts represent AI R&D, 2 of them working for tech companies compared to 3 experts working for institutes affiliated with universities or independent research institutions. In the think-aloud interview, 13 experts reflect on an ongoing AINC project, 4 experts reflect on a completed project, and 1 expert reflects on a project in planning.

Think-Aloud Interview and Expert Group Discussion

Individual 30-minute think-aloud interviews and subsequent 35-minute group discussions were conducted online by 1 to 3 members of the study team in an overall 90-minute session. All participants received an overview of all AINCRA pilot version dimensions and attributes in advance. However, detailed level descriptions were only presented during the interview, along with the verbal prompt to evaluate an ongoing, in planning, or past AINC project they were familiar with. They were asked to do so by using the attributes and level descriptions of one or more selected AINCRA dimensions, referring to a recalled point in time (eg, planning, implementation, or evaluation phase), while verbalizing all their thoughts aloud. Since we anticipated the pilot version to be a complex and time-consuming assessment, participants were only presented with selected dimensions, never all of them. The selection was made with knowledge of the participants' areas of expertise, with a slight emphasis on process-related requirements. This focus was chosen because these dimensions included attributes that are described in the literature as being often neglected in AI maturity models [20], and which, according to our preliminary study, are known to significantly influence project flow and success [13].

We obtained expert statements and comments on 64 of 67 attributes with a total of 70 suggestions for changes to the AINCRA pilot version (Multimedia Appendix 1). Changes suggested included adding definitions, or explanations, or examples for selected terms (eg, "data champion" and "ontological representation"), simplification of terms, reordering and grouping of attributes within 1 dimension or across 2 dimensions, splitting of attributes, and adding of examples in level descriptions (eg, sponsor role as a required responsibility when compliance with the EU Medical Device Regulation is mandatory for the AI system). Some comments highlighted the need for adding information to the AINCRA user manual without resulting in changes to dimensions, attributes, or level descriptions. Overall, 37 of 70 (53%) changes suggested were realized in the final AINCRA version.

In the group discussions, experts rated and discussed the AINCRA dimensions, attributes, and level descriptions they had applied to their own AINC projects in the think-aloud interview. An overview of the discussed evaluation criteria with their respective guiding questions for the discussion can be found in the "Methods" section. In summary, AINCRA was assessed as a valuable tool for AINC project development, management, and evaluation by the experts. However, some

XSL•FO
RenderX

experts pointed out that not all influencing aspects of AINC projects can be foreseen using AINCRA, and that a project process can also be set up without applying AINCRA. Further, general aspects of good practice of project or data management that still need to be considered in each project should not be overlooked when focusing on AINCRA as a project management tool. Multimedia Appendix 1 displays expert ratings and feedback for each evaluation criteria. Results for each evaluation criteria are summarized below and highlighted with selected expert statements.

Benefits and Consequences of Applying AINCRA

Considering benefits and consequences of applying the AINCRA to AINC projects, experts generally see added value. Advantages of applying the complex AINCRA are seen by some as being more suited to larger research projects rather than smaller initiatives or in-house projects, while others also recognize its value for those smaller endeavors. Added value was also noted for project consortia with little prior experience in conducting AINC projects. For clinical partners, AI R&D, and nursing or health science research partners in AINC projects, AINCRA is seen as useful for decision-making, planning, grant applications, coordination among partners, identifying resources and expertise, aligning goals, managing projects, and conducting evaluations. Experts also pointed out AINCRA's usefulness for AINC project funding bodies, as AINCRA may support documentation, monitoring, evaluation, and comparison of AINC projects. Expert participants representing tech companies pointed out that AINCRA may help with their project management activities, client consulting, and long-term support of AI implementations.

> *I think another advantage is that with [AINCRA], you can already identify project risks [...] during project planning, for example the area of social and ethical success factors [...] which is very helpful if such a dimension is already represented here.* [IP2, 00:48:18]

Results of AINCRA may lead to various consequences for AINC projects: positive consequences entail advantages in grant applications for high-scoring applicants, identification of improvement areas and necessary changes, support for go or no-go AINC project decisions, and early reflection on and evaluation of an AI system's practical value. However, an AINCRA result may evoke emotional effects in the individuals conducting the assessment, influencing their motivation to engage with AINC readiness attributes or to push an AINC project forward. Low AINCRA scores may cause frustration or disappointment but can also spark motivation and ambition to improve. However, low-level placement may demotivate clinical partners or discourage participation if initial AINC readiness appears too low and difficult to improve.

> *Especially if the requirements are not met at the beginning, it can quickly give the impression that one doesn't want to do it and doesn't see the benefit for the effort that has to be put in, even though there actually is one.* [IP14, 00:54:16]

Rater Entity

When asked who should carry out the assessment, experts agree that AINCRA is unlikely to be effectively carried out by a single individual. Depending on the dimension and attribute, different individuals should be involved, or AINCRA should be conducted collaboratively, either in a group or in a tandem (eg, between clinical and AI R&D partner) within an interprofessional team. For assessing processual and translational requirements and aspects, suitable raters include experienced personnel from the clinical partner with nursing and digitalization expertise, and IT staff of the clinical partner in collaboration with nursing management with a focus on joint assessments by AI R&D and clinical partners. Conducting AINCRA as a guided, participatory assessment involving frontline nurses was stated as a suitable approach. However, some experts caution against including certain roles (eg, nursing assistants) and emphasize that nurses should be supported during their involvement. Experts also raise the question of who holds responsibility for the AINCRA result, especially if the result impacts, for example, project funding decisions.

> *I think it would be good if the evaluation is done at the start by the project consortium, not by the individual specialists [...] it would be good if everyone in the consortium, with its different disciplines, is clear about [...] all five dimensions [...] so that together they have a clearer idea of what lies ahead.* [IP3, 00:52:05]

Attainability of Levels

Assessing the attainability of AINCRA levels, overall, experts view the levels as realistic and reflective of the diversity among clinical partners and AINC projects. However, not all levels appear to be achievable. They emphasize that this should not necessarily be the expectation for AINC projects, as attainability of levels depends on the dimension assessed and the project context. Level 5 is ambitious and currently hard to reach for most clinical partners. But experts emphasize that level 5 should be a long-term goal and is seen as achievable, especially for clinical partners with extensive AINC project experience, but is considered unrealistic for R&D projects. Level 4 is regarded as the upper realistic limit for R&D projects and a desirable target for clinical partners. Level 3 is also seen as a satisfactory goal for clinical partners. While level 2 was not discussed in detail, for level 1, experts point out that some clinical partners already start AINC projects while scoring above level 1.

> *I would definitely say that we have developed. There were some levels that, from our perspective, were not achievable. But the question is also whether that's the goal. Maybe it's enough to say that you get through with a 3.* [IP11, 00:51:18]

Achievability of Levels

Reflecting on past AINC projects that, from the experts' point of view, were successful, experts discussed the achievement of the AINCRA levels regarding the level at the start of a project and the evolution of levels throughout the project process until completion. Overall, the achievement of AINCRA levels is rated as depending on individual goals and structures of an

XSL•FO
RenderX

AINC project. Progression is inherent to AINC projects, as no project meets all requirements from the start. Participation in an AINC project can lead to the development of new structures and processes at the clinical partner organization that go beyond AINCRA attributes. In contrast, some AINC projects may not show improvement despite great efforts to advance AINC readiness attributes. All types of development trajectories are possible, with advancement, remaining, or regression of levels. Entry levels and basic conditions may help industry (or R&D) partners to more easily identify suitable clinical partners, with level 1 being rated not as a general exclusion criterion, but a knockout criterion for certain attributes (eg, data availability).

> *I do believe that there can be deal-breaker criteria in attributes. Off the top of my head, I thought of data quality and data availability in our project. Especially structural things – if those aren't in place, we don't need to start an AI project. If my data are just scattered on paper.* [IP13, 00:53:24]

Step Size of Levels and Influenceability of Attributes

Step size of AINCRA levels is rated as sufficient, appropriate, and distinct with little in-depth discussion. Influenceability of the AINCRA attributes was discussed in depth: most attributes are considered influenceable within AINC projects. However, some less influenceable attributes are still important, such as regulatory requirements or legal framework changes, which are typically beyond a project's control. Experts highlight influenceability as a reflection criterion, especially when a project receives a low maturity level rating. Attributes considered particularly influenceable are related to attitudes, acceptance, knowledge, and willingness to change of stakeholders at the clinical partners' organization. Limited influenceability is assigned to attributes requiring fundamental corporate decisions (eg, developing and committing to an internal AI strategy) as well as attributes dependent on external conditions of clinical partners (eg, staff shortages, turnover, and financial resources).

> *[AINCRA attributes are] all topics that I can address within the project.* [IP11, 00:55:39]

Availability of Information, Comprehensibility, and Applicability

Discussion of the availability of the information or data needed to carry out the AINCRA revealed that the required information is rated as available or obtainable with reasonable effort for AINC project stakeholders by the experts. Involving additional stakeholders is emphasized, supporting the previous assessment that AINCRA should be conducted as an interprofessional assessment. However, the availability of additional stakeholders may be limited, for example, if IT staff of clinical partners have competing responsibilities and carrying out AINCRA is considered a lower priority compared to other tasks.

> *All [attributes] I have seen so far [.] are definitely such that you can either answer them yourself or at least know who to turn to in order to get the information.* [IP12, 01:14:26]

Regarding the comprehensibility of the AINCRA, some experts consider the instrument to be low-threshold and emphasize its comprehensibility for AINC project leaders. However, some wording is not familiar to all nursing professionals. There is also concern about whether nonnative speakers in nursing leadership or administrative roles will understand the expressions and terminology used in AINCRA. Referring to statements made in the think-aloud interview, experts provide suggestions for logically ordering, rephrasing, or explaining selected attributes, as details of the attributes and level descriptions were rated as not quickly to be grasped.

> *So, I would assess it as very understandable [...] of course, during the [think aloud interview], I naturally picked out the things I understand in order to be able to answer them [...] the comprehensibility depends on who from the consortium is conducting the assessment [...] regarding nursing topics, the IT partners probably would have said, 'I don't understand anything about that at all' [...] it depends on who is sitting in front of it and which professional domain it is.* [IP2, 01:03:33]

For the applicability of AINCRA, experts assess the AINCRA pilot version as a complex, text-heavy instrument that, when applied thoroughly, can take several hours to complete depending on the dimension rated. This long application duration may reduce acceptance of AINCRA. Initial use in particular requires time to understand the instrument and is not necessarily intuitive. However, experts expect that the time required will decrease over time and that the application will become easier with continued use. Conducting AINCRA in an interprofessional team enhances its applicability. Some experts attest to the very good usability of AINCRA, though it should be noted that no expert applied the entire instrument. Application seems challenging for stakeholders in long-term care facilities due to lack of expertise in certain attributes. Experts discussed whether a shortened AINCRA version could be provided.

Comparability and Completeness

Comparability of AINCRA results between AINC projects is assessed heterogeneously between the experts, with some pointing out that comparability is difficult, and others emphasizing that comparability of individual dimensions and attributes is possible. Further, multiple practice partners within one AINC project may be compared with AINCRA, as well as multiple AINC projects within the same funding program, and comparability can be ensured through external evaluation.

> *Of course, it's a self-assessment tool [...] the question is whether it's answered honestly and whether that allows for comparability. If it's evaluated in a review process by external parties, then I do believe that good comparability can be achieved.* [IP8, 00:21:20]

Completeness of AINCRA overall is rated as given, with some experts highlighting suggestions for additions or change. These address the inclusion of staff representatives of clinical partners as important stakeholders, allowing for custom attributes to be added by users, and enabling mapping of multiple clinical partners within a single AINC project.

XSL•FO
RenderX

Overview of AINCRA Dimensions, Attributes, and Levels

AINCRA consists of 69 AINC readiness attributes across five dimensions: (1) regulatory requirements and aspects (9 attributes), (2) processual and translational requirements and aspects (40 attributes), (3) technical requirements and aspects (6 attributes), (4) ethical and social requirements and aspects (11 attributes), and (5) community building requirements and aspects (3 attributes). Table 3 provides an overview of the 5 dimensions. For each attribute, 5 maturity levels are described, a recommendation for stakeholders considered as most suitable as assessors is given, along with a reference to the source from which the attribute or the justification for the relevance of the attribute is derived (Multimedia Appendix 2).

Table . Overview of AINCRA[a] dimensions.

Dimension	Description
Regulatory requirements and aspects (9 attributes)	• This dimension takes into account the necessary legal frameworks and data protection regulations that require an analysis of the data and an examination of data-sharing models to enable the use of AI[b] systems in everyday nursing practice.
Processual and translational requirements and aspects (40 attributes)	• This dimension captures the human, material, temporal, and intangible resources available for AI in nursing care projects in care facilities or hospitals, as well as their general and AI-specific level of digitalization. • It also assesses the maturity of overarching strategies for AI, data governance, and IT governance, as well as plans for long-term external support and evaluation of AI implementation. • Engagement with the implementation of needs-based research and development of AI systems—focused on practical benefits and added value—as well as the existing AI-related knowledge and competencies within care facilities and clinics, is intended to support the seamless integration of AI systems into existing care processes. • This dimension also includes reflection on attitudes toward AI, acceptance of AI technologies, and trust in their use in daily nursing practice, to foster adoption and address concerns of those affected by the deployment of AI.
Technical requirements and aspects (6 attributes)	• This dimension addresses the availability and functionality of the technical infrastructure, including aspects of data protection and data security, which are essential for reliable data use and analysis. • To ensure smooth, secure, and efficient data integration and exchange between different systems and stakeholders in everyday nursing practice, this dimension also considers the integration of AI systems into data infrastructures or data platforms, as well as the use of technical interoperability standards and nomenclatures.
Social and ethical requirements and aspects (11 attributes)	• This dimension encompasses ethical and methodological considerations related to voluntariness, privacy, fairness, and transparency, as well as the ethical and normative value orientations of the field and individual practice partners. • Engaging with the impacts of AI use at the micro, meso, and macro levels, along with a responsible approach to data handling, is intended to support the use of AI systems in everyday nursing practice.
Community building requirements and aspects (3 attributes)	• This dimension aims to promote a network that strengthens knowledge exchange and the collective advancement of AI systems in everyday nursing practice. Researchers, developers, and stakeholders from nursing practice and nursing management are involved in this network.

[a]AINCRA: artificial intelligence nursing care readiness assessment.

[b]AI: artificial intelligence.

AINCRA dimension 1, regulatory requirements and aspects, involves the assessment of how detailed and specific methodological and regulatory decisions, definitions, and limitations are being addressed by an AINC project. This includes, among other things, attributes reported by Pumplun et al [25] for analyzing the dataset, as well as the operationalization of data-sharing models [30,41,49]. In the latter case, for example, the degree of finalization and the progress of the implementation of the model serve as reflection criteria. This dimension also relates to the engagement with requirements arising from the EU AI Act. AINC project stakeholders recommended for assessing dimension 1 are mainly

XSL•FO
RenderX

individuals representing AI R&D, supplemented by a clinical partners' and nursing science perspective for some attributes. An example of the attributes 1.6 Data-sharing Models and 1.9 EU AI Act, corresponding level descriptions, recommended rater entity, and sources of attribute origin is shown in Table 4.

AINCRA dimension 2, processual and translational aspects and requirements, evaluates the resources, general as well as AI-specific digital readiness, and strategic planning in place for implementing AINC. It considers infrastructure, staff skills, governance, and external support, as well as how well AI aligns with care needs and existing workflows. It also examines staff attitudes, acceptance, and trust in AI to support successful integration into everyday practice. Most attributes are targeted at the clinical partner, whose expertise is crucial for assessing this dimension and which needs to be supplemented by nursing science and AI R&D expertise to arrive at a comprehensive assessment. This is reflected, for example, in attributes targeting the degree of digitization, which includes the use of data quality standards [25], the availability of data champions [30] as a personnel resource, or the systematic demonstration of practical benefit and added value of an AI system (Table 4).

AINCRA dimension 3, technical requirements and aspects, focuses on the technical infrastructure needed for reliable and secure data use in AINC projects. It includes data protection, system integration, and the use of interoperability standards to ensure smooth data exchange and effective AI implementation in clinical practice. The main AINC project stakeholder to assess dimension 3 is AI R&D. While the technical infrastructure at clinical partner sites is an essential and priorly described prerequisite for AINC project success [30,41,49], we recommend especially assessing IT infrastructure with a focus on AI compute capabilities and dedicated hardware, for which requirements are often underestimated or not clearly communicated between AI R&D and clinical partners during initial AINC project planning (Table 4).

AINCRA dimension 4, ethical and social requirements and aspects, covers ethical and methodological aspects such as voluntariness, privacy, fairness, and transparency. It also considers the values of the nursing field and its practitioners, as well as the broader impacts of AI use and responsible data handling to support ethical AI integration in daily nursing care. Dimension 4 should be assessed from a multifaceted perspective, including nursing science and clinical partners. In collaboration with clinical partners and AI R&D, nursing scientists should drive the assessment of whether and how structured engagement with ethical-normative values of nursing and care [49] is operationalized in an AINC project. The same holds true for reflecting on the impact of AI on the nursing profession [49], and how this impact informs AI system development, implementation, and evaluation (Table 4).

XSL•FO
RenderX

Table . Examples of attributes and level descriptions across the AINCRA[a] dimensions.

Dimension and attribute number	Attribute	Level 1 (initial)	Level 2 (assessing)	Level 3 (determined)	Level 4 (managed)	Level 5 (optimized)	Not applicable	(Joint) Assessment by (rater entity)	Source
Regulatory requirements and aspects									
1.6	Data-sharing models	• Data-sharing models relevant to the AINC[b] project are unknown or unclear. • No considerations regarding appropriate models and approaches exist.	• Initial considerations and steps toward defining and implementing a data-sharing model are being taken. • However, the model is not finalized, and necessary people and procedures are not yet fully clarified.	• The data-sharing model and approach are defined, and most legal, organizational, and technical requirements are identified. • Contact persons and responsibilities are known, but detailed negotiation and implementation have not started.	• The data-sharing model and approach are defined and known to all project participants. • All legal, organizational, and technical requirements are fully identified and documented in a structured manner. • Responsible contacts are involved and took part in detailed negotiations. • Implementation includes appropriate regulatory, organizational, and technical safeguards	• The model and processes of data sharing have been implemented with all necessary regulatory, organizational, and technical safeguards. • Fulfillment of requirements is proven (eg, through internal and external certification). • Data sharing and usage are fully documented, and documentation is regularly reviewed. • Processes and structures evolve based on audit results.	• Not applicable	• AI[c] R&D[d]	[30,41,49]

XSL•FO
RenderX

Dimension and attribute number	Attribute	Level 1 (initial)	Level 2 (assessing)	Level 3 (determined)	Level 4 (managed)	Level 5 (optimized)	Not applicable	(Joint) Assessment by (rater entity)	Source
1.9	EU[e] AI Act	• No consideration of the EU AI Act in the AINC project. • No awareness of which parts of the project are affected.	• Initial steps to comply with the EU AI Act are underway. • A basic overview exists, but the requirements are inconsistently implemented. • Responsibilities and external actors are not fully clarified.	• The AINC project is committed to compliance with the EU AI Act. • A broad overview exists, responsibilities are defined. • Many requirements are implemented or planned, but coverage and documentation are not complete.	• The AINC project is committed to AI Act compliance and has a structured overview of legal requirements. • Responsibilities are clear and implementation follows a defined process standard. • All requirements are planned or fulfilled with explicit deadlines and sufficient resources (time, personnel, and other). • Implementation is well-documented, tested, and regularly monitored.		• Not applicable	• AI R&D • Nursing Science • Clinical partner	Emerged from this research

XSL•FO
RenderX

Dimension and attribute number	Attribute	Level 1 (initial)	Level 2 (assessing)	Level 3 (determined)	Level 4 (managed)	Level 5 (optimized)	Not applicable	(Joint) Assessment by (rater entity)	Source
						• Full compliance with the EU AI Act is ensured. A complete and structured legal framework exists and is shared with all participants. • Responsibilities are clear, and implementation follows international standards (eg, ISO/IEC 42001). • All requirements are planned or fulfilled with explicit deadlines and sufficient resources (time, personnel, other), documented, tested, externally certified, and continuously			

XSL•FO
RenderX

Dimension and attribute number	Attribute	Level 1 (initial)	Level 2 (assessing)	Level 3 (determined)	Level 4 (managed)	Level 5 (optimized)	Not applicable	(Joint) Assessment by (rater entity)	Source
						improved.			
Processual and translational aspects and requirements									

XSL•FO
RenderX

Dimension and attribute number	Attribute	Level 1 (initial)	Level 2 (assessing)	Level 3 (determined)	Level 4 (managed)	Level 5 (optimized)	Not applicable	(Joint) Assessment by (rater entity)	Source
2.9	Clinical partner: personnel resources: available data champions[f]	• There are no data champions (at the clinical partner) in the AINC project, and the use of data is not actively promoted or supported.	• First data champions (at the clinical partner) are being identified, but their role is not yet formalized or recognized.	• Some data champions (at the clinical partner) are active, their role is defined, and they promote a data culture in selected areas.	• Data champions (at the clinical partner) are established and work across departments to promote and support data use.	• Data champions are fully integrated into the organizational structure (of the clinical partner) and actively drive a data-driven culture forward.	Not applicable	• Clinical partner with IT staff • AI R&D	[30]
2.32	• Practical benefit and added value of the AI system	• The practical benefit or added value of the AI system is unknown, or no clear practical benefit or added value is evident. • The AINC project is predominantly theoretical.	• The potential benefit of the AI system is recognized, but concrete indicators or criteria for capturing the benefit and added value are unclear. • Initial steps to evaluate the added value are being taken but are not consistently recorded and documented.	• The AI system shows a recognizable practical benefit, which is represented by concrete indicators or criteria. • However, end points are only partially captured, or the benefit is only partially realized.	• The practical benefit of the AI system is clearly defined and is systematically implemented and demonstrated within the AINC project.	• The AI system generates significant practical benefit and added value, which is continuously measured and optimized using clearly defined criteria and indicators.	Not applicable	• AI R&D • Nursing Science • Clinical Partner	[30,41]
Technical requirements and aspects									

XSL•FO
RenderX

Dimension and attribute number	Attribute	Level 1 (initial)	Level 2 (assessing)	Level 3 (determined)	Level 4 (managed)	Level 5 (optimized)	Not applicable	(Joint) Assessment by (rater entity)	Source
3.2	Use of technical interoperability standards and nomenclatures	• No use of interoperability standards or nomenclatures.	• Initial steps toward using and considering interoperability standards and nomenclatures in the development of the AI system within the AINC project, but not yet fully implemented.	• Interoperability standards and nomenclatures are partially considered and used in the development of the AI system within the AINC project, but not consistently.	• Extensive use and integration of interoperability standards and nomenclatures in the development of the AI system within the AINC project.	• Complete and optimized use of interoperability standards and nomenclatures in the development of the AI system within the AINC project, regularly updated and applied across all relevant areas.	Not applicable	• AI R&D • Nursing Science	[30,41,49]
3.6	IT infrastructure: AI compute: hardware	• No dedicated AI hardware available.	• Dedicated AI hardware is available but not yet connected, eg, to data interfaces.	• Dedicated AI hardware is available and connected (eg, to data interfaces). It is possible to transfer data into the runtime environment for model development.	• Dedicated AI hardware is available and connected. Standardized interfaces are in place for data transfer and for offering model predictions. • Continuous testing and maintenance of the AI system is possible.	• Fully optimized and up-to-date dedicated (hardware and software) runtime environments are available. • Model development and maintenance are implemented according to modern CI or CD components.	Not applicable	• AI R&D	Emerged from this research

XSL•FO
RenderX

Dimension and attribute number	Attribute	Level 1 (initial)	Level 2 (assessing)	Level 3 (determined)	Level 4 (managed)	Level 5 (optimized)	Not applicable	(Joint) Assessment by (rater entity)	Source
Social and ethical requirements and aspects									
4.4	Engagement with ethical-normative values of nursing and care, and individual clinical partners	• No engagement with ethical values (has taken place) in the AINC project.	• Initial ideas on how to engage with ethical-normative values in the AINC project exist but are not concrete or systematic.	• Methods to engage with ethical-normative values are defined but are only partially implemented or exclude some stakeholders.	• Regular, systematic, and documented engagement with ethical-normative values including all stakeholders is part of the AINC project.	• Regular, systematic, and documented engagement with ethical-normative values including all stakeholders is part of the AINC project, and is adapted as needed.	Not applicable	• AI R&D • Nursing Science • Clinical Partner	Emerged from our prior research [49]
4.6	Reflection on the impact of AI on the nursing profession	• The general influence of AI on nursing is acknowledged, but there is no project-level reflection.	• Initial reflections on the impact of AI on the nursing profession are present. • However, reflections are unsystematic and without an effect on system development, implementation, or evaluation.	• Systematic reflection on the impact of AI on the nursing profession is part of the AINC project but lacks influence on development, implementation, or evaluation.	• Comprehensive reflection on the impact of AI on the nursing profession that informs system development, implementation, and evaluation is part of the AINC project.	• Fully integrated reflection on the impact of AI on the nursing profession is included in the AINC project results and documentation.	Not applicable	• AI R&D • Nursing Science • Clinical Partner	Emerged from our prior research [49]
Community building requirements and aspects									

XSL•FO
RenderX

Dimension and attribute number	Attribute	Level 1 (initial)	Level 2 (assessing)	Level 3 (determined)	Level 4 (managed)	Level 5 (optimized)	Not applicable	(Joint) Assessment by (rater entity)	Source
5.3	Strategic partnerships	• Existing strategic partnerships are unknown or there are no strategic partnerships in the AINC project or among individual project partners.	• Initial strategic partnerships are being explored, but have not yet been fully concluded or formalized.	• Strategic partnerships are partially established, but not all partners in the AINC project have strategic partnerships. • Clinical partners in particular do not have strategic partnerships outside of the project consortium.	• Comprehensive strategic partnerships are established. • Clinical partners also have strategic partnerships outside of the project consortium. • Existing strategic partnerships are regularly reviewed and expanded, especially on a national level.	• Proven and effective strategic partnerships are established for all partners in the AINC project. • Existing strategic partnerships are continuously and internationally expanded.	Not applicable	• AI R&D • Nursing • Science • Clinical partner	[42]

[a]AINCRA: artificial intelligence nursing care readiness assessment.

[b]AINC: artificial intelligence in nursing care.

[c]AI: artificial intelligence.

[d]R&D: research and development.

[e]EU: European Union.

[f]Data Champions understand the types of data generated at the clinical partner, advocate for proper data handling, and mediate between nursing staff and IT. They ensure data are correct, complete, and up to date, identify and resolve data collection issues, enhance staff data literacy through training, and raise awareness about the importance of data in nursing care and research.

AINCRA dimension 5, community building requirements and aspects, promotes a collaborative network to support knowledge exchange and the shared development of AI systems for nursing care, involving researchers, developers, and professionals from nursing practice and management, ideally reaching beyond single-project boundaries and striving for national and international networks and shared knowledge. All AINC project stakeholders can assess dimension 5 with AI R&D and nursing science, especially reflecting on their involvement in technical knowledge transfer [49], while clinical partners should also reflect on existing strategic partnerships and how to foster these [42] (Table 4).

Applying AINCRA

AINCRA is primarily designed as a self-assessment tool for AINC project partners. It addresses topics such as the use and benefit of AINC, data representativeness and sharing, participatory design, and ethical and professional implications of AI implementation. AINCRA can be used at any project stage: a single assessment provides a snapshot of current AI readiness, while repeated assessments enable monitoring progress over time. It may also serve as an external evaluation tool. In all cases, AINCRA should be applied by an interprofessional team, ideally involving representatives from

XSL•FO
RenderX

AI R&D, nursing science, and clinical partner organizations. The tool can be used for an entire AINC project or for selected dimensions or sub-areas.

Interpreting AINCRA

In the absence of an internationally agreed consensus on which factors most strongly determine the success of AINC projects, AINCRA does not produce an overall score. Instead, it prompts users to reflect systematically on all dimensions and attributes, which are consistently considered important for project success. AINCRA results provide a snapshot of an AINC project's current readiness with regard to known prerequisites for successful AI development and implementation in nursing care. The tool can be used in context analysis, project planning, and formative evaluation to identify areas for further development or prioritization, and in summative evaluation to reflect on factors that influenced project outcomes. Across all project phases, AINCRA supports alignment, justification, and documentation of strategies while fostering constructive, content-focused dialogue among project partners.

Discussion

Principal Findings

Despite AI's high potential to support nursing care, many use cases in acute and long-term care are still underexplored [13], and R&D projects face multifaceted challenges when designing, implementing, or evaluating AI systems in everyday nursing care. In this study, we systematically developed dimensions, attributes, and levels of AI readiness based on published empirical evidence as well as insights from experts with practical experience in conducting AINC projects. Integrating findings from prior work, a systematic literature review and expert discussions directly shaped the development of AINCRA. Attributes and dimensions identified in earlier work provided an initial conceptual structure that was operationalized in a first AINCRA version. This version served as the basis for think-aloud interviews and discussions with experts during which the relevance, clarity, and completeness of individual attributes were critically examined. Feedback from these discussions led to the refinement and consolidation of attributes and informed adjustments. As a result, the final AINCRA reflects a synthesis of theoretical grounding, empirical evidence, and practical expert insight. The resulting AINCRA comprises 5 dimensions and 69 AI readiness attributes across 5 readiness levels that can guide AINC project planning, implementation, and evaluation. While experts with practical experience in conducting AINC projects rated the readiness level 5 (optimized) as ambitious and often unrealistic for AINC projects, they recognized AINCRA's value and overall benefit to support AINC projects, especially with respect to early identification and thus avoidance of common pitfalls of these projects.

A major strength of AINCRA is its comprehensive inclusion of evidence-based, previously established AI readiness factors, with a focus on processual and translational aspects. Prior research on a sociotechnical perspective for responsible AI maturity models has shown that organizations often focus on technical and business assessments when developing and deploying AI systems, while overlooking critical ethical, social, and human dimensions [20]. We argue that this holds true for AINC projects as well, so that they will benefit from recognizing and systematically addressing the importance of both technical and social aspects [20]. In addition, AINCRA sets itself apart from other freely available or commercial instruments, such as the analytics maturity assessment model promoted by the Healthcare Information and Management Systems Society through its consistent focus on the specific application context of nursing care. Many of the models existing before the AINCRA development have not been thoroughly reviewed by the scientific community. Therefore, they were not identified during the systematic literature review. AINCRA goes beyond preexisting models, offering a framework that is both more specialized and scientifically validated for use in AINC projects. As most attributes were derived from and may be transferred to a global context, we see overall generalizability of our results, even though AINCRA was developed within the European legal framework for AI in health care settings. While experiences of AINC experts stem from a German health care and research system perspective that may differ from other countries' prerequisites and available resources, for example, for funding AINC project staff, technical infrastructure or knowledge transfer and dissemination activities, as well as structures, processes, and outcomes of nursing care, we believe that European-centric attributes can be easily adapted to other regional or national requirements.

Implications for AI R&D, Nursing Science, and Clinical Partner Organizations

Given the widely reported issue that actors from health care and information science often do not share the same domain language in R&D projects and frequently work past one another [50], creating transparency and a common understanding of project aims and key concepts should be a shared goal among AINC project partners. AINCRA contributes to achieving this goal by raising awareness of relevant aspects. It also provides a foundation for decision-making at the project initiation phase, helping to determine which partners need to be involved, at what stages, and with what level of resource commitment. Additionally, it enables a better estimation of the risks and costs associated with AI development and deployment. By operationalizing and making the AINC project's progress more visible across different domains and project partners, it fosters a shared basis for discussion and terminology, promoting more effective collaboration.

Leveraging the potential of AI in health and nursing care requires robust estimates of a variety of factors relevant for such an investment, including implementation time, associated costs, regulatory risks, and technology acceptance among potential users and patients [21,25,30]. These estimates call for a thorough understanding of the status quo of health care institutions with respect to a variety of dimensions as detailed in AINCRA. The structure provided with AINCRA can support strategic decisions on investments by providing reliable insights on the efforts required to implement AI technology in clinical practice. The attributes developed in this study can also help to identify gaps and future areas of improvement with regard to AI technology in nursing practice. In addition, the insights gained in expert interviews help to manage expectations in translating and

XSL•FO
RenderX

disseminating AI research to nursing practice. It is evident that an assessment of AI readiness will be challenging or even impossible without transdisciplinary teams that combine expertise in technical, regulatory, and ethical questions with expertise in nursing care.

Although AINCRA is extensive and complex, by advocating for an interprofessional approach, it enables a comprehensive evaluation of AINC projects, due to its high level of detail, ensuring that no blind spots are overlooked. Furthermore, AINCRA is applicable to a wide range of projects, as the term AI is defined broadly, making it relevant across various areas within the field.

Comparison With Design Recommendations and Common Criticism of Maturity Models

Eight requirements for designing maturity models [23] were applied throughout the AINCRA development process. Table 5 outlines how these requirements were addressed and implemented in this study. Above all, the comprehensive comparison with existing maturity models, alongside an iterative approach and the sequential expert evaluation of attribute and level descriptions, strengthens the presented research.

Table . Incorporation of design recommendations in the AINCRA[a] development process.

Requirement (R) derived from [23]	Incorporation of requirement through...
R1: Comparison with existing maturity models	• Systematic literature review • Further development of own prior work • Focus on AI[b] readiness or maturity models in health or nursing care
R2: Iterative procedure	• Research process with 3 iterations
R3: Evaluation	• Application and rating of AINCRA pilot version through experts
R4: Multimethodological procedure	• Sequential exploratory multimethods design
R5: Identification of problem relevance	• Rating and confirmation of AINCRA's relevance in expert group discussions
R6: Problem definition	• Results of preliminary work [13]
R7: Targeted presentation of results	• Feedback from the target group of intended users • Use of discipline-specific terminology • Development of a standard vocabulary for level distinctions together with the intended users
R8: Scientific documentation	• Study documents • Ethics application • Mandatory reporting for funding body • Open access publication of AINCRA

[a]AINCRA: artificial intelligence nursing care readiness assessment.

[b]AI: artificial intelligence.

Three main criticisms of maturity models [24] should be reflected on regarding AINCRA's internal and external validity: (1) lack of a solid theoretical foundation without grounding elements such as maturity levels and dimensions in established academic literature [24], (2) lack of sufficient empirical support for the selection of dimensions or variables, and (3) lack of a clear definition of how maturity should be measured [24]. While we rigorously aimed to address the first 2 criticisms during development by review of the academic literature and iterative evaluation and consensus building with domain experts, a degree of ambiguity in the measurement of AINC readiness remains. This arises partly from the nature of AINCRA as a self-assessment instrument. On the other hand, it is worth noting that AINCRA provides qualitative but not quantitative criteria for assigning a maturity level. Hence, an answer to the question, for example, whether simply participating in one regional network is sufficient to meet the requirements for community building attributes, or whether involvement in at least 3 different networks is needed, remains open to interpretation. Users may assess individual attributes with varying degrees of realism and flexibility. Further operationalization is therefore strongly recommended.

Limitations and Future Research

Our results have limitations that call for further research. First, experts commented on 66 of 69 readiness attributes only. Attributes 2.39 strategies for long-term external support: software and hardware: upgrades, 2.40 strategies for long-term external support: software and hardware: maintenance, and 2.29 financial resources and investments: exploring alternative financing models for including clinical partners remained without comments. Why these attributes were not addressed, even though they have been identified as relevant for AI [41], is not obvious. Most likely is the limited time span of the think-aloud task, which may have constrained responses. Toward the end of the task, experts focused particularly on attributes

XSL•FO
RenderX

they found interesting or especially relevant. Besides, not all invited experts participated. The findings are based on the expertise of a subset of participants. Although key stakeholders were represented, it cannot be assumed that data saturation was fully achieved, as not all invited experts took part in the study, and data collection concluded after interviewing all experts who had agreed to participation.

Second, as research on organizational AI readiness has previously highlighted as a relevant limitation [21], while AINCRA was developed to systematize key attributes of AINC readiness, overlaps with readiness or frameworks for adopting and implementing other digital technologies can be recognized, such as stakeholder acceptance as a key attribute for technology adoption [51]. However, the project-specific perspective of AINCRA expands preexisting frameworks, which predominantly adopt an organizational perspective, thus providing AINC projects with a reflective tool and instrument to support project management.

Third, we decided against substantially shortening or prioritizing AINC readiness attributes, despite suggestions from some experts. This makes AINCRA more time-consuming and potentially less practical for everyday project work. However, this decision is informed by our experiences with multiple AINC projects over the past 5 years, which show that many projects underestimate their complexity and are affected by domain-specific blind spots that hinder collaboration. At the same time, our findings do not fully clarify how individual attributes interact or which should be addressed together when strengthening AINC readiness. While we applied an established method for maturity model evaluation that is being used in information science for decades [48], future research may strive for validating our results by applying AINCRA to real process improvement activities [48] in future AINC projects.

Conclusion

In a field where many use cases remain underexplored and R&D projects often face substantial barriers, AINCRA offers structured guidance to enhance AI readiness for research, development, and implementation of AI projects in nursing care. AINCRA provides a balanced integration of technical, organizational, ethical, and social aspects of AI readiness, addressing a common shortcoming of existing maturity models that tend to prioritize technical or business considerations. Its nursing-specific focus further distinguishes it from other generic or commercially available assessment instruments, making it both context-sensitive and scientifically grounded. Overall, AINCRA offers both a practical tool for current projects and a foundation for future research and refinement in the evolving field of AI systems in nursing care.

Acknowledgments

The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy, the following tasks were delegated to GAI tools under full human supervision: Translation. The GAI tool used was: ChatGPT 5. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by: Kathrin Seibert Additional note: We used ChatGPT-5 to assist with translating the German language full AINCRA version and user manual to English language. We used for GPT-4o to generate initial level descriptions based on the AI readiness vocabulary for level distinction developed by the authors for this study.

Funding

This study was funded by the German Ministry for Research, Technology and Space, Grant ID 16SV8835, the Einstein Center Digital Future, Berlin, and the German Research Foundation (Project number: 528483508 - FIP 12). The funders played no role in study design, data collection, analysis, and interpretation of data, or the writing of this manuscript.

Data Availability

All data generated or analyzed for the systematic literature review are included in this published article and its supplementary information files. The datasets generated and analyzed based on think-aloud interviews and focus group discussions are not publicly available, to ensure participant privacy but are available from the corresponding author upon reasonable request. Data are stored in a controlled access data storage at the University of Bremen.

Authors' Contributions

Conceptualization: KS, DD, FB, DF, DB, KW
Data curation: KS
Formal analysis: KS, JA, AN, JP, LB, DW
Funding acquisition: FB, DF, DB, KW
Investigation: KS, JA, AN, RG, LB
Methodology: KS, DD, AN, DF
Project administration: KS, DD
Supervision: KS, FB, DF, KW
Validation: KS, SJ, FB, AN, RG, MS, DF, JP, LB, KB, KW
Visualization: KS, LB

XSL•FO
RenderX

Writing – original draft: KS
Writing – review & editing: DD, JA, SJ, FB, AN, RG, MS, DF, JP, LB, DW, KB, DB, KW

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional methods information and results.
[PDF File, 968 KB - nursing_v9i1e84148_app1.pdf]

Multimedia Appendix 2
Artificial intelligence nursing care readiness assessment full version.
[PDF File, 1459 KB - nursing_v9i1e84148_app2.pdf]

Multimedia Appendix 3
Artificial intelligence nursing care readiness assessment user manual.
[PDF File, 6038 KB - nursing_v9i1e84148_app3.pdf]

Checklist 1
PRISMA 2020 checklist.
[PDF File, 68 KB - nursing_v9i1e84148_app4.pdf]

Checklist 2
COREQ checklist.
[PDF File, 297 KB - nursing_v9i1e84148_app5.pdf]

References

1. Hassanein S, El Arab RA, Abdrbo A, et al. Artificial intelligence in nursing: an integrative review of clinical and operational impacts. Front Digit Health 2025;7:1552372. [doi: 10.3389/fdgth.2025.1552372] [Medline: 40124108]
2. O'Connor S, Yan Y, Thilo FJS, Felzmann H, Dowding D, Lee JJ. Artificial intelligence in nursing and midwifery: a systematic review. J Clin Nurs 2023 Jul;32(13-14):2951-2968. [doi: 10.1111/jocn.16478] [Medline: 35908207]
3. Seibert K, Domhoff D, Bruch D, et al. Application scenarios for artificial intelligence in nursing care: rapid review. J Med Internet Res 2021 Nov 29;23(11):e26522. [doi: 10.2196/26522] [Medline: 34847057]
4. Approval of the content of the draft communication from the commission - commission guidelines on the definition of an artificial intelligence system established by regulation (EU) 2024/1689 (AI act). : European Commission; 2025 URL: https://ec.europa.eu/newsroom/dae/redirection/document/118559 [accessed 2026-05-08]
5. Raza MM, Venkatesh KP, Kvedar JC. Generative AI and large language models in health care: pathways to implementation. NPJ Digit Med 2024 Mar 7;7(1):62. [doi: 10.1038/s41746-023-00988-4] [Medline: 38454007]
6. Majjouti K, Priester V, Tapp-Herrenbrueck M, et al. Nursing-centered development of an AI-based decision support system in pressure ulcer and incontinence-associated dermatitis management - a mixed methods study. BMC Nurs 2025 Jul 1;24(1):808. [doi: 10.1186/s12912-025-03448-4] [Medline: 40598248]
7. Kopanz J, Mader JK, Donsa K, et al. Digital algorithm-guided insulin therapy in home healthcare for elderly persons with type 2 diabetes: a proof-of-concept study. Front Clin Diabetes Healthc 2022;3:986672. [doi: 10.3389/fcdhc.2022.986672] [Medline: 36992745]
8. Majjouti K, Tapp-Herrenbrueck M, Pinnekamp H, et al. KIADEKU: identification of wound types with AI. In: Yurish SY, editor. Presented at: Proceedings of the 1st International Conference on AI in Medicine and Healthcare (AiMH' 2025); Apr 8-10, 2025 URL: https://sensorsportal.com/AIMH_2025/AIMH_2025_Proceedings.pdf [accessed 2026-05-08]
9. Alves J, Azevedo R, Marques A, Encarnação R, Alves P. Pressure injury prediction in intensive care units using artificial intelligence: a scoping review. Nurs Rep 2025 Apr 9;15(4):126. [doi: 10.3390/nursrep15040126] [Medline: 40338264]
10. Gallo RJ, Shieh L, Smith M, et al. Effectiveness of an artificial intelligence-enabled intervention for detecting clinical deterioration. JAMA Intern Med 2024 May 1;184(5):557-562. [doi: 10.1001/jamainternmed.2024.0084] [Medline: 38526472]
11. Schwabe K, Ferizaj D, Neumann S. Reducing nurses' workload with an AI speech assistant for documentation. In: Yurish SY, editor. 2025 Presented at: First International Conference on AI in Medicine and Healthcare (AiMH' 2025); Apr 8-10, 2025. [doi: 10.13140/RG.2.2.20387.77606]
12. Ferizaj D, Neumann S. Assessing perceptions and experiences of an AI-driven speech assistant for nursing documentation: a qualitative study in german nursing homes. 2024 Presented at: Human-Computer Interaction: Thematic Area, HCI 2024,

XSL•FO
RenderX

Held as Part of the 26th HCI International Conference; Jun 29 to Jul 4, 2024; Washington, DC, USA p. 17-34. [doi: 10.1007/978-3-031-60449-2_2]
13. Seibert K, Domhoff D, Fürstenau D, Biessmann F, Schulte-Althoff M, Wolf-Ostermann K. Exploring needs and challenges for AI in nursing care – results of an explorative sequential mixed methods study. BMC Digit Health 2023;1(1). [doi: 10.1186/s44247-023-00015-2]
14. von Gerich H, Moen H, Block LJ, et al. Artificial Intelligence -based technologies in nursing: a scoping literature review of the evidence. Int J Nurs Stud 2022 Mar;127:104153. [doi: 10.1016/j.ijnurstu.2021.104153] [Medline: 35092870]
15. Feuerriegel S, Frauen D, Melnychuk V, et al. Causal machine learning for predicting treatment outcomes. Nat Med 2024 Apr;30(4):958-968. [doi: 10.1038/s41591-024-02902-1] [Medline: 38641741]
16. State of the world's nursing 2025. : World Health Organization; 2025 URL: https://www.who.int/publications/i/item/9789240110236 [accessed 2026-05-08]
17. EFN policy statement on improving frontline nurses' time for direct patient care with digitalisation and responsible AI. : European Federation of Nurses Association; 2024 URL: https://efn.eu/wp-content/uploads/2024/10/EFN-PS-improving-frontline-nurses-time-for-direct-patient-care-with-digitalisation-responsible-AI-Oct.-2024.pdf [accessed 2026-05-08]
18. Ball Dunlap PA, Michalowski M. Advancing AI data ethics in nursing: future directions for nursing practice, research, and education. JMIR Nurs 2024 Oct 25;7:e62678. [doi: 10.2196/62678] [Medline: 39453630]
19. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on the domains of nursing: scoping review. JMIR Nurs 2020 Dec 17;3(1):e23939. [doi: 10.2196/23939] [Medline: 34406963]
20. Akbarighatar P, Pappas I, Vassilakopoulou P. A sociotechnical perspective for responsible AI maturity models: findings from a mixed-method literature review. Int J Inf Managment Data Insights 2023 Nov;3(2):100193. [doi: 10.1016/j.jjimei.2023.100193]
21. Jöhnk J, Weißert M, Wyrtki K, Not AC. Ready or not, AI comes— an interview study of organizational AI readiness factors. Bus Inf Syst Eng 2021 Feb;63(1):5-20. [doi: 10.1007/s12599-020-00676-7]
22. Fountaine T, McCarthy B, Saleh T. Building the AI-Powered Organization. Technology isn't the biggest challenge. Culture is. Harv Bus Rev 2019:62-73 [FREE Full text]
23. Becker J, Knackstedt R, Pöppelbuß J. Developing maturity models for IT management. Bus Inf Syst Eng 2009 Jun;1(3):213-222. [doi: 10.1007/s12599-009-0044-5]
24. Lasrado LA, Vatrapu R, Andersen KN. Maturity models development in IS research: a literature review. 2015 Presented at: System design for, with and by users; Aug 9-12, 2015 URL: https://aisel.aisnet.org/iris2015/6/ [accessed 2026-05-08]
25. Pumplun L, Fecho M, Wahl N, Peters F, Buxmann P. Adoption of machine learning systems for medical diagnostics in clinics: qualitative interview study. J Med Internet Res 2021 Oct 15;23(10):e29301. [doi: 10.2196/29301] [Medline: 34652275]
26. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
27. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
28. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
29. Seibert K, Domhoff D, Altona J, et al. ProKIP: rationale and development of the AI-nursing-care-readiness-assessment (AINCRA). In: Yurish SY, editor. 2025 Presented at: In First International Conference on AI in Medicine and Healthcare (AiMH' 2025); Apr 8-10, 2025; Innsbruck, Austria p. 8-10 URL: https://sensorsportal.com/AIMH_2025/AIMH_2025_Proceedings.pdf [accessed 2026-05-08]
30. Chang A. Implementation of artificial intelligence in medicine, in intelligence-based medicine artificial. In: Artificial Intelligence and Human Cognition in Clinical Medicine and Healthcare: Academic Press; 2020:397-412. [doi: 10.1016/B978-0-12-823337-5.00009-3]
31. Boren SA, Moxley D. Systematically reviewing the literature: building the evidence for health care quality. Mo Med 2015;112(1):58-62. [Medline: 25812277]
32. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
33. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544. [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
34. Willis GB. Analysis of the Cognitive Interview in Questionnaire Design: Oxford University Press; 2015.
35. Doctor E, Eymann T, Fürstenau D, et al. A maturity model for assessing the digitalization of public health agencies. Bus Inf Syst Eng 2023 Oct;65(5):539-554. [doi: 10.1007/s12599-023-00813-y]
36. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]

XSL•FO
RenderX

37. AI nursing care readiness assessment (AINCRA): a tool for structured reflection and preparation of AI projects in nursing care. ProKIP. 2026. URL: https://aincra.prokip.care/en [accessed 2026-08-04]
38. Abuzaid MM, Elshami W, Tekin H, Issa B. Assessment of the willingness of radiologists and radiographers to accept the integration of artificial intelligence into radiology practice. Acad Radiol 2022 Jan;29(1):87-94. [doi: 10.1016/j.acra.2020.09.014] [Medline: 33129659]
39. Andersson J, Nyholm T, Ceberg C, et al. Artificial intelligence and the medical physics profession - a Swedish perspective. Phys Med 2021 Aug;88:218-225. [doi: 10.1016/j.ejmp.2021.07.009] [Medline: 34304045]
40. Weinert L, Müller J, Svensson L, Heinze O. Perspective of information technology decision makers on factors influencing adoption and implementation of artificial intelligence technologies in 40 German hospitals: descriptive analysis. JMIR Med Inform 2022 Jun 15;10(6):e34678. [doi: 10.2196/34678] [Medline: 35704378]
41. Alami H, Lehoux P, Denis JL, et al. Organizational readiness for artificial intelligence in health care: insights for decision-making and practice. J Health Organ Manag 2020 Dec 3;ahead-of-print(ahead-of-print). [doi: 10.1108/JHOM-03-2020-0074] [Medline: 33258359]
42. Wiljer D, Hakim Z. Developing an artificial intelligence-enabled health care practice: rewiring health care professions for better care. J Med Imaging Radiat Sci 2019 Dec;50(4 Suppl 2):S8-S14. [doi: 10.1016/j.jmir.2019.09.010] [Medline: 31791914]
43. Jennett P, Yeo M, Pauls M, Graham J. Organizational readiness for telemedicine: implications for success and failure. J Telemed Telecare 2003;9 Suppl 2:S27-S30. [doi: 10.1258/135763303322596183] [Medline: 14728753]
44. Pumplun L, Tauchert C, Heidt M. A new organizational chassis for artificial intelligence - exploring organizational readiness factors. 2019 Presented at: 27th European conference on insformation systems (ECIS) URL: https://aisel.aisnet.org/ecis2019_rp/106/ [accessed 2026-05-08]
45. Someren MW, Barnard YF, Sandberg JAC. The Think Aloud Method A Practical Guide to Modelling Cognitive Processes: Academic Press; 1994.
46. Wolcott MD, Lobczowski NG. Using cognitive interviews and think-aloud protocols to understand thought processes. Curr Pharm Teach Learn 2021 Feb;13(2):181-188. [doi: 10.1016/j.cptl.2020.09.005] [Medline: 33454077]
47. Peute LWP, de Keizer NF, Jaspers MWM. The value of retrospective and concurrent think aloud in formative usability testing of a physician data query tool. J Biomed Inform 2015 Jun;55:1-10. [doi: 10.1016/j.jbi.2015.02.006] [Medline: 25733166]
48. Helgesson YYL, Höst M, Weyns K. A review of methods for evaluation of maturity models for process improvement. J Software Evolu Process 2012 Jun;24(4):436-454 [FREE Full text] [doi: 10.1002/smr.560]
49. Wolf-Ostermann K, Fürstenau D, Theune S, et al. Concept for embedding AI systems in nursing care: exploratory project on AI in nursing care (sokip). : University of Bremen URL: https://media.suub.uni-bremen.de/entities/publication/7f4e91a6-6598-4b1c-bec7-fcbf513dd8eb [accessed 2026-05-08]
50. Nyström ME, Karltun J, Keller C, Andersson Gäre B. Collaborative and partnership research for improvement of health and social services: researcher's experiences from 20 projects. Health Res Policy Syst 2018 May 30;16(1):46. [doi: 10.1186/s12961-018-0322-0] [Medline: 29843735]
51. Greenhalgh T, Abimbola S. The NASSS framework - a synthesis of multiple theories of technology implementation. Stud Health Technol Inform 2019 Jul 30;263:193-204. [doi: 10.3233/SHTI190123] [Medline: 31411163]

Abbreviations

AI : artificial intelligence
AINC : artificial intelligence in nursing care
AINCRA : artificial intelligence nursing care readiness assessment
COREQ: Consolidated Criteria for Reporting Qualitative Research
EU AI: European Union artificial intelligence
ML: machine learning
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
R&D : research and development

XSL•FO
RenderX

Edited by E Borycki; submitted 15.Sep.2025; peer-reviewed by O Nguyen, Q Wang; revised version received 08.Apr.2026; accepted 21.Apr.2026; published 02.Jun.2026.

Please cite as:
Seibert K, Domhoff D, Altona J, Jäger S, Bießmann F, Nowak A, Gubser R, Schulte-Althoff M, Fürstenau D, Pohle J, Bergmann L, Walter D, Beier K, Borchers D, Wolf-Ostermann K
Readiness Assessment for AI in Nursing Care Projects: Multimethods Study
JMIR Nursing 2026;9:e84148
URL: https://nursing.jmir.org/2026/1/e84148
doi:10.2196/84148

© Kathrin Seibert, Dominik Domhoff, Janissa Altona, Sebastian Jäger, Felix Bießmann, Alessia Nowak, Rahel Gubser, Matthias Schulte-Althoff, Daniel Fürstenau, Jörg Pohle, Lea Bergmann, David Walter, Kathi Beier, Dagmar Borchers, Karin Wolf-Ostermann. Originally published in JMIR Nursing (https://nursing.jmir.org), 2.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Nurses' Evaluation of a Service Robot for Inpatient Care: Technology Acceptance Study

Christopher Friese[1], MSc; Robert Klebbe[1], MA; Anika Heimann-Steinert[2], Dr rer medic

[1]Department of Geriatrics and Medical Gerontology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Augustenburger Platz 1, Berlin, Germany

[2]Institute of Public Health, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

Corresponding Author:
Christopher Friese, MSc
Department of Geriatrics and Medical Gerontology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Augustenburger Platz 1, Berlin, Germany

Abstract

Background: The integration of robotic systems into nursing practice is increasingly discussed as a potential strategy to alleviate workload and support care processes in response to demographic changes and staffing shortages. However, the acceptance of nursing staff as primary end users remains a critical determinant for successful implementation. Despite technological advances, the practical requirements and perspectives of nursing staff have not been adequately considered in research and development efforts to date.

Objective: Building on the user-centered development approach applied, this study aimed to examine nursing staff's evaluation of a service robot designed to assist with routine tasks in inpatient care, as well as their intention to use it, while accounting for technology-specific and psychological determinants of acceptance.

Methods: A total of 30 nurses tested the robot across 3 application scenarios (information service, item delivery, and beverage delivery) in a simulated care setting, alternating between the roles of nurse and care recipient. Acceptance-related constructs, including intention to use, were measured using the Technology Usage Inventory. General attitudes toward robots were assessed via the General Attitudes Towards Robots Scale. Participants' prior experience with robotics was also documented. Spearman rank correlations and Mann-Whitney *U* tests were used for analysis.

Results: The robot was rated positively across all dimensions. Usability was high (median 20, IQR 18-21; scale range 3 - 21), as was perceived usefulness (median 21, IQR 16-24; range 4 - 28). Skepticism was low (median 10.5, IQR 7-12; scale range 4 - 28), and accessibility was moderate (median 10, IQR 8-13; scale range: 3 - 21). Intention to use was strong (median 224.5, IQR 157-248; scale range 0 - 300) and correlated positively with usability (r_s(28)=0.505; *P*=.004), perceived usefulness (r_s(28)=0.74; *P*<.001), and accessibility (r_s(28)=0.628; *P*<.001), and negatively with skepticism (r_s(28)=–0.516; *P*=.004). More positive personal attitudes toward robots were also associated with higher perceived usefulness (r_s(28)=0.549; *P*=.002) and greater intention to use (r_s(28)=0.483; *P*=.007). No significant differences in intention to use were found between participants with and without prior robotics experience (*U*=83.5; *P*=.62).

Conclusions: The findings indicate a high level of acceptance among the participating nursing staff for the developed service robot within the tested scenarios. Considering the chosen user-centered development approach, they further underscore the need for strategies that combine participatory design, transparent communication of system capabilities and limitations, and structured opportunities for hands-on experience. Such measures, together with proactive knowledge transfer and skills development, are essential to sustainably leverage the practical potential of service robotics in nursing practice.

Trial Registration: OSF Registries osf.io/ems7b; https://osf.io/ems7b

(JMIR Nursing 2026;9:e86824) doi:10.2196/86824

KEYWORDS

nursing; inpatient care; digital health; assistive technology; robotics; service robotics; technology acceptance; human-robot interaction; user-centered design

XSL•FO
RenderX

Introduction

Background

Against the backdrop of an aging population and a shortage of nursing staff [1,2], robotic assistance is being discussed as a promising form of support in inpatient and hospital care [3,4]. A recent scoping review by Ohneberg et al [5] identified assistive robotic systems in nursing and categorized applications into domains, including information and data-related support, monitoring and safety functions, communication and telepresence, and task-oriented assistance in everyday care processes, alongside socially assistive functions targeting emotional care and cognitive support. Against this broader landscape, this paper focuses on service robotics as one class of assistive robotic systems in inpatient care.

Service Robotics in Nursing: Scope and Status Quo

Service robotics denotes technical systems that support people in the partially or fully automated execution of services and tasks. In addition to information-processing and sensory functions, service robots can move autonomously or carry out complex tasks comprising several work steps and materials [5]. Current developments in the nursing context aim at deploying service robots to support administrative tasks such as digital patient data collection and assistance during medical rounds [6,7]; logistical functions, including the transport of medication, supplies, and laundry [8-10]; as well as physically demanding care activities like patient lifting, transferring, and hygiene-related procedures involving both personal care and environmental cleaning tasks [11-14]. At present, the use of assistive robotic systems in nursing remains heterogeneous and occurs primarily during development phases and pilot implementations, although wider practical implementation is anticipated in the coming years [5,15,16].

Early evidence suggests that diverse stakeholder groups, such as nursing staff, are generally open to the integration of robotic assistance, provided that the technology complements rather than replaces human care [4,17]. The expectation is for robots to act as collaborative assistants, reducing the high workload without fully substituting nurses' tasks [18,19]. Their introduction, however, entails a fundamental transformation of nursing work processes [3,5,20]. Consequently, in light of these systemic changes, end user acceptance emerges as a critical factor in the development and deployment of assistive robotic systems in nursing, as it has been consistently identified as a key enabler for successful implementation, as well as for their effective and efficient long-term use [21,22].

Conceptual Background: Technology Acceptance and Intention to Use

In technology research, acceptance is commonly defined as users' willingness to use a technology for the tasks it is intended to support [23]. This broad definition is typically operationalized in technology-acceptance frameworks through users' perceptions of a technology's characteristics and its suitability for intended tasks and conditions of use. In turn, these perceptions shape intention to use, with intention to use serving as the most proximal predictor of actual system use [24,25].

A prominent intention-based framework is the technology acceptance model (TAM) by Davis [26]. The TAM explains intention to use primarily through perceived usefulness and perceived ease of use, linking these perceptions to attitudes toward using the technology and, ultimately, to behavioral intention. While the TAM has been widely applied in health care contexts, its comparatively narrow focus has been criticized for failing to capture the full complexity of technology acceptance in such settings [27]. This critique aligns with a growing body of research highlighting limitations of established acceptance models and calling for broader conceptual frameworks that account for both technical attributes and individual-level factors, a perspective supported by recent reviews [28,29] and empirical studies [30-32].

In contrast to TAM's parsimonious focus, broader implementation-oriented models such as the unified theory of acceptance and use of technology extend the explanatory scope by incorporating social and organizational determinants relevant to real-world implementation [25]. Diffusion-of-innovation perspectives, in turn, address adoption dynamics over time at the system level [33]. These approaches, as well as more general behavioral and health-behavior frameworks such as the theory of reasoned action, the theory of planned behavior, or the health belief model, were not adopted because this study targets technology-specific acceptance perceptions in a workplace context rather than implementation processes or health behavior change [34-36].

Nursing Staff Perspectives on Robotic Assistance

From this perspective, nursing staff perceptions and acceptance criteria are particularly relevant given their role as the primary end users of assistive robotic systems. Although nursing staff play a central role in implementation and everyday use of such systems, their perspectives have long been underrepresented in technology acceptance research [3,37]. Recent studies have begun to address this gap, offering initial insights into nurses' acceptance criteria. However, findings to date point to a complex and multifactorial picture.

Consistent with prior technology acceptance research, the robot's perceived usefulness remains pivotal. Nurses are more likely to adopt a system that demonstrably reduces workload by assisting with physically demanding tasks such as lifting, transferring, or routine deliveries [3,38,39]. Ease of use is equally critical, as increased operational complexity or system failure can offset intended benefits [40]. Acceptance is further shaped by perceptions of reliability and safety. Only a system that performs dependably and poses no risk to care recipients is likely to be entrusted with responsibilities in patient care [41].

Beyond these technology-related considerations, individual-level factors such as attitudes, perceptions, and prior experiences play a critical role. A key factor is nurses' general attitude toward technology and robots in particular. Nurses who are technologically inclined or have had positive experiences with new technologies tend to be more receptive to robotic assistance [21,32,42,43]. Multiple studies, including a large-scale survey of nursing staff, have repeatedly found that prior experience and knowledge specifically related to robotics positively influence acceptance. Nurses with first-hand experience and

XSL•FO
RenderX

greater confidence in operating such systems consistently evaluate their usefulness and applicability more positively [31,38,44].

Taken together, this evidence suggests that acceptance is not a static reaction to technological features, but a dynamic process shaped by individual attitudes, prior experiences, and progressive familiarization. Moreover, empirical findings highlight the importance of early end user involvement in fostering acceptance. When nurses are engaged in planning, adaptation, and decision-making processes, they are more likely to perceive the robot as useful and aligned with their workflow, which in turn strengthens their intention to use the system in everyday nursing practice [45].

Rationale and Objectives

Building on the limited routine deployment of assistive robotic systems in nursing and the multifactorial acceptance determinants outlined above, further evidence is needed from hands-on evaluations of concrete prototypes that are aligned with nursing workflows and professional requirements. Involving nursing staff in participatory development and providing opportunities for direct interaction are widely considered crucial to avoid technology-driven implementation and to ensure that robotic systems fit the practical realities, values, and needs of nursing practice [5,19,20,41].

The research and development project RoMi (Roboterunterstützung bei Routineaufgaben zur Stärkung des Miteinanders in Pflegeeinrichtungen), funded by the German Federal Ministry of Education and Research, aimed to develop a mobile service robot to support routine nonclinical tasks in inpatient care, such as information provision and logistical activities. Nursing staff were actively involved in key phases of the project, including the examination of contextual and ethical prerequisites for robot use [46], investigations of preferences regarding robot design and interaction characteristics [47-49], and the codevelopment of application scenarios, capability requirements, and evaluation criteria for context-sensitive human-robot collaboration [19].

Following this user-centered development approach, the resulting functional prototype was evaluated in an end user study with nursing staff. The study examined participants' general attitudes toward robots, their evaluation of the developed service robot, and their intention to use it. Accordingly, we conceptualized acceptance in terms of intention to use and its multifactorial determinants, integrating both technology-related appraisals and individual-level factors. This paper aimed to advance a differentiated understanding of the usage potential of service robots in nursing, as well as the factors influencing their acceptance and prospective integration within inpatient care settings.

Methods

Procedure

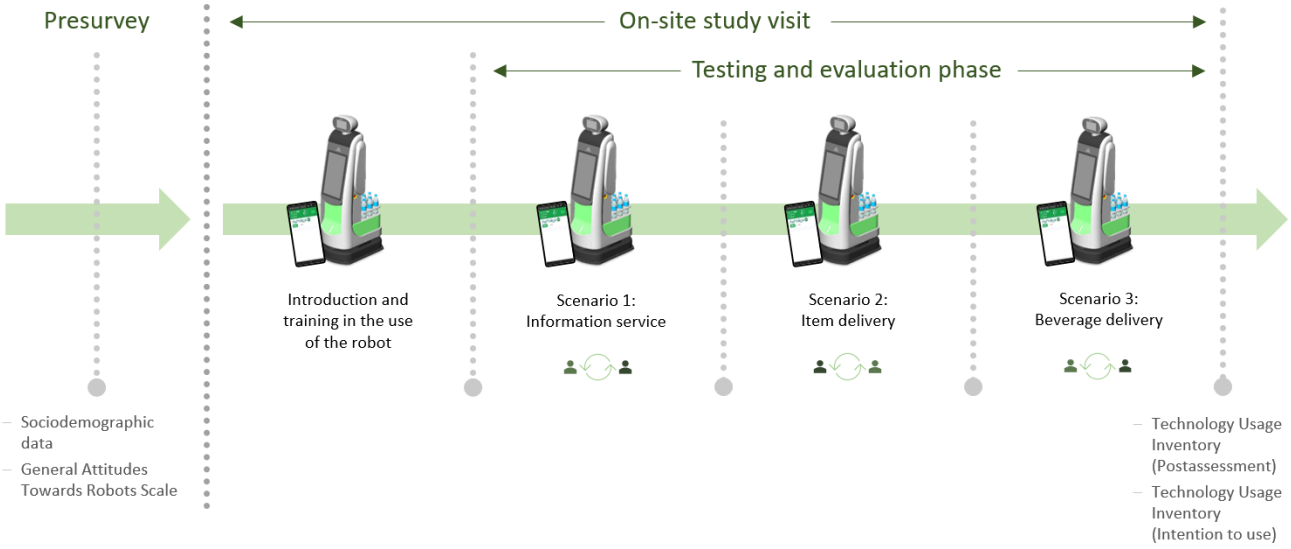
The study consisted of an online presurvey and a subsequent testing and evaluation phase of the developed service robot. To recruit participants, various inpatient care facilities were informed about the study, enabling interested individuals to contact the study team directly. Once interest was expressed, potential participants received the study information and consent form, followed by the scheduling of a telephone consultation. During this call, the study's objectives, design, and procedures were reiterated, and the inclusion criteria were assessed. These included (1) minimum age of 18 years; (2) current or past occupation as a nurse, nursing assistant, geriatric nurse, geriatric nursing assistant, pediatric nurse, occupational therapist, remedial therapist, or physical therapist in inpatient care or daycare; and (3) sufficient proficiency in the German language.

After providing informed consent, participants received a link to the online presurvey, which gathered sociodemographic data and included the General Attitudes Towards Robots Scale (GAToRS). Appointments were then arranged for the testing and evaluation phase, which took place at the research laboratory of the Geriatrics Research Group at Charité—Universitätsmedizin Berlin. Each session involved 2 participants who alternated between the roles of nurse and care recipient. Participants first received structured instruction and hands-on training covering safety procedures as well as the use of the robotic system and its smartphone app for task management and status monitoring.

The service robot was then tested and evaluated across three predefined application scenarios: information service, item delivery, and beverage delivery. These scenarios had been developed and refined in earlier empirical studies conducted within the RoMi project [19,46]. Each scenario was completed twice by every participant, once in each role. To ensure consistency, all received standardized written instructions outlining the required interaction procedures. After performing both roles within a scenario, the session proceeded to the next. Following the test phase, participants completed the posttest section of the Technology Usage Inventory (TUI), including the intention to use (ITU) scale, for the evaluation of the service robot. Each session lasted approximately 90 minutes. Figure 1 illustrates the overall study procedure.

XSL•FO
RenderX

Figure 1. Study procedure.



Robot

The service robot evaluated in this study was developed within the RoMi research project and represents an iteratively adapted prototype. A semihumanoid mobile service robot was selected to match the RoMi focus on routine, nonclinical support tasks, including information provision and logistics. Other systems such as patient-handling, telepresence, or socially assistive robots were not considered within RoMi because they address different task classes.

The developed service robot is based on the workerbot6 platform by pi4_robotics. Throughout the project, the platform was iteratively refined for the care setting, with technical implementation led by pi4_robotics. Iterative refinements were informed by empirical studies on robot design and interaction preferences led by Humboldt-Universität zu Berlin and by industrial design concepts contributed by Hochschule für Technik und Wirtschaft Berlin [47-49]. The authors affiliated with Charité—Universitätsmedizin Berlin contributed the user-centered workstream by conducting studies with nursing staff on practical and ethical considerations of robot use and translating the findings into the development and refinement of application scenarios, interaction sequences, and the associated capability requirements and evaluation criteria for the service robot [19,46].

The research prototype ultimately resulting from the RoMi project and evaluated in this study is shown in Figure 2. The service robot has a height of 175 cm and a weight of 133 kg.

XSL•FO
RenderX

Figure 2. Developed service robot (image used with permission from pi4_robotics).



Based on initial design investigations showing that a human-like appearance can contribute to the acceptance and perceived competence of robots, an anthropomorphic embodiment was chosen [46,47]. The robot's head unit features a screen displaying animated eyes and a mouth, as well as an integrated camera that detects human presence and enables context-sensitive autonomous interaction (eg, in patient rooms or at loading stations). A microphone for speech recognition is also embedded in the head section to enable low-effort voice input, which aligns with caregiver preferences for speech-based interaction [46-48].

The torso is equipped with a front-mounted 21.5-inch touchscreen display that provides graphical feedback on the robot's status and serves as an additional interface for human-robot interaction. This visual channel was intended to enhance interaction transparency and comprehensibility while accommodating potential age- or disease-related changes in the communicative abilities of care recipients [46]. Loudspeakers and an emergency stop button are positioned on both sides of the robot to support clear feedback and immediate shutdown in case of safety concerns. Lateral, arm-like structures function as rails that support a fold-out tray mounted at the rear, enabling the transport and handover of small items without requiring complex manipulation. The back of the robot also houses a beverage dispenser, which is capable of holding twelve 0.5 liter bottles and is equipped with integrated sensors and compartment lighting to support structured loading and ensure correct retrieval during beverage delivery.

The robot is mounted on a mobile platform that enables autonomous navigation based on a map configurable for the ward environment. The robot's navigation was configured with a safety-oriented low-speed profile (maximum 1.4 km/h) appropriate for shared indoor environments, consistent with RoMi project findings that caregivers perceive slower approach speeds as more comfortable and safer [49]. Moreover, the system incorporates a range of obstacle detection sensors. Upon encountering an obstacle, the robot first reduces speed and attempts to navigate around it. If avoidance is not possible, it stops and waits until the path is clear, thereby prioritizing safe behavior in the presence of uncertainty.

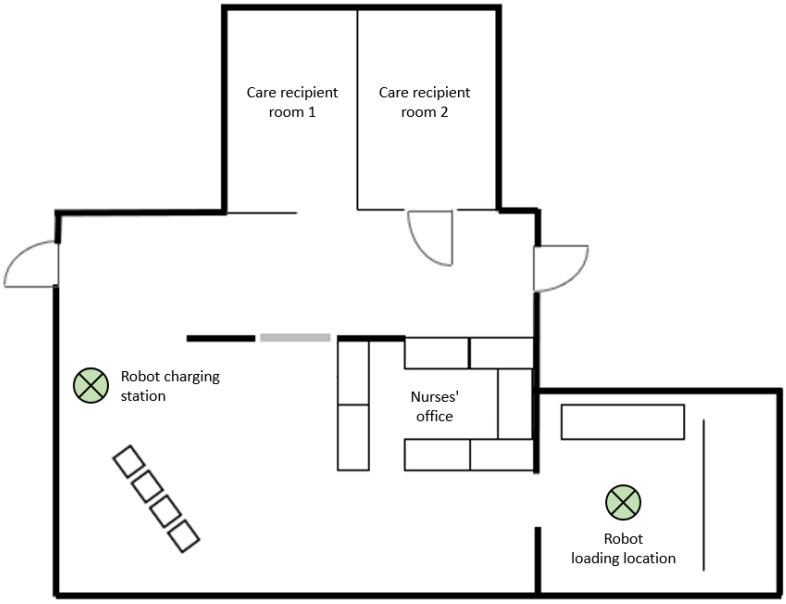
Throughout the study, participants used a smartphone app as the primary tool for giving instructions to the robot. The app also provided posttask status information.

Setting

To align the interaction with the robot more closely with the professional care context, a simulated nursing ward environment was set up in the research laboratory (Figure 3). The setup included a nurses' office, 2 patient rooms, and a storeroom serving as the loading station. To minimize mutual influence, partition elements were installed to ensure that participants could not observe each other's interactions while alternating roles.

XSL•FO
RenderX

Figure 3. Setup within the laboratory setting.



Application Scenarios

To standardize exposure to the robot, participants completed 3 scripted application scenarios that operationalized typical nonclinical support tasks and specified role-dependent interactions between the participant in the nurse role, the participant in the care recipient role, and the robot. Tasks were assigned via the robot's instruction and status-monitoring app.

Information Service

The participant in the nurse role created a notification task in the app by selecting the room and entering a free-text message. The robot navigated to the room, signaled arrival with a bell sound, and waited until a person entered its detection area. After person detection, the participant in the care recipient role confirmed readiness to receive the notification via voice or touchscreen. The robot then delivered the message via audio and on-screen output and offered an option to repeat it. If no repetition was requested, the robot returned to the charging station.

Item Delivery

The participant in the nurse role initiated a delivery task in the app by selecting the room and entering a free-text item list. The robot navigated to a predefined loading area and, after person detection, provided voice and on-screen loading instructions. The participant in the nurse role placed the items on the rear tray and started execution on the touchscreen. The robot delivered to the room, signaled arrival, and prompted a confirmation-based handover. The participant in the care recipient role confirmed via voice or touchscreen and retrieved the items, after which the robot returned to the charging station.

Beverage Delivery

This scenario comprised a standard and an extended sequence (phases 1 and 2). In phase 1, the participant in the nurse role configured deliveries for 2 rooms in the app using drop-down selection, loaded beverages into illuminated bottle holders at the loading area following robot guidance, and started execution on the touchscreen. The robot delivered to the room of the participant in the care recipient role using the standard arrival, person detection, confirmation, and pickup sequence. The second room delivery was conducted by study staff to complete a second standardized delivery sequence while keeping participant roles constant. In phase 2, 8 additional deliveries alternated between the 2 rooms. The participant in the care recipient role remained in one room and performed predefined interaction variations for deliveries to their room. These included a misunderstanding prompting repetition followed by rejection, as well as acceptance while using a walker and a wheelchair. In phase 2, the remaining deliveries were again handled by study staff. Rejected deliveries were documented in the app as not completed.

Across all scenarios, the participant in the nurse role performed typical documentation tasks in parallel to autonomous robot operation to reflect realistic time management practices. They also verified task completion in the app and documented the outcomes in accordance with the study protocol. Full scripts and step tables are provided in the supplementary material (Multimedia Appendix 1).

Measures

With regard to sociodemographic variables, the following data were collected: age, gender, highest educational qualification, current occupation, and previous experience with robots.

In addition, the GAToRS was applied. The GAToRS is a validated multidimensional instrument that accounts for both positive and negative attitudes toward robots. It comprises four subscales: (1) personal level positive attitude (P+), (2) personal level negative attitude (P–), (3) societal level positive attitude (S+), and (4) societal level negative attitude (S–). Thus, the scale captures (1) perceived comfort and pleasure and (2) discomfort and fears toward robots on a personal level, along with (3) rational hopes and (4) rational fears on a societal level. The questionnaire shows Cronbach α coefficients ranging from 0.74 to 0.84 across the 4 scales [50].

XSL•FO
RenderX

Furthermore, the TUI was used. The TUI is a validated questionnaire used in technology acceptance research to capture evaluations and selected determinants of intention to use for emerging technologies, encompassing technology-related perceptions and user-related factors. Subscales were selected a priori in line with the study focus on the immediate postinteraction evaluation of the service robot. Accordingly, we administered the postuse subscales usability, usefulness, skepticism, and accessibility together with the separate ITU scale. Reported internal consistencies (Cronbach α) for these subscales range from 0.70 to 0.81 [51].

Participants

A total of 30 participants took part in the study. The mean age was 43.2 (SD 11.2; range 22 - 67) years. The sample consisted of 20 women and 10 men.

At the time of participation, 11 individuals were employed as geriatric nurses and 2 as geriatric nursing assistants. Additionally, 3 participants worked as nurses and another 3 as pediatric nurses. Six were employed as nursing assistants. Five participants had previously worked in nursing and were currently employed in health care delivery-related fields, including nursing education (n=1), case management (n=1), health care controlling (n=1), and research (n=2).

Almost one-third of participants (n=9) reported prior experience with robotics, either in private (n=7), professional (n=1), or other contexts (n=1).

Data Analysis

Data were analyzed using IBM SPSS Statistics (version 29), applying both descriptive and inferential statistical methods. Within the GAToRS dataset, 10 missing items were identified without any discernible pattern. No variable exhibited more than 6.7% missing values. Given that previous studies have shown minimal differences in the performance of various imputation techniques when only 5% - 10% of values are missing [52,53], the use of simple imputation was deemed methodologically appropriate. Accordingly, mean imputation (series mean) was applied in SPSS.

Descriptive statistics are reported as means with SDs or, where appropriate, as medians with IQRs. As normal distribution could either not be assumed or was clearly absent, nonparametric tests were applied. Spearman rank correlation coefficient was used to assess statistical associations, and the Mann-Whitney *U* test was conducted for group comparisons. Statistical significance was set at α=.05, with correlations significant at the 0.01 level additionally highlighted.

Ethical Considerations

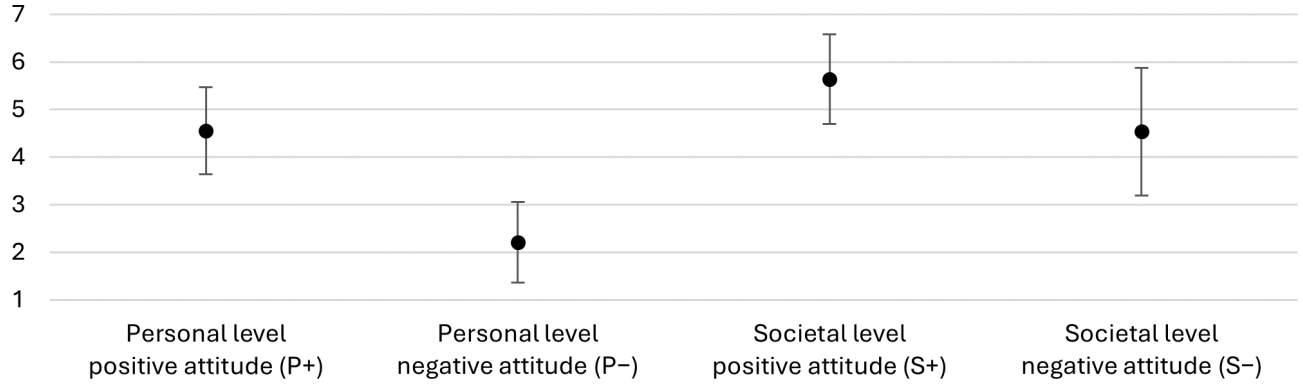
The study was conducted in accordance with the Declaration of Helsinki and the International Council for Harmonisation Guideline for Good Clinical Practice [54]. The study protocol was reviewed and approved by the Ethics Committee at the Institute of Psychology, Humboldt-University, Berlin (2023 - 14). All participants provided written informed consent prior to participation. Upon completion of the study, participants received financial compensation of EUR 30 (US $35). Privacy and confidentiality were ensured through data pseudonymization, secure storage of study data, and access restricted to authorized members of the research team.

Results

Participants' General Attitude Toward Robots

To assess participants' baseline attitudes toward robots, the GAToRS was applied (Figure 4). Each subscale yields scores from 1 to 7. Higher scores on P+ and S+ indicate stronger positive attitudes, whereas higher scores on P– and S– indicate stronger negative attitudes.

Figure 4. Results of the General Attitudes Towards Robots Scale subscales. The scale captures attitudes on two levels: personal (positive attitude=P+, negative attitude=P–) and societal (positive attitude=S+, negative attitude=S–). The x-axis displays the 4 subscales, and the y-axis shows the subscale scores. The dots represent the means, and the error bars indicate the SDs. The subscales range from 1 to 7. Higher values on P+ and S+ indicate stronger positive attitudes, whereas higher values on P– and S– indicate stronger negative attitudes.



At the personal level, positive attitudes ranged from 2.6 to 6.6 (mean 4.6, SD 0.9), while negative attitudes ranged from 1 to 4 (mean 2.2, SD 0.8).

At the societal level, positive attitudes showed the highest mean value across all subscales, with scores ranging from 3.2 to 7 (mean 5.6, SD 0.9). In contrast, negative attitudes showed the greatest variability, with scores between 1.8 and 7 (mean 4.5, SD 1.3).

Evaluation of the Developed Service Robot

The general evaluation of the developed service robot was conducted using the TUI, focusing on the immediate postintervention responses. The results included the subscales

XSL•FO
RenderX

usability, usefulness, skepticism, and accessibility (Figure 5), as well as the separate ITU scale (Figure 6). Higher scores indicate greater endorsement of the respective construct.

Figure 5. Results of the Technology Usage Inventory subscales. The x-axis displays the four subscales (usability, usefulness, skepticism, and accessibility), and the y-axis shows the subscale scores. The dots represent the medians, the error bars indicate the IQRs, and horizontal dotted lines mark the scale limits (usability and accessibility: 3 - 21; usefulness and skepticism: 4 - 28). Higher values indicate greater endorsement of the respective construct.

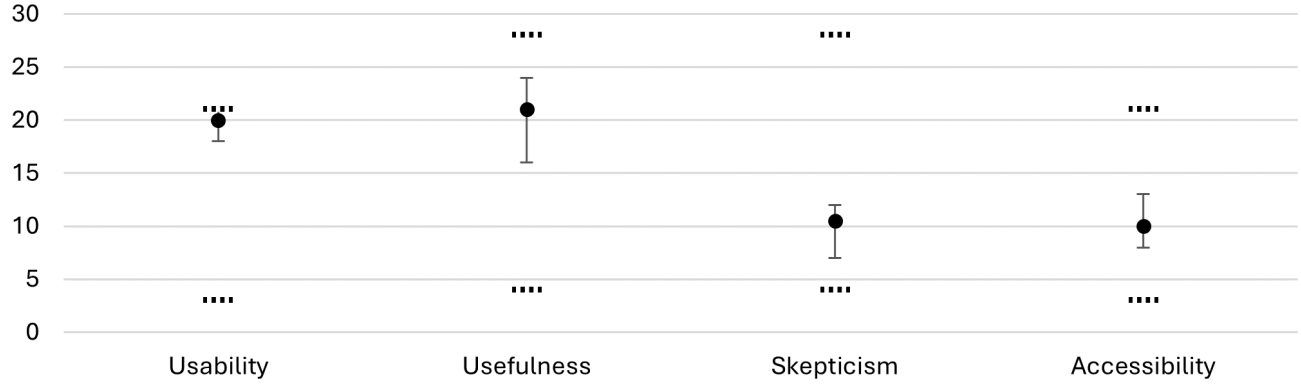
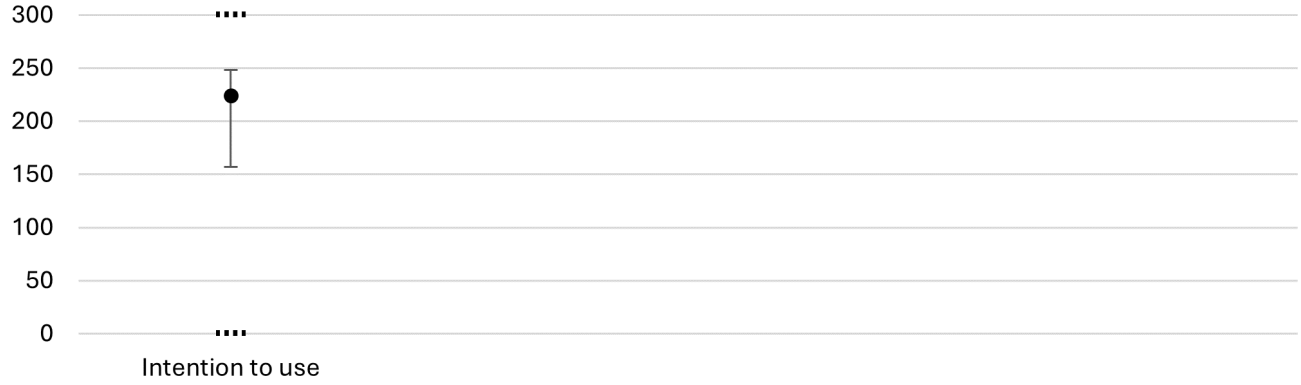


Figure 6. Results of the Technology Usage Inventory intention to use scale. The x-axis displays the scale, and the y-axis shows the score. The dot represents the median, the error bar indicates the IQR, and horizontal dotted lines mark the scale limits (0 - 300). A higher value indicates a stronger intention to use the technology.



Usability scores were between 14 and 21, with a median of 20 (IQR 18-21). Usefulness showed the widest spread among the 4 subscales, with values from 6 to 28, a median of 21 (IQR 16-24). Scores for skepticism varied from 4 to 24, with a median of 10.5 (IQR 7-12). Accessibility was rated between 3 and 21, with a median of 10 (IQR 8-13). For the ITU scale, scores spanned from 0 to 300, with a median of 224.5 (IQR 157-248).

Analysis of Dependencies and Influencing Factors

To measure the strength of the linear relationship between the participants' general attitude toward robots, based on the GAToRS, and the dimensions of the TUI, a Spearman correlation analysis was computed (Table 1).

XSL•FO
RenderX

Table . Results of the Spearman correlation analysis between the General Attitudes Towards Robots Scale and the Technology Usage Inventory.

	Usability	Usefulness	Skepticism	Accessibility	Intention to use
Personal level positive attitude					
r_s	0.057	0.549[a]	–0.317	0.457[b]	0.483[a]
P value	.76	.002	.09	.01	.007
N	30	30	30	30	30
Personal level negative attitude					
r_s	–0.270	–0.359	0.436[b]	–0.309	–0.226
P value	.15	.052	.02	.10	.23
N	30	30	30	30	30
Societal level positive attitude					
r_s	0.122	0.187	0.085	–0.072	0.179
P value	.52	.32	.66	.71	.34
N	30	30	30	30	30
Societal level negative attitude					
r_s	–0.153	–0.387[b]	0.479[a]	–0.576[a]	–0.159
P value	.42	.03	.007	<.001	.40
N	30	30	30	30	30

[a]Correlation is significant at the 0.01 level (2-tailed).

[b]Correlation is significant at the 0.05 level (2-tailed).

P+ showed a significant positive correlation with perceived usefulness ($r_s(28)=0.549$; $P=.002$), indicating a strong effect size [55]. Significant positive correlations were also found with accessibility ($r_s(28)=0.457$; $P=.01$) and intention to use ($r_s(28)=0.483$; $P=.007$), corresponding to medium effect sizes [55]. No significant correlations were observed with usability or skepticism.

P– correlated significantly positively with skepticism ($r_s(28)=0.436$; $P=.02$), suggesting a medium effect size [55]. No significant correlations were found with usability, perceived usefulness, accessibility, or intention to use.

S+ did not exhibit any significant correlations with the dimensions of the TUI.

S– was significantly negatively correlated with perceived usefulness ($r_s(28)=-0.387$; $P=.03$), indicating a medium effect size [55]. Additionally, a significant positive correlation was found with skepticism ($r_s(28)=0.479$; $P=.007$), also reflecting a medium effect size [55]. A significant negative correlation was further observed with accessibility ($r_s(28)=-0.576$; $P<.001$), corresponding to a strong effect size [55]. No significant correlations were found with usability or with intention to use.

To further examine associations between the TUI subscales (usability, usefulness, skepticism, and accessibility) and its ITU scale, an additional Spearman correlation analysis was conducted (Table 2).

XSL•FO
RenderX

Table . Results of the Spearman correlation analysis between the Technology Usage Inventory subscales (usability, usefulness, skepticism, and accessibility) and its intention to use scale.

	Intention to use
Usability	
r_s	0.505[a]
P value	.004
N	30
Usefulness	
r_s	0.740[a]
P value	<.001
N	30
Skepticism	
r_s	−0.516[a]
P value	.004
N	30
Accessibility	
r_s	0.628[a]
P value	<.001
N	30

[a]Correlation is significant at the 0.01 level (2-tailed).

Usability ($r_s(28)=0.505$; $P=.004$), perceived usefulness ($r_s(28)=0.740$; $P<.001$) and accessibility ($r_s(28)=0.628$; $P<.001$) showed significant positive correlations with intention to use. Skepticism, however, yielded a significant negative correlation ($r_s(28)=-0.516$; $P=.004$). All 4 correlation coefficients exceeded the threshold of 0.5, indicating strong effect sizes according to Cohen's criteria [55].

Table 3 presents the comparison of the TUI subscale scores between participants with and without prior robotics experience, analyzed using the Mann-Whitney *U* test.

Table . Results of the Technology Usage Inventory by robotics experience.

TUI[a] subscales and robotics experience	Participants, n	Median (IQR)	*U* test	*z*-score	*P* value
Usability					
No	21	20 (18-21)	73	−1.019	.31
Yes	9	21 (19-21)			
Usefulness					
No	21	22 (16-23)	82	−0.567	.57
Yes	9	21 (18-26)			
Skepticism					
No	21	11 (7-13)	75.5	−0.865	.39
Yes	9	8 (7-11)			
Accessibility					
No	21	11 (8-13)	93.5	−0.045	.96
Yes	9	9 (8-12)			
Intention to use					
No	21	211 (157-247)	83.5	−0.498	.62
Yes	9	241 (166-269)			

[a]TUI: Technology Usage Inventory.

XSL•FO
RenderX

Usability was rated slightly higher by participants with prior experience in robotics (median 21, IQR 19-21) compared to those without (median 20, IQR 18-21). Yet, this difference did not reach statistical significance (U=73; P=.31).

Perceived usefulness was also rated comparably across both groups, with slightly higher scores among participants without prior robotics experience (median 22, IQR 16-23) than among those with experience (median 21, IQR 18-26), but this difference also failed to achieve statistical significance (U=82; P=.57).

Regarding skepticism, participants with prior robotics experience reported lower levels (median 8, IQR 7-11) compared to those without (median 11, IQR 7-13). Similarly, this difference was not statistically significant (U=75.5; P=.39).

Accessibility was perceived as slightly higher by participants without prior robotics experience (median 11, IQR 8-13) relative to those with experience (median 9, IQR 8-12). Once again, this difference did not attain statistical significance (U=93.5; P=.96).

Finally, the intention to use was higher among participants with prior robotics experience (median 241, IQR 166-269) than among those without (median 211, IQR 157-247). However, the difference on this scale was not statistically significant either (U=83.5; P=.62).

Discussion

Summary of Key Findings

This study investigated nursing staff's general attitudes toward robotics, as well as their evaluation of and intention to use a newly developed service robot intended to support routine tasks in inpatient care settings. The quantitative analyses revealed a differentiated pattern of findings.

The GAToRS was used to assess participants' perceptions of robots. Ratings on the P+ subscale, which measures trust in robot developers, trust in robots, and comfort in interacting with robots, were moderately high across the sample. Consistent with this, ratings on the P– subscale, assessing discomfort with using robots, apprehension about misunderstandings, and unease in physical proximity to robots, were low and showed little variability. Responses on the S+ subscale, which captures perceived societal benefits of robots such as task delegation, safety, and assistance in meaningful activities, were consistently high. The S– subscale, which includes concerns about job displacement, reduced human interaction, and the necessity for regulatory oversight, yielded moderate ratings with a broader distribution of responses compared to S+.

The evaluation of the developed service robot was carried out using the TUI, focusing specifically on the subscales usability, perceived usefulness, skepticism, accessibility, and the ITU scale. Participants reported high ratings on usability, which measures the extent to which the robot is perceived as user-friendly and easy to operate. Similarly, ratings on the usefulness subscale, reflecting the perceived practical benefit of the robot in supporting everyday activities, were consistently high across the sample. The skepticism subscale, which captures distrust, perceived risks, and critical attitudes toward the technology, received low ratings. In contrast, accessibility, which assesses perceived affordability and the ease with which the robot can be acquired, was rated lower compared to the other dimensions. Participants' ratings on the ITU scale, measuring the reported willingness to acquire and use the robot, were high.

Correlation analyses integrating the GAToRS demonstrated that individually positive attitudes toward robots were associated with higher perceived usefulness and accessibility, as well as a stronger intention to use the system. Individually negative attitudes were primarily linked to elevated skepticism. On the societal level, negative perceptions were associated with increased skepticism and diminished ratings of usefulness and accessibility. Notably, perceived societal optimism toward robots showed no statistically significant associations with any of the TUI dimensions.

Further correlational analyses identified multiple statistically significant linear relationships between dimensions of the TUI and the behavioral intention to use the service robot. Usability, usefulness, and accessibility were positively associated with intention to use, whereas skepticism correlated negatively.

In addition, prior experience with robotics did not exert a significant influence on either the evaluation of the developed system or the intention to use it.

Interpretation of Key Findings

The positive evaluation and high intention to use the developed service robot may be attributable to the user-centered development approach adopted. Within the framework of the RoMi project, the objective of designing the system through the active participation of nursing staff was successfully achieved. The robot was iteratively refined, both technically and with regard to its application scenarios, based on insights from prior empirical studies [19,46-49]. The early and continuous integration of end users facilitated a development trajectory that was aligned with the actual demands of everyday nursing practice, in contrast to predominantly technology-driven design strategies. The favorable assessments of usability and usefulness, as well as the pronounced intention to use the system, thus provide empirically grounded support for the effectiveness of participatory design approaches, consistent with findings from previous research [5,19,41,46].

Analyses of the TUI dimensions further demonstrate that usability, usefulness, accessibility, and skepticism all serve as strong predictors of the intention to use. These findings extend beyond the core constructs of the TAM and underscore the necessity of adapting and expanding the model for health care contexts through the inclusion of additional context-sensitive factors [29].

The strong positive correlation identified between usability and the intention to use the developed robot aligns with established findings in technology acceptance research. Usability has consistently been shown to be a critical determinant of adoption in health care contexts, particularly among nursing staff operating under time constraints [56]. Technologies that are intuitive and easy to integrate into nursing workflows are more likely to be accepted and used consistently [5,57].

XSL•FO
RenderX

Similarly, the strong positive correlation between perceived usefulness and intention to use reflects the importance of demonstrable benefits in practice. In nursing, adoption decisions are often contingent on clear advantages to patient care and workflow efficiency [28]. Robots that reduce workload and improve care outcomes are more likely to gain acceptance [21]. These findings indicate that pilot implementations and use-case demonstrations play a key role in fostering both perceived usefulness and adoption in nursing environments.

The strong positive correlation between perceived accessibility and intention to use further highlights the relevance of economic and logistical considerations in adoption decisions. The data suggest that nursing staff are more likely to engage with robotic systems when these are perceived as affordable and readily available. This finding is consistent with previous research identifying cost-related factors, such as initial investment and maintenance expenses, as potential barriers to implementation [21,58]. Strategies that improve economic accessibility, including subsidies, leasing models, or evidence of long-term cost-effectiveness, may therefore be critical to facilitate broader integration into nursing practice.

The strong negative correlation identified between skepticism and the intention to use the developed service robot underscores the importance of psychological and attitudinal barriers in shaping acceptance. In this study, skepticism was operationalized through items addressing perceived risks, potential disruption to daily routines, and a negative balance of expected benefits and drawbacks. These findings are consistent with previous research showing that concerns related to safety, reliability, and professional roles can hinder the acceptance of robotic systems [22]. Addressing such concerns requires more than technical refinement: regulatory clarity, comprehensive training, and transparent communication about system capabilities and limitations are considered essential [21,22,59], with particular emphasis on framing robotic systems as supportive rather than substitutive.

The findings related to the GAToRS clearly differentiate between societal and individual levels. The subscale S+ reflects an abstract acknowledgment of the potential societal benefits of assistive robotic systems within the surveyed sample. However, S+ is not significantly associated with the evaluation of the developed service robot or with the individual intention to use it. This dissociation may suggest that an intellectual consensus on the societal benefits of robotics is not sufficient to translate into concrete behavioral readiness within the everyday practice of nursing.

In contrast, S– exhibits a substantially stronger association with the evaluation of the service robot. The subscale is significantly negatively correlated with the perceived usefulness and accessibility of the system and positively correlated with skepticism. Similar to S+, S– does not exert a direct influence on the individual intention to use. However, this asymmetry between S+ and S– may reflect a negativity bias, whereby societal-level concerns impair specific perceptions of the robot, while positive societal-level attitudes remain largely ineffective in shaping individual evaluations [60]. Accordingly, implementation strategies should not only emphasize potential benefits but also proactively address negative societal narratives—for example, through transparent communication that highlights the robot's role in supporting nursing staff, as well as clear information on potential risks and the corresponding safety measures [59].

At the individual level, a contrasting pattern emerges. P+ is positively associated with perceived usefulness and accessibility of the service robot, and also correlates with a stronger intention to use. As expected, nursing staff who reported a generally favorable disposition toward robotics evaluated the developed service robot more positively and demonstrated a higher intention to use it in future practice.

Although scores on P– were generally low in this sample, the subscale showed a significant positive correlation with skepticism. However, no significant associations were observed with usability, perceived usefulness, accessibility, or intention to use. Once again, an asymmetry becomes apparent. While positive dispositions are associated with more favorable evaluations across multiple outcome variables, negative dispositions are primarily linked to increased skepticism. Notably, this asymmetry contrasts with the pattern observed at the societal level, where negative attitudes exert a broader influence. In this sense, the expected negativity bias appears attenuated at the individual level, suggesting that favorable personal attitudes toward robotics may be more impactful in shaping evaluative outcomes than negative ones.

In summary, the data suggest that societal-level attitudes (S+ or S–) may shape the broader cognitive context in which robotic systems are perceived, whereas individual-level factors (P+ or P–) and concrete user experiences are ultimately more decisive for actual readiness to adopt such technologies. From a practical perspective, these findings underscore the importance of proactively counteracting negative societal narratives surrounding robotics—especially by addressing concerns about job displacement and by ensuring transparent communication of potential risks and existing safety measures, particularly those concerning care recipients [21,41]. In addition, positive individual experiences with robotic systems should be facilitated at an early stage, for example, through hands-on sessions or pilot implementations at the ward level [5,41]. A combined approach that integrates risk-sensitive communication with experience-based benefit framing constitutes a promising strategy for fostering the acceptance of service robots in nursing practice.

Another noteworthy observation is the unexpectedly limited influence of prior experience with robotics on the intention to use the developed service robot. In contrast to previous studies that describe prior exposure as a facilitator of technology acceptance [31,38,44], the present results show no statistically significant differences in the evaluation of usability, usefulness, skepticism, accessibility, or intention to use. Although minor trends in the expected direction were observable, the absence of significant group differences may be attributable to several explanatory factors, including structural characteristics of the research and development project and aspects of its methodological implementation. First, the user-centered and participatory development process likely enhanced the system's

XSL•FO
RenderX

approachability across experience levels. The continuous involvement of nursing staff throughout the design and refinement phases may have resulted in a solution that is intuitively operable, useful, and closely aligned with everyday nursing practice—even for users without prior exposure. Second, the selected application scenarios should have addressed tasks that are broadly accepted and perceived as meaningful within the professional nursing context. When the service robot's intended use is seen as relevant and beneficial, prior technical familiarity may become less decisive for acceptance. Third, the standardized, hands-on training provided to all participants ensured a uniform introduction to the system and likely helped reduce disparities in user confidence and system understanding. As a controlled intervention, this training could have mitigated differences in prior experience, thereby contributing to a more consistent evaluation of the system across groups.

These considerations indicate that a lack of prior experience does not necessarily constitute a major barrier to acceptance, especially when supportive conditions are present. Such conditions may involve user-centered system and scenario design that enables experience-independent access, along with targeted training efforts to reduce uncertainty and strengthen user competence.

Limitations

The interpretations of these findings are informed by methodological considerations. The sample size of 30 participants was adequate for an exploratory study and allowed for the identification of relevant patterns. However, it inherently limits the generalizability of the results. Subtle effects, such as potential associations with prior robotic experience, may not have been captured.

Additionally, a potential selection bias cannot be ruled out. Given that participation was voluntary, individuals with a greater affinity for technology may have been overrepresented, which could have contributed to a more favorable evaluation of the developed service robot. While such bias cannot be entirely avoided in applied research, it should be considered when interpreting the observed positive ratings.

Furthermore, the controlled laboratory setting, while beneficial for standardization and internal validity, only partially captures the complexity and situational variability of everyday nursing practice. Similarly, the relatively brief interaction period with the service robot prevents conclusions about sustained use or long-term integration into professional routines.

While acknowledging the study's limitations, this paper lays a foundation for future research by demonstrating the system's readiness for field testing and offering valuable insights into the factors influencing nursing staff's acceptance of service robots.

Conclusion

This study investigated a service robot developed to support routine tasks in inpatient care. A total of 30 nurses tested the system in 3 application scenarios (information service, item delivery, and beverage delivery), alternating between the roles of nurse and care recipient. Overall, the robot was evaluated positively, with high ratings for usability and perceived usefulness, low levels of skepticism, and moderate accessibility. The reported intention to use the system was high, indicating a substantial level of acceptance and a likely readiness for practical use in the tested application scenarios. The findings thus confirm the relevance of both the selected scenarios and the robot's potential to support these tasks.

The significant correlations observed between acceptance-related dimensions and the intention to use underline the necessity of a multilayered strategy that extends beyond technical optimization. User-centered design processes, transparent communication regarding system capabilities and limitations, as well as opportunities for personal orientation and hands-on experience, are essential to align robotic systems with the practical requirements, expectations, and concerns of nursing staff.

As the field of robotics in nursing advances, these aspects should be systematically integrated into future research, development, and nursing education. Ultimately, such an approach is essential to ensuring not only that the deployment of service robots in nursing is technically feasible but also that their potential to alleviate the physical and cognitive burden on nursing staff is fully realized, thereby enabling their effective and sustainable integration into the future of nursing practice.

Acknowledgments

The authors would like to thank all study participants for their time and valuable contributions, as well as Kim Klüber, Carolin Heizmann, Dr Peter Remmers, Claudius Lotz, and Prof Dr Linda Onnasch for their support in conducting the study. The authors also thank Dr Diana Rothe and her team at pi4_robotics for their technical support throughout the study.

Regarding the preparation of the manuscript, ChatGPT based on GPT-5.2 was used solely for linguistic refinement of text drafted by the authors. It was not used to generate scientific content, analyses, interpretations, or conclusions. The authors take full responsibility for the final manuscript.

Funding

The German Federal Ministry of Education and Research funded this research within the project "*Roboterunterstützung bei Routineaufgaben zur Stärkung des Miteinanders inPflegeeinrichtungen*"—grant number 16SV8436.

XSL•FO
RenderX

Data Availability

The datasets generated and analyzed during the study are available from the corresponding author on reasonable request.

Authors' Contributions

CF and RK conceptualized the study design. CF and RK recruited the study participants. CF and RK supervised the study visits. CF and RK collected, analyzed, and interpreted the data. The manuscript draft was written by CF and RK, with critical revisions and editing provided by AH-S. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Application scenarios.
[PDF File, 224 KB - nursing_v9i1e86824_app1.pdf]

References

1. State of the world's nursing 2020: investing in education, jobs and leadership. : World Health Organization; 2020 URL: https://iris.who.int/server/api/core/bitstreams/5d564226-7418-4c10-9fb3-08d7e09626ec/content [accessed 2024-05-07]
2. Influencing health policy to protect and respect nurses' rights and improve health. : International Council of Nurses; 2023 URL: https://www.icn.ch/sites/default/files/2023-05/ICN_RA%202022_ANGLAIS.pdf [accessed 2026-03-21]
3. Trainum K, Liu J, Hauser E, Xie B. Nursing staff's perspectives of care robots for assisted living facilities: systematic literature review. JMIR Aging 2024 Sep 16;7:e58629. [doi: 10.2196/58629] [Medline: 39283666]
4. Rebitschek FG, Wagner GG. Akzeptanz von assistiven Robotern im Pflege- Und Gesundheitsbereich [Article in German]. Z Für Gerontol Geriatr 2020;53:637-643. [doi: 10.1007/s00391-020-01780-9]
5. Ohneberg C, Stöbich N, Warmbein A, et al. Assistive robotic systems in nursing care: a scoping review. BMC Nurs 2023 Mar 18;22(1):72. [doi: 10.1186/s12912-023-01230-y] [Medline: 36934280]
6. Ahn HS, Lee MH, MacDonald BA. Healthcare robot systems for a hospital environment: carebot and receptionbot. Presented at: 2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN); Aug 31 to Sep 4, 2015; Kobe, Japan p. 571-576. [doi: 10.1109/ROMAN.2015.7333621]
7. Tasaki R, Kitazaki M, Miura J, Terashima K. Prototype design of medical round supporting robot "Terapio". Presented at: 2015 IEEE International Conference on Robotics and Automation (ICRA); May 26-30, 2015. [doi: 10.1109/ICRA.2015.7139274]
8. Jacobs T, Graf B. Practical evaluation of service robots for support and routine tasks in an elderly care facility. Presented at: 2012 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO); May 21-23, 2012; Munich, Germany p. 46-49. [doi: 10.1109/ARSO.2012.6213397]
9. Dasanayake D, Gunasekara PS, Wickramasinghe HD, Fernando S, Kulasekera AL. Automated hospital ward management system interacting with mobile robot platform WDBOT. Presented at: 2018 IEEE International Conference on Mechatronics and Automation (ICMA); Aug 5-8, 2018. [doi: 10.1109/ICMA.2018.8484356]
10. Nambiappan HR, Arboleda SA, Lundberg CL, Kyrarini M, Makedon F, Gans N. MINA: a robotic assistant for hospital fetching tasks. Technologies (Basel) 2022 Mar 12;10(2):41. [doi: 10.3390/technologies10020041]
11. Mukai T, Hirano S, Nakashima H, et al. Development of a nursing-care assistant robot RIBA that can lift a human in its arms. Presented at: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems; Oct 18-22, 2010. [doi: 10.1109/IROS.2010.5651735]
12. Greenhalgh M, Landis JM, Brown J, et al. Assessment of usability and task load demand using a robot-assisted transfer device compared with a hoyer advance for dependent wheelchair transfers. Am J Phys Med Rehabil 2019 Aug;98(8):729-734. [doi: 10.1097/PHM.0000000000001176] [Medline: 31318755]
13. Beedholm K, Frederiksen K, Lomborg K. What was (also) at stake when a robot bathtub was implemented in a Danish elder center: a constructivist secondary qualitative analysis. Qual Health Res 2016 Aug;26(10):1424-1433. [doi: 10.1177/1049732315586550] [Medline: 25987583]
14. Begić A. Application of service robots for disinfection in medical institutions. In: Hadžikadić M, Avdaković S, editors. Advanced Technologies, Systems, and Applications II: Springer International Publishing; 2018:1056-1065. [doi: 10.1007/978-3-319-71321-2_89]
15. Graf B. Serviceroboter in stationären Gesundheitseinrichtungen [Article in German]. ÖKZ 2024 Nov;65(4):78. [doi: 10.1007/s43830-024-0489-3]
16. Bardaro G, Antonini A, Motta E. Robots for elderly care in the home: a landscape analysis and co-design toolkit. Int J of Soc Robotics 2022 Apr;14(3):657-681. [doi: 10.1007/s12369-021-00816-3]

XSL•FO
RenderX

17. Madi M, Nielsen S, Schweitzer M, et al. Acceptance of a robotic system for nursing care: a cross-sectional survey with professional nurses, care recipients and relatives. BMC Nurs 2024 Mar 14;23(1):179. [doi: 10.1186/s12912-024-01849-5] [Medline: 38486244]
18. Soljacic F, Law T, Chita-Tegmark M, Scheutz M. Robots in healthcare as envisioned by care professionals. Intel Serv Robotics 2024 May;17(3):685-701. [doi: 10.1007/s11370-024-00523-8]
19. Klebbe R, Friese C. Exploring the role of robots in inpatient care: caregivers´ perspectives on the development and evaluation of collaborative robot applications – a qualitative research approach. Int J of Soc Robotics 2025 Sep;17(9):1707-1724. [doi: 10.1007/s12369-025-01270-1]
20. Trainum K, Tunis R, Xie B, Hauser E. Robots in assisted living facilities: scoping review. JMIR Aging 2023 Mar 6;6:e42652. [doi: 10.2196/42652] [Medline: 36877560]
21. Servaty R, Kersten A, Brukamp K, Möhler R, Mueller M. Implementation of robotic devices in nursing care. Barriers and facilitators: an integrative review. BMJ Open 2020 Sep 21;10(9):e038650. [doi: 10.1136/bmjopen-2020-038650] [Medline: 32958491]
22. Christoforou EG, Avgousti S, Ramdani N, Novales C, Panayides AS. The upcoming role for nursing and assistive robotics: opportunities and challenges ahead. Front Digit Health 2020;2:585656. [doi: 10.3389/fdgth.2020.585656] [Medline: 34713058]
23. Dillon A, Morris MG. User acceptance of information technology: theories and models. Annu Rev Inf Sci Technol 1996;31 [FREE Full text]
24. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manage Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
25. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
26. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
27. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172. [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
28. AlQudah AA, Al-Emran M, Shaalan K. Technology acceptance in healthcare: a systematic review. Appl Sci (Basel) 2021 Nov 9;11(22):10537. [doi: 10.3390/app112210537]
29. Lee AT, Ramasamy RK, Subbarao A. Understanding psychosocial barriers to healthcare technology adoption: a review of TAM technology acceptance model and Unified Theory of Acceptance and Use of Technology and UTAUT frameworks. Healthcare (Basel) 2025 Jan 27;13(3):250. [doi: 10.3390/healthcare13030250] [Medline: 39942440]
30. Turja T, Taipale S, Kaakinen M, Oksanen A. Care workers' readiness for robotization: identifying psychological and socio-demographic determinants. Int J of Soc Robotics 2020 Jan;12(1):79-90. [doi: 10.1007/s12369-019-00544-9]
31. Turja T, Van Aerschot L, Särkikoski T, Oksanen A. Finnish healthcare professionals' attitudes towards robots: reflections on a population sample. Nurs Open 2018 Jul;5(3):300-309. [doi: 10.1002/nop2.138] [Medline: 30062023]
32. Turja T, Aaltonen I, Taipale S, Oksanen A. Robot acceptance model for care (RAM-care): a principled approach to the intention to use care robots. Inf Manag 2020 Jul;57(5):103220. [doi: 10.1016/j.im.2019.103220]
33. Rogers EM. Diffusion of Innovations, 5th edition: Free Press; 2003. URL: https://www.simonandschuster.com/books/Diffusion-of-Innovations-5th-Edition/Everett-M-Rogers/9780743222099 [accessed 2026-03-21]
34. Fishbein M, Ajzen I. Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research: Addison-Wesley; 1975. URL: https://www.researchgate.net/publication/233897090_Belief_attitude_intention_and_behaviour_An_introduction_to_theory_and_research [accessed 2026-03-21]
35. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
36. Rosenstock IM. Historical origins of the health belief model. Health Educ Monogr 1974 Dec;2(4):328-335. [doi: 10.1177/109019817400200403]
37. Honekamp I, Sauer L, Wache T, Honekamp W. Akzeptanz von Pflegerobotern im Krankenhaus: eine quantitative Studie [Article in German]. TATuP - Z Für Tech Theor Prax 2019 Jul 8;28(2):58-63. [doi: 10.14512/tatup.28.2.s58]
38. Merda M, Schmidt K, Kähler B. Pflege 4.0 – einsatz moderner technologien aus der sicht professionell pflegender [Report in German]. : Berufsgenossenschaft für Gesundheitsdienst und Wohlfahrtspflege (BGW); 2017 BGW 09-14-002 URL: https://www.bgw-online.de/resource/blob/20346/3989e17a6bd235a1637fc16a4c71d84e/bgw09-14-002-pflege-4-0-einsatz-moderner-technologien-data.pdf [accessed 2026-03-26]
39. Papadopoulos I, Koulouglioti C, Ali S. Views of nurses and other health and social care workers on the use of assistive humanoid and animal-like robots in health and social care: a scoping review. Contemp Nurse 2018;54(4-5):425-442. [doi: 10.1080/10376178.2018.1519374] [Medline: 30200824]
40. Wong KLY, Hung L, Wong J, et al. Adoption of artificial intelligence-enabled robots in long-term care homes by health care providers: scoping review. JMIR Aging 2024 Aug 27;7:e55257. [doi: 10.2196/55257] [Medline: 39190455]
41. Adeyemo A, Coffey A, Kingston L. Utilisation of robots in nursing practice: an umbrella review. BMC Nurs 2025 Mar 4;24(1):247. [doi: 10.1186/s12912-025-02842-2] [Medline: 40038679]

XSL•FO
RenderX

42. Liang HF, Wu KM, Weng CH, Hsieh HW. Nurses' views on the potential use of robots in the pediatric unit. J Pediatr Nurs 2019;47:e58-e64. [doi: 10.1016/j.pedn.2019.04.027] [Medline: 31076190]
43. Neyer FJ, Felber J, Gebhardt C. Entwicklung und Validierung einer Kurzskala zur Erfassung von Technikbereitschaft [Article in German]. Diagnostica 2012 Apr;58(2):87-99. [doi: 10.1026/0012-1924/a000067]
44. Melkas H, Hennala L, Pekkarinen S, Kyrki V. Impacts of robot implementation on care personnel and clients in elderly-care institutions. Int J Med Inform 2020 Feb;134:104041. [doi: 10.1016/j.ijmedinf.2019.104041] [Medline: 31838291]
45. Manavi M, Carros F, Schwaninger I, et al. "Making robots usable": empowering care professionals to customize social robots in care. Int J of Soc Robotics 2025 Sep;17(9):1671-1692. [doi: 10.1007/s12369-025-01238-1]
46. Klebbe R, Klüber K, Dahms R, Onnasch L. Caregivers' perspectives on human–robot collaboration in inpatient elderly care settings. Machines 2023 Dec 28;11(1):34. [doi: 10.3390/machines11010034]
47. Klüber K, Onnasch L. Appearance is not everything - preferred feature combinations for care robots. Comput Human Behav 2022 Mar;128:107128. [doi: 10.1016/j.chb.2021.107128]
48. Klüber K, Onnasch L. When robots fail—a VR investigation on caregivers' tolerance towards communication and processing failures. Robotics 2022 Oct 7;11(5):106. [doi: 10.3390/robotics11050106]
49. Klüber K, Onnasch L. Keep your distance! Assessing proxemics to virtual robots by caregivers. Presented at: HRI '23: Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction; Mar 13-16, 2023. [doi: 10.1145/3568294.3580070]
50. Koverola M, Kunnari A, Sundvall J, Laakasuo M. General Attitudes Towards Robots Scale (GAToRS): a new instrument for social surveys. Int J of Soc Robotics 2022 Sep;14(7):1559-1581. [doi: 10.1007/s12369-022-00880-3]
51. Kothgassner OD, Felnhofer A, Hauk N, Kastenhofer E, Gomm J, Kryspin-Exner I. Technology Usage Inventory (TUI) – manual. : ICARUS Research Team; 2013 URL: https://www.ffg.at/sites/default/files/allgemeine_downloads/thematische%20programme/programmdokumente/tui_manual.pdf [accessed 2026-03-21]
52. Tabachnick BG, Fidell LS. Using Multivariate Statistics, 7th edition: Pearson; 2013. URL: https://www.pearsonhighered.com/assets/preface/0/1/3/4/0134790545.pdf [accessed 2026-03-21]
53. Shrive FM, Stuart H, Quan H, Ghali WA. Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. BMC Med Res Methodol 2006 Dec 13;6(1):57. [doi: 10.1186/1471-2288-6-57] [Medline: 17166270]
54. ICH harmonised guideline: guideline for good clinical practice E6(R3). : International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; 2025 URL: https://database.ich.org/sites/default/files/ICH_E6%28R3%29_Step4_FinalGuideline_2025_0106_ErrorCorrections_2025_1024.pdf [accessed 2026-03-26]
55. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Lawrence Erlbaum Associates; 1988. [doi: 10.4324/9780203771587]
56. Babamiri M, Heidarimoghadam R, Ghasemi F, Tapak L, Mortezapour A. Insights into the relationship between usability and willingness to use a robot in the future workplaces: studying the mediating role of trust and the moderating roles of age and STARA. PLoS One 2022;17(6):e0268942. [doi: 10.1371/journal.pone.0268942] [Medline: 35657928]
57. Silvera-Tawil D. Robotics in healthcare: a survey. SN Comput Sci 2024 Jan 11;5(1):189. [doi: 10.1007/s42979-023-02551-0]
58. Yuan S, Coghlan S, Lederman R, Waycott J. Ethical design of social robots in aged care: a literature review using an ethics of care perspective. Int J of Soc Robotics 2023 Oct;15(9-10):1637-1654. [doi: 10.1007/s12369-023-01053-6]
59. Johansson-Pajala RM, Thommes K, Hoppe JA, et al. Care robot orientation: what, who and how? Potential users' perceptions. Int J of Soc Robotics 2020 Nov;12(5):1103-1117. [doi: 10.1007/s12369-020-00619-y]
60. Frank D, Chrysochou P, Mitkidis P. The paradox of technology: negativity bias in consumer adoption of innovative technologies. Psychol Mark 2023 Mar;40(3):554-566. [doi: 10.1002/mar.21740]

Abbreviations

GAToRS: General Attitudes Towards Robots Scale
ITU: intention to use
P+: personal-level positive attitudes
P–: personal-level negative attitudes
RoMi: *Roboterunterstützung bei Routineaufgaben zur Stärkung des Miteinanders in Pflegeeinrichtungen* (translation: robotic support for routine tasks to strengthen collaboration in care facilities)
S+: societal-level positive attitudes
S–: societal-level negative attitudes
TAM: technology acceptance model
TUI: Technology Usage Inventory

XSL•FO
RenderX

Edited by L Donelle; submitted 30.Oct.2025; peer-reviewed by C Corrigan, J Leoste; revised version received 09.Feb.2026; accepted 19.Feb.2026; published 14.Apr.2026.

Please cite as:
Friese C, Klebbe R, Heimann-Steinert A
Nurses' Evaluation of a Service Robot for Inpatient Care: Technology Acceptance Study
JMIR Nursing 2026;9:e86824
URL: https://nursing.jmir.org/2026/1/e86824
doi:10.2196/86824

© Christopher Friese, Robert Klebbe, Anika Heimann-Steinert. Originally published in JMIR Nursing (https://nursing.jmir.org), 14.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Virtual Nursing Pilot in the Inpatient Setting: Qualitative Evaluation

Ksenia Gorbenko[1,2,3], PhD; Maria Bailon[4], MPH; Jordan Randall[5], MS; James Viskochil[4], MLiS; Clara Cualing[4], RN; Patrick Healy[6], MBA; Madhu Mazumdar[1,3,7], PhD; Robert Freeman[4], DNP

[1]Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, One Gustav Levy Place, Box 1077, New York, NY, United States

[2]Windreich Department of Artificial Intelligence and Human Health, Icahn School of Medicine at Mount Sinai, New York, NY, United States

[3]Institute for Healthcare Delivery Science, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, NY, United States

[4]Medical/Surgical Unit, Mount Sinai Hospital, Mount Sinai Health System, New York, NY, United States

[5]Clinical Operations and Analytics, New York–Presbyterian Hospital, New York, NY, United States

[6]Technology, Innovation, and Transformation, Rush University System for Health, Chicago, IL, United States

[7]Quantitative Data Sciences, Mount Sinai Tisch Cancer Center, New York, NY, United States

Corresponding Author:
Ksenia Gorbenko, PhD
Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, One Gustav Levy Place, Box 1077, New York, NY, United States

Abstract

Background: Global nursing shortages require innovative care delivery models. Virtual nursing is a cutting-edge model being explored.

Objective: This study aimed to examine the perspectives of clinical and administrative staff involved in a virtual nursing pilot on a medical-surgical unit and to identify best practices for future adoption.

Methods: We conducted a qualitative evaluation using individual semistructured interviews with virtual and bedside nurses, nurse executives, and IT project managers implementing a virtual nursing pilot program at a medical-surgical unit with 35 private and semiprivate beds at the Mount Sinai Hospital, a 1110-bed acute care hospital in New York, NY. Interviews took place in the spring of 2024 and were completed via phone or Zoom, audio recorded, and professionally transcribed. Participants were selected using purposive sampling. The authors applied an iterative thematic analysis to transcripts using Dedoose. Claude.ai was used to generate code summaries for select codes.

Results: Of 33 individuals approached, 16 (48.5%) consented to participate. Nine participants were clinical staff (virtual and bedside nurses and nurse managers), and 7 were executive leaders or managers in nursing and informatics. Our analysis identified the following themes: (1) staff attitudes toward virtual nursing shifted from resistance to acceptance over time, (2) direct communication channels between virtual and bedside nurses were critical for efficient care coordination and model adoption, (3) admission and discharge processes evolved throughout the pilot implementation, and (4) adaptable staffing allocations were necessary to accommodate fluctuating patient census and unit demands.

Conclusions: The main beneficiaries of this intervention, bedside nurses, found their virtual counterparts helpful following a few adjustments. Participants reported a perceived reduction in administrative burden, uninterrupted completion of clinical tasks, and they felt their overtime was reduced, which all increased their buy-in for this care model. There are several opportunities for improvement, such as real-time communication, unit-specific virtual nurse training, and flexible staffing for high-volume units. Our findings suggest that virtual nursing can address staffing challenges. Calculating the true return on investment for virtual nursing programs will require comprehensive mixed methods evaluations of such outcomes as care team and patient satisfaction, length of stay, readmission prevention, completion of nursing tasks, and reduction in overtime.

(JMIR Nursing 2026;9:e76994) doi:10.2196/76994

KEYWORDS

virtual nursing; evaluation; qualitative; systems integration; change management; digital transformation; care delivery model; care coordination; staff satisfaction

XSL•FO
RenderX

Introduction

Technological advancements, staffing challenges, and evolving patient expectations are driving a digital transformation that is rewriting how we deliver health care. The COVID-19 pandemic accelerated the adoption of electronic communication platforms to help reduce the spread of the deadly virus. A workforce crisis has emerged globally with the United States facing a shortage of between 1.8 and 3 million nurses according to different projections [1,2]. The excessive workload placed on existing health care personnel has resulted in heightened stress, burnout, psychosomatic disorders, anger, and anxiety [1]. To address this challenge, health care leaders are exploring a variety of digital health tools and integrated care models. The use of telehealth for outpatient visits increased 38 times in the first 2 years of the pandemic [3]. The use of digital tools also increased in the inpatient setting [4]. The majority of these interventions have involved a physician connecting with a patient or bedside staff (a nurse) in a patient's home or a clinical setting [5,6].

The pandemic's toll on the health care workforce, nurses in particular, has put new pressure on hospitals to identify nursing tasks that could be completed remotely. Among these are administrative tasks such as admissions and discharges [7], patient education [8], and 1:1 sitting [9]. Of these, virtual sitting for fall prevention has the most evidence, with 12 studies conducted between 2009 and 2021 [9]. Using virtual nurses for admissions and discharges is more recent, with only three virtual nursing pilots on the topic published to date [7,8,10]. Other formats of virtual nursing have also been explored. For example, colleagues from the Mayo Clinic reported on an inpatient virtual registered nurse (RN) program, in which experienced nurses (often retired or nearing retirement) provided guidance to bedside RNs and proactively monitored patients, with no difference in patient safety and care outcomes [11,12].

In a collaborative effort to address staffing challenges and improve patient care, the Mount Sinai Hospital (MSH) partnered with Banyan Medical Solutions, a third-party vendor, to implement a virtual nursing platform on a selected inpatient unit. The implementation process involved installing compatible large-screen smart TVs, wall-mounted cameras, and pillow speakers, enabling virtual nurses to perform critical roles such as patient education, admissions, and discharges. This qualitative research study aimed to examine partner perspectives on this virtual nursing pilot at MSH with a focus on operational implementation and to identify best practices and recommendations for further refinement.

Methods

Setting

In January 2024, a virtual nursing pilot was initiated at a 35-bed medical-surgical unit at MSH. The fast-paced unit with high daily discharges accepts various postoperative patients. Virtual nurses, credentialed by the state but working remotely from a command center, have full access to the hospital's electronic health records. These nurses joined the company in 2021 2022, are relatively junior with experience in other settings, and chose virtual work for personal reasons. Privacy protocols included "virtual knock" protocol, verbal consent requirements, and headphone availability for patients in semiprivate rooms.

Study Design and Sampling

We conducted a qualitative study using semistructured interviews to examine staff and leaders' perspectives and experiences related to the virtual nursing pilot program. We used purposive sampling to recruit participants from diverse professional backgrounds, aiming to gain a comprehensive understanding of the program from the frontline to leadership viewpoints.

Recruitment and Participants

Between February and June 2024, we invited 33 individuals to participate through email outreach, in-person shadowing, distributing flyers in common areas, and collaborating with hospital leadership. Data saturation was reached after completing about two-thirds of the nurse interviews, when new interviews yielded little additional information, following accepted standards in qualitative research [13]. We continued to recruit nonclinical personnel to ensure diverse perspectives.

Data Collection

Semistructured interviews were conducted via Zoom using a MSH account, lasting about 30 minutes each. Topics included patient and staff safety, privacy, technology reliability, program acceptability and usefulness, and improvement opportunities. The interview guide was developed based on research team discussions, particularly input from nursing leaders attuned to staff hopes and expectations. Our interview guide did not specifically explore formal clinical governance structures, focusing instead on operational implementation experiences. Using an interview guide helped ensure consistency and reliability of collected data across participants [14]. Interviewers (JR, MB, KG) used probes for clarification and conducted clinical shadowing (JR, MB) with field notes.

Data Analysis and the Use of Artificial Intelligence

Interview data were professionally transcribed and analyzed using iterative thematic analysis [15]. The team used Dedoose [16] for data analysis and management, and Claude.ai [17] to generate drafts of data summaries for selected codes. Three researchers (KG, MB, JR) read and coded the transcripts manually, using both deductive and inductive codes, refining codes through weekly discussion. Two coders applied the final codebook to each transcript in Dedoose. Major codes included care coordination, impact on workload, opportunities for improvement, and inclusion in decision-making. Minor codes included staff satisfaction, timing, best practices, and limitations of virtual nursing. After coding, the team identified key themes, downloaded relevant excerpts, and used AI to generate comparative summaries for each subtheme. Three analysts (KG, JR, MB) combined, iterated, and edited these summaries to create drafts of Results sections for further team refinement.

Ethical Considerations

The study was determined to be exempt by the Program for Protection of Human Subjects at the Icahn School of Medicine at Mount Sinai on January 17, 2024 (STUDY-23-01378). Informed consent was obtained after the nature and possible

XSL•FO
RenderX

consequences of the study were explained. The study adhered to local, national, regional, and international law and regulations regarding protection of personal information, privacy, and human rights. The study was deemed exempt by MSH's Human Subjects Protection Program, with all participants consenting to be interviewed and recorded, and receiving a US $25 gift card as compensation.

Results

Participants

Of the 33 individuals invited, 16 (48.5%) staff members consented and completed interviews (Table 1). The participants included 9 clinical staff (RNs, nurse managers, and virtual nurses) and 7 nonclinical personnel (executives in nursing and informatics, and IT managers). Based on the analysis of the data, we identified the following 4 major themes: (1) staff attitudes toward virtual nursing shifted from resistance to acceptance over time; (2) direct communication channels between virtual and bedside nurses were critical for efficient care coordination and model adoption; (3) admission and discharge processes evolved throughout the pilot implementation; and (4) adaptable staffing allocations were necessary to accommodate fluctuating patient census and unit demands.

Table . Participant characteristics and professional background (N=16).

Participant	Count, n
Professional background	
Registered nurse, including virtual nurses and executives with nursing background	12
Information technology or managers	3
Executives, including MSHS[a] executives and Banyan	2
Professional role	
Registered nurses at bedside	7
Virtual nurses	3
Information Technology	2
Operations personnel	2
Executives, including MSHS executives and Banyan	2
Gender	
Female	14
Male	2

[a]MSHS: Mount Sinai Health System.

Theme 1: Implementation Process: Staff Perspectives Adapt From Resistance to Acceptance

The implementation of virtual nursing revealed a clear evolution in staff attitudes, moving from initial skepticism to overall acceptance. Some of it had to do with the timing and depth of staff involvement in virtual nursing implementation decisions. Staff accounts indicate early concerns about role displacement and workflow disruption (Table 2, example 1a), which were exacerbated by the fact that early-stage participation was limited and staff were brought in late in the planning process. Staff identified missed opportunities for leadership to proactively engage frontline workers: "I would have loved to have had that interaction with Banyan earlier, because we had a bunch of questions that our people didn't have answers to" (Bedside RN, CL01; Table 2, example 1b). This delayed engagement led to accumulated questions and uncertainty among staff, though many issues were eventually resolved through concentrated information sessions. Nurse participants also reported that once the training was complete, they were relieved and realized that virtual nurses could be helpful (Table 2, example 1c).

XSL•FO
RenderX

Table . Key themes and representative quotes from participant interviews. Theme 1: Implementation process.

Example No.	Subtheme	Quote and participant role, ID
1a	Initial fears	“So there was a lot of trepidation. what is this actually going to look like. like physically how are they going to do this? And they’d be like, oh, we’ll get back to you.” (Bedside RN, CL01)
1b	Staff involvement	“.if the leadership is like. we’d like to go talk to the staff before we officially decide this so that we can let them know exactly how this works, maybe they’ll be a little bit more comfortable with it” (Bedside RN, CL02, Table 2, 1b).
1c	Relief	"We were thinking it was going to be a thing that we would have to worry about. But once you know all you have to do is a few clicks in the Banyan board, then I think that is definitely really helpful.” (Bedside RN, CL03)
1d	Decisions made at leadership level	“Because we are a union hospital, we did have to be very specific. That came from the CNO and the union.” (Nursing executive, NCL06)
1e	Leadership and change management	"It requires multiple conversations and sessions in advance of the model being implemented for people to first understand why we’re doing this. It's not for the heck of it. It's to help you" (Nursing executive, NCL07)
1f	Peer-to-peer learning	"I think if also there was an opportunity for people from [our unit] to go talk to the other units about it and how it worked, like to hear from peer-to-peer, that would be really helpful" (Bedside RN, CL03).

Several bedside nurses noted they wanted to be involved in technology evaluation and workflow decisions. This, however, was not always possible, as we learned from the interviews with nursing executives and nurse managers. One consideration was that the union leaders’ agreement on what tasks would be delegated to virtual nurses. Union considerations shaped decision-making authority, with decisions made at the leadership level (Table 2, examples 1c and 1d).

Our analysis also demonstrated real-world challenges in achieving meaningful staff participation while maintaining unit operations. Leaders described efforts to include staff while acknowledging limitations: "We do pull nursing staff off the floor to assist with any projects to get their insight, but we often know that they’re not able to stay every time for the whole hour, if at all” (Executive leader, NCL06). This resulted in a selective approach to staff involvement, with few staff members being able to go onsite to another hospital to observe Banyan operations.

During interviews, leaders emphasized their commitment to address staff concerns and support adoption of the new system. As one operations leader explained, "I would do anything I could to get leadership to the unit, to recognize that we appreciate you... we’re doing this as a resource for you” (Executive leader, NCL04). The implementation process required careful messaging and consistent support (Table 2, example 1e).

Finally, peer influence played a crucial role in shifting perceptions, particularly as floating nurses shared positive experiences across units: "When we get nurses from other units that float to us, they’re like oh my god, this is amazing. They all love it” (Bedside RN, CL06). Initial resistance transformed through hands-on experience and peer testimonials, with staff gradually recognizing virtual nursing as supportive rather than replacement technology (Table 2, example 1f).

Theme 2: Optimizing Technology Integration: Opportunities and Best Practices

System reliability and integration emerged as critical factors in implementing a virtual nursing platform. Here we describe *several operational opportunities for improvement* that can be used to establish best practices for implementation. These included technical integration to eliminate delays in communication, resolving audio quality issues, remote printing capability for virtual nurses, and volume control in shared rooms. These technical challenges collectively highlight the importance of robust infrastructure and seamless integration between systems for optimal integration into operational workflows.

Integration with existing hospital systems presented ongoing challenges. Multiple nurses noted issues with electronic health record connectivity and documentation access. These integration issues required bedside staff to navigate between multiple platforms, creating workflow inefficiencies (Table 3, example 2a). Bedside nurses desired a more nuanced communication tool with their virtual nurse partners in which they could specify discharge timing, for example, when transportation for the patient was already prearranged, and a direct communication

XSL•FO
RenderX

line between virtual nurses and providers. This highlighted the need for systems that could accommodate real-time coordination between virtual and bedside teams.

Table . Key themes and representative quotes from participant interviews. Theme 2: Optimizing technology integration.

Example No.	Subtheme	Quote and participant role, ID
2a	Systems integration	"Opening a separate screen, it sounds crazy, but it's often a pain... We basically have to open up a separate page. We have to go out of our Epic and open up a separate page to move Banyan." (Bedside RN, CL04)
2b	Audio quality	"It's loud... depending on the age group, depending on how sedated the patient is... Some of it is a [local] issue, that some of the TVs aren't coming through the call bell." (Bedside RN, CL04)
2c	Printing connectivity	"...if [virtual nurses] have completed all the discharge… it's just… one extra button to print it. Whereas if we have to print it, we have to go in, we have to look at all the things, then print it, and then gather it again. So it just saves a couple extra minutes with it being like oh, you're already in that section, you press the button." (Bedside RN, CL02)
2d	Patients with special accommodations	"So at that point I would probably do the interpretive services myself in person because on top of it that there's a new barrier of virtual, we have to then find an interpreter, and then we have the technology issues of whether the phone in the room is actually working... But even when it is working, I've seen that admission, instead of taking probably like a half an hour, it takes like over an hour." (Bedside RN, CL05)
2e	Workarounds	"And they themselves, taxi services, most of the time they're rude… they're like no, if they're not here in the next two minutes I have to leave and pick up the next person. They don't care. And then we're like let's just do the discharge ourselves and let's just… like let's not even wait for transport, let's just roll the patient down in the wheelchair on our own before they lose their taxi." (Bedside RN, CL05)

The interoperability challenges between the virtual nursing platform and the hospital's EHR system affected *care plan documentation and continuity*. Virtual nurses documented within both the Banyan system and Epic, but the lack of seamless integration meant that care plan updates were not always immediately visible to bedside staff without actively switching between systems. When virtual nurses completed admission assessments or discharge education, bedside nurses reported needing to actively seek out these updates rather than receiving automatic notifications of completed tasks. Remote printing connectivity for virtual nurses represented another technical hurdle. Bedside nurses explained that efficiency could be gained through improved printing functionality (Table 3, example 2c). According to one nurse, "it just saves a couple extra minutes… [if] you're already in that section, you press the button" (Bedside RN, CL02). While virtual nurses tried to send paperwork to print locally on the unit, this functionality did not work at the time of the interviews, making bedside nurses sign into the electronic medical record each time for the (virtually) discharged patient, find all the information, and send it to the printer again.

Audio quality issues particularly affected patients' willingness to accept a virtual nurse and staff efficiency, with problems being more pronounced during night shifts. Volume control concerns in shared rooms further complicated the situation, creating potential privacy issues and patient dissatisfaction, highlighting the need for better audio privacy solutions such as mandatory headphone use or room-specific volume controls (Table 3, example 2b).

While adaptations in the virtual nursing program showed promise for standard patient interactions, additional complexities arose with specific patient populations. Extended processing times for patients requiring interpreter services emerged as a particular challenge. Our bedside nurse participants reported foregoing virtual support altogether to save time and ensure effective communication with patients who needed interpreter services (Table 3, example 2d).

In response to these challenges, staff developed various *workarounds* to maintain efficient patient care. For example, bedside nurses described how transportation constraints sometimes necessitated bypassing virtual nursing protocols.

XSL•FO
RenderX

One nurse explained that taxi drivers often gave ultimatums about immediate departure, stating they would leave if patients weren't ready. This led staff to expedite discharges independently rather than risk patients losing their transportation (Table 3, example 2e). This adaptive approach allowed staff to balance the benefits of virtual nursing with practical operational needs. Another workaround—the implementation of Epic's secure chat functionality—marked a significant improvement in real-time coordination, with staff reporting it helped them coordinate plans with their virtual counterparts. Developing clear communication protocols, particularly for urgent situations, proved essential for maintaining seamless patient care across virtual and bedside teams.

IT specialists in our study highlighted the importance of technical considerations, such as future *interoperability*, beyond the immediate virtual nursing pilot. One IT leader pointed out that it is important to consider the use/reuse of technology components, such as smart TVs, for new nursing workflows that could be covered remotely: "how do we utilize a single or a smaller set of devices?" (NCL03). This perspective underscores the importance of strategic planning for technology integration that considers both current needs and future applications.

Theme 3: Care Coordination: Benefits of Virtual Nursing on Teamwork

The implementation of virtual nursing transformed admission and discharge processes, creating new team dynamics between bedside and virtual nurses. This theme explores how these workflows evolved and the benefits that emerged from this collaborative approach.

Transformation of the Discharge Process

Before the virtual nursing pilot, discharges required bedside nurses to manage multiple competing priorities simultaneously. One nurse described the traditional discharge process prior to the virtual nursing pilot implementation that involved educating patients, who were often frightened, about their discharge, trying to be supportive, "all while maybe a call bell's ringing, you're trying to get somebody their meds, you're hopefully not getting an admission because your workload is full" (Bedside RN, CL04; Table 4, example 3a). With the introduction of virtual nursing, the process became more streamlined. The same nurse explained that now they just switch a certain setting and "Banyan will often just start doing the discharges" (Bedside RN, CL04), freeing bedside staff to perform other tasks.

XSL•FO
RenderX

Table . Key themes and representative quotes from participant interviews. Theme 3: Care coordination.

Example No.	Subtheme	Quote and participant role, ID
3a	Old workflows (previrtual nursing)	"…prior to Banyan, you waited for a discharge order… You printed out your paper and we went to the bedside… we're a postsurgical unit, many people are absolutely frightened about going home with a new hip… and you would spend a tremendous amount of time educating, and teaching, and explaining their discharge. Then we wait for transport. And this is all while maybe a call bell's ringing, you're trying to get somebody their meds, you're hopefully not getting an admission because your workload is full." (Bedside RN, CL04)
3b	Good catch	"The only time I remember that it didn't go well is we had a patient who needed… [Lovenox] education... And the only way I was able to figure that out is because he was leaving and then his wife came up to me and she was like we never got the Lovenox education… So I think the nurse should have picked it up, like oh, you're going home on Lovenox injection, do you know how to administer it?.. He could have just walked out, went home, and not know how to do it. Good thing his wife [asked]… And I was like oh, okay, let's go back into the room and we will go through the steps on how to administer the Lovenox injection." (Bedside RN, CL06)
3c	Division of patient education responsibilities	"Especially with our ortho patients, they are usually prescribed a certain blood thinner so that they don't develop any clots, and if they're already taking blood thinners at home, and there's an issue with the medication or any interaction, they could definitely help somebody, you know, prevent them from having an adverse reaction..." (Bedside RN, CL05) "The hands-on teaching, plus the discharge, can take a long time. So for the discharge paperwork to be reviewed, and then I come in and just have to do Foley teaching, for example, with the leg bag, that just takes a lot of time off of my plate for me to focus on other things." (Bedside RN, CL03)
3d	No direct communication with providers	"[Virtual nurses] are not able to communicate with the frontline providers, and if… the patient has any questions, and they can't answer, and then they'll message us, and then I have to message the doctor, and then I have to relay that information to the virtual nurse… So, you know what I mean, so she still has to go through me…" (Bedside RN, CL06)

The discharge process transformation remains a work in progress, particularly regarding the division of educational responsibilities between virtual and bedside nurses. This was evident when a patient nearly left without critical Lovenox self-injection education, discovered only when his wife proactively inquired as they were leaving (Table 4, example 3B). Following this learning opportunity, the nurse implemented a new protocol of personally reviewing the After Visit Summary before handoff, proactively identifying teaching responsibilities that fall within the bedside nurse's scope rather than the virtual nurse's capabilities, thus ensuring patients receive all necessary education before discharge.

Emerging Communication Patterns and Task Coordination

As the unit grew more comfortable with virtual nursing, new communication patterns emerged between bedside and virtual nurses. Bedside nurses noticed that virtual nurses began to proactively update them on task completion. Virtual nurses reported developing strategies for prioritization through communication with bedside staff. As one virtual nurse explained: "Communication is key. You just let the bedside nurse know what's going on… could you please… let us know the order that you want us to prioritize the patients, whose transportation is coming first… It makes all the difference in

XSL•FO
RenderX

care" (Virtual RN, CL08). Virtual nurses reported initial lack of clarity about their role among patients, with one participant noticing an improvement in that respect: "the patients [are] just being educated more… regarding our services" (Virtual RN, CL07).

The division of educational responsibilities between virtual and bedside teams emerged as a key strength of the program. The system proved particularly valuable for complex cases, including newly diagnosed patients requiring extended education sessions. One bedside nurse highlighted the importance of virtual nurses' role in patient education, particularly for high-risk patients. She explained how orthopedic patients on blood thinners need to have their medications at home reviewed in case they were already taking blood thinners to prevent any interactions or adverse reactions (Table 4, example 3c).

A gap that remained unresolved at the time of the interviews had to do with virtual nurses' inability to message providers. For example, if a patient had a question during discharge that the virtual nurse could not answer, they had to message a bedside nurse who would then message a provider and report back to the virtual nurse who would relay it to the patient (Table 4, example 3d). Being unable to communicate directly with providers made this process unnecessarily involved and inefficient.

Time Efficiency and Focus on Direct Patient Care

A significant benefit reported by bedside nurses was how virtual nursing saved them time and allowed them to focus on hands-on teaching and direct patient care. One bedside nurse mentioned that discharges traditionally could take up to an hour and were now taken care of by their virtual counterparts (Table 4, example 3c). Another nurse echoed: "We're not stuck in a room for an hour doing admission questions and then having four other patients call on the Vocera and then get interrupted, going back and forth. That's what was happening before" (Bedside RN, CL06).

The ability to handle multiple tasks simultaneously was particularly valuable during busy periods. As one nurse explained: "Because if you're getting two admissions and you have a discharge, it's just a lot… So it's really helpful because then you're able to actually focus on hands-on patient care as opposed to…the paperwork aspect of it" (Bedside RN, CL03).

Improved Nursing Satisfaction and Work-Life Balance

The collaboration between virtual and bedside nursing staff resulted in positive outcomes for staff satisfaction and work-life balance. One nurse remarked, "It's nice to just like have that thought taken off your plate, as well as the admissions, just not having to stay late to do an admission when they come at… poor timing." (Bedside RN, CL02). Overall, the evolution of teamwork between virtual and bedside nurses created a more efficient workflow for admissions and discharges, allowing nurses to dedicate more time to direct patient care while ensuring comprehensive education and documentation.

Theme 4: Timing Is Key: Staffing Capacity and Patient Flow

Workflow integration and staffing alignment emerged as critical factors in the uptake and efficiency of virtual nursing, revealing the importance of proper coordination between virtual and bedside teams in managing patient flow.

Managing High-Volume Patient Movement

The partnership between virtual and bedside teams impacted patient workflow management, particularly during high-volume periods. One bedside nurse explained that their unit had "a colossal amount of movement… [with] 14 to 20 [daily] discharges and admissions" (CL04). Effective workforce distribution proved essential for managing such volume, as staff needed to adapt to fluctuating demands throughout the day. As another bedside nurse explained: "Sometimes they could say 10 [discharges], so they have, let's say, 5 [virtual] nurses on board, but then as the day goes on, the discharges do happen, so instead of having 10 we'll have 18" (Bedside nurse, CL06). This unpredictability highlighted the ongoing need to develop flexible staffing management strategies that could respond to rapidly changing conditions. Ascertaining the right staffing levels for virtual took some tweaking, as is evidenced by one bedside nurse report about initial challenges during the overnight shift, when there was only one virtual nurse available (Table 5, example 4A).

XSL•FO
RenderX

Table . Key themes and representative quotes from participant interviews. Theme 4: Timing is key.

Example No.	Subtheme	Quote and participant role, ID
4a	Initial insufficient staffing	"…it also comes down to how many people, how many nurses is Banyan hiring for our unit per shift... Because the thing is, [my supervisor] said oh, we noticed that things are running slow. [The virtual nurse said] I'm sorry, I apologize, it's just [I am] the only nurse with them at that shift… That was when Banyan was first implemented… I actually think Banyan can work in a fast-paced, but if there's enough staff, virtual nurse staff, and also to make things run faster." (Bedside RN, CL05)
4b	Impact on hospital throughput	"Once we discharge this patient we can pull more patients coming from the PACU and from the ED, so it helps them, like don't hold the OR schedule nor have a critical diversion in the ED" (Bedside RN, NCL05).
4c	Nurses more readily accept admissions at end of shift	"I think now that we have the Banyan nurses, nurses are not... reluctant on taking the admission from PACU. They'll just pick up because they know that when they come onto the floor that they don't have to deal with these admission questions, that they're going to be taken care of" (Bedside RN, CL06).

Impact on Hospital-Wide Operations

The virtual nursing system's impact extended beyond individual patient care to affect broader hospital operations. The most positive impact of the pilot, as perceived by our participants, was what they described as improved patient throughput. During peak periods, virtual nurses supported bedside teams, creating capacity to process patient admissions and discharges simultaneously. High volume peaks required different coordination compared to slower-paced environments, while night shift operations needed special adaptations to maintain patient comfort and prevent disruption during quiet hours. According to one nurse, by facilitating more efficient discharges, the system improved patient throughput across multiple departments (Table 5, example 4b).

This improvement in patient flow was also reflected in staff attitudes toward accepting new admissions. Nurses used to be reluctant to accept an admission from PACU, for example, if it meant they would have to stay longer after the end of their shift to complete it. According to several participants, this was no longer the case (Table 5, example 4c).

Time Investment in Patient Education

Virtual nurses reported dedicating significant time to patient education, highlighting the value of this investment despite its potential impact on short-term throughput. One virtual nurse reflected on the depth of engagement required: "I would say I think I'm long with discharges personally… I take a while. They have so many questions… The family's in there, the mom, the dad. So yeah, I feel like my discharges are long, like an hour… I feel bad. Some of them are newly diagnosed, so that's even longer, like an hour and a half" (Virtual RN, CL09). This time investment, while potentially extending the discharge process itself, ultimately supported more comprehensive patient education and potentially better postdischarge outcomes, demonstrating that flexibility in staffing and timing considerations needs to balance efficiency with quality of care.

Discussion

Principal Findings

This is the first qualitative evaluation of a virtual nursing pilot on an inpatient unit. Our findings demonstrate that virtual nursing implementation requires careful attention to technological infrastructure, operational readiness, and change management. We identified four major themes that characterized the virtual nursing implementation in our pilot: (1) achieving staff buy-in takes time and can benefit from earlier engagement strategies; (2) seamless technology integration, especially direct communication channels, is critical for smooth operation; (3) recognition of virtual nurses as integral team members, rather than auxiliary support, enhances collaborative care delivery; and (4) virtual nursing coverage must meet the demand to maintain relevance within the organizational or unit context.

Our study offers insights into implementing virtual nursing on inpatient units, which other hospitals can use when considering similar models. The unique unit or patient population characteristics—high volume, fast pace, surgical patients with diverse diagnoses—fit well with the virtual nursing pilot seeing benefits in admission or discharge support and patient education. If virtual support was unavailable or inefficient, the bedside nurses reverted to their traditional workflows. With the exception of some initial issues, the vendor was able to meet the challenge and provide sufficient support. Virtual nursing implementation on slower-paced units might prioritize functions like virtual sitters, which could be particularly useful on observation units.

XSL•FO
RenderX

Our findings highlight the critical importance of seamless EHR integration for virtual nursing models. The fragmented documentation workflow we observed, where staff navigated between multiple systems, created inefficiencies that could impact care continuity. The care plan coordination issues, exemplified by the Lovenox education incident, demonstrate how documentation gaps can directly affect patient safety. Virtual nursing programs must establish clear protocols for documenting care plan handoffs, particularly for complex patients requiring both virtual and bedside interventions and prioritize robust interoperability that allows direct documentation within the primary care record.

Technical and system reliability directly impacted operation readiness. The integration between EHR and the virtual nursing platform, audio quality, and physical environment considerations influenced staff adoption and clinical workflows. Virtual nursing adoption hinges on infrastructure investment that could be a barrier for smaller hospitals.

The pilot used out-of-state junior RNs. Other staffing models could be considered, such as floating bedside nurses as virtual for a fraction of their time. Implementers should weigh the efforts needed to educate new virtual staff vis-à-vis managing virtual nursing in-house.

Flexible staffing in response to patient flow emerged as one of the most critical factors for optimal operation. Implementation followed an overall positive trend—initial increases in workload during rollout, followed by productivity improvements once frontline staff adapted to their new workflow. Participants reported positive downstream effects such as reducing operating room holds and emergency department diversions through improved time to discharge and bed turnover.

This pilot aligns well with the cutting-edge innovations in health care operations. Our findings demonstrate that implementation depends on structured approaches to change management [18] while minimizing disruption to care delivery. Key insights from our analysis—albeit not new [19]—reveal the importance of early staff involvement in the decision-making process.

Staff adoption of the virtual nursing platform followed a progression from initial skepticism to approval. Clinical champions were powerful influencers of technology acceptance. Clear role separation created effective partnerships, particularly in educational delivery. Virtual nursing could leverage experienced nurses' expertise post COVID-19 pandemic, like programs at other institutions have shown [11,12].

Limitations and Strengths

Our study has the following limitations. We interviewed hospital personnel but not patients or caregivers. Understanding how both patients and caregivers view this intervention is essential for improving it further. Researchers need to examine not just how the interactions between patients and virtual nurses are designed to work, but also how patients and caregivers actually experience and respond to these interactions. It is a single-center pilot study in an academic medical center; our findings might not apply to a different health care setting or patient population. With about a 50% response rate, a nonrespondent bias is possible. Our study did not include quantitative data on reduced overtime, improved throughput, or quality indicators such as patient satisfaction, medication errors, or readmission rates. All comments related to these important metrics were based on the impressions of bedside and virtual staff and executive leaders. We also did not systematically examine formal clinical governance models or committee structures that may have guided the implementation process, nor did we evaluate preimplementation preparation or formal support models during rollout. However, this is the first published qualitative study of virtual nursing programs to our knowledge, capturing valuable nurse and leadership perspectives through its qualitative approach.

Recommendations for Planning Virtual Nursing Implementation

Based on our findings, implementation of virtual nursing requires attention across several domains. Organizations considering adoption should prioritize early infrastructure investment, unit-specific workflow design, clear governance structures, and intentional patient-facing communication. The following checklist sums up key considerations to guide planning and implementation.

- Technical infrastructure
- Invest in technical and communication infrastructure upfront to ensure seamless system integration between virtual and bedside nurses and between virtual nurses and providers
- Develop or adopt AI tools to support automatic patient prioritization for admission and discharge

- Workflow
- Implement dynamic staffing allocations responsive to patient volume fluctuations
- Adapt and optimize workflows to reflect specific demands of each unit
- Develop clear standard operating procedures for communication between virtual and bedside staff
- Conduct regular review of staffing allocations based on utilization patterns

- Clinical governance
- Establish formal governance structures with designated clinical champions

- Patient-facing communication
- Use nurses' insights to create standard messaging on introducing the virtual nursing program to patients and setting expectations
- Develop privacy standard operating procedures for shared rooms, including headphone distribution and consent procedures for virtual interactions

Conclusions

Virtual nursing is a cutting-edge care delivery innovation that goes beyond adoption of new technology. This virtual team model enhances care coordination, patient throughput, and staffing while keeping care patient-focused. Program success requires resources, technical infrastructure, and clear, unit-specific communication. Health care should see it as bedside care extension, with careful planning and ongoing adjustments based on data and feedback.

XSL•FO
RenderX

This study provides valuable insights for hospital leadership and management to understand the impact on clinical workload, staff acceptability and satisfaction, and patient care quality to support informed implementation decisions. Future research should explore several questions that emerged during this implementation: (1) How do virtual nursing models impact nurse retention and recruitment in a high turnover environment? (2) How might AI integration enhance the efficiency of virtual nursing documentation and communication? (3) How do patients perceive virtual nursing and how can their experience be improved? (4) What does virtual nursing affect quality and safety outcomes such as patient satisfaction, readmissions, and medication errors? Addressing these questions will further develop the evidence for this emerging care delivery model.

Funding

KG received salary support from the Office of Clinical Innovation to help fund this study.

Conflicts of Interest

The authors have the following conflicts to disclose. KG received salary support from the Office of Clinical Innovation. MB and JR are part of the Digital Experience team in the Office of Clinical Innovation.

References

1. Tamata AT, Mohammadnezhad M. A systematic review study on the factors affecting shortage of nursing workforce in the hospitals. Nurs Open 2023 Mar;10(3):1247-1257. [doi: 10.1002/nop2.1434] [Medline: 36303066]
2. Scheffler RM, Arnold DR. Projecting shortages and surpluses of doctors and nurses in the OECD: what looms ahead. Health Econ Policy Law 2019 Apr;14(2):274-290. [doi: 10.1017/S174413311700055X] [Medline: 29357954]
3. Bestsennyy O, Gilbert G, Harris A, Rost J. Telehealth: a quarter-million-dollar post-COVID-19 reality? : McKinsey & Company; 2020 May URL: https://www.mckinsey.com/~/media/McKinsey/Industries/Healthcare%20Systems%20and%20Services/Our%20Insights/Telehealth%20A%20quarter%20trillion%20dollar%20post%20COVID%2019%20reality/Telehealth-A-quarter-trilliondollar-post-COVID-19-reality.pdf [accessed 2026-02-12]
4. Gorbenko K, Mohammed A, Ezenwafor EII, et al. Innovating in a crisis: a qualitative evaluation of a hospital and Google partnership to implement a COVID-19 inpatient video monitoring program. J Am Med Inform Assoc 2022 Aug 16;29(9):1618-1630. [doi: 10.1093/jamia/ocac081] [Medline: 35595236]
5. Lepre B, Job J, Martin Z, Kerrigan N, Jackson C. The Queensland Virtual Integrated Practice (VIP) partnership program pilot study: an Australian-first model of care to support rural general practice. BMC Health Serv Res 2023 Oct 31;23(1):1183. [doi: 10.1186/s12913-023-10189-0] [Medline: 37907917]
6. Buchanan S, Peixoto C, Belanger C, Archibald D, Bjerre L, Fitzsimon J. Investigating patient experience, satisfaction, and trust in an Integrated Virtual Care (IVC) model: a cross-sectional survey. Ann Fam Med 2023;21(4):338-340. [doi: 10.1370/afm.2978] [Medline: 36944507]
7. Roberson AE, Martinez HR, Holligan SA, Carlson M. An evolving virtual nursing model: enhancing care delivery through advanced workflow integration. Nurs Outlook 2025;73(3):102421. [doi: 10.1016/j.outlook.2025.102421] [Medline: 40252518]
8. Ransford J, Tidwell T, Johnson L, Gitney N, Morgan A, Hauch R. Implementing a virtual discharge nurse pilot: utilizing the Pathway to Excellence® Framework. J Nurs Adm 2024 Nov 1;54(11):605-611. [doi: 10.1097/NNA.0000000000001498] [Medline: 39475890]
9. Hogan Quigley B, Renz SM, Bradway C. Fall prevention and injury reduction utilizing virtual sitters in hospitalized patients: a literature review. Comput Inform Nurs 2021 May 28;39(12):929-934. [doi: 10.1097/CIN.0000000000000773] [Medline: 34050057]
10. Schwartz RL, Hamlin SK, Vozzella GM, et al. Utilizing telenursing to supplement acute care nursing in an era of workforce shortages: a feasibility pilot. Comput Inform Nurs 2024 Feb 1;42(2):151-157. [doi: 10.1097/CIN.0000000000001097] [Medline: 38252545]
11. Roberson AE, Carlson M, Kohler CM, Harris PA, Volkmann CL. Initiating virtual nursing in general inpatient care. Am J Nurs 2023 Jun 1;123(6):48-54. [doi: 10.1097/01.NAJ.0000938736.42266.5e] [Medline: 37233140]
12. Savitz ST, Frederick RK, Sangaralingham LR, et al. Evaluation of safety and care outcomes after the introduction of a virtual registered nurse model. Health Serv Res 2023 Oct;58(5):999-1013. [doi: 10.1111/1475-6773.14208] [Medline: 37525521]
13. Guest G, Bunce A, Johnson L. How many interviews are enough?: an experiment with data saturation and variability. Field Methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
14. Colorafi KJ, Evans B. Qualitative descriptive methods in health science research. HERD 2016 Jul;9(4):16-25. [doi: 10.1177/1937586715614171] [Medline: 26791375]

XSL•FO
RenderX

15. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
16. Dedoose. 2021. URL: https://www.dedoose.com/ [accessed 2026-02-12]
17. Claude. 2025. URL: https://claude.ai/login [accessed 2026-02-12]
18. Campbell RJ. Change management in health care. Health Care Manag (Frederick) 2008;27(1):23-39. [doi: 10.1097/01.hcm.0000285028.79762.a1] [Medline: 18510142]
19. Gorbenko K, Romanchuk K, Sagliocca F, Mazumdar M. A changing supply chain for a changing health care system: barriers and facilitators of implementing enterprise resource planning. Work 2023;74(3):977-990. [doi: 10.3233/WOR-210081] [Medline: 36404564]

Abbreviations

MSH: Mount Sinai Hospital
RN: registered nurse

Edited by A Virani; submitted 05.May.2025; peer-reviewed by B Cross, M Tadiello; revised version received 09.Oct.2025; accepted 23.Nov.2025; published 19.Mar.2026.

Please cite as:
Gorbenko K, Bailon M, Randall J, Viskochil J, Cualing C, Healy P, Mazumdar M, Freeman R
Virtual Nursing Pilot in the Inpatient Setting: Qualitative Evaluation
JMIR Nursing 2026;9:e76994
URL: https://nursing.jmir.org/2026/1/e76994
doi:10.2196/76994

© Ksenia Gorbenko, Maria Bailon, Jordan Randall, James Viskochil, Clara Cualing, Patrick Healy, Madhu Mazumdar, Robert Freeman. Originally published in JMIR Nursing (https://nursing.jmir.org), 19.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Developing a Cross-Device Platform for Robotic Systems in Nursing Care: Mixed Methods Feasibility Study

Pascal Müller[1], MSc; Ruven Veit[1], BA; Sebastian Hofstetter[1,2], PhD; Patrick Jahn[1], Prof Dr

[1]Health Service Research Working Group, Acute Care, Department of Internal Medicine, Faculty of Medicine, University Medicine Halle (Saale), Martin-Luther-University Halle-Wittenberg, Magdeburger Straße 12, Halle (Saale), Germany

[2]Dorothea-Erxleben-Learning-Centre, Faculty of Medicine, University Medicine Halle (Saale), Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

Corresponding Author:
Pascal Müller, MSc
Health Service Research Working Group, Acute Care, Department of Internal Medicine, Faculty of Medicine, University Medicine Halle (Saale), Martin-Luther-University Halle-Wittenberg, Magdeburger Straße 12, Halle (Saale), Germany

Abstract

Background: Aging populations and rising chronic illness prevalences are increasing demands for nursing care, while staff shortages threaten care quality. Robotics offer potential support, yet usability, workflow integration, and user acceptance remain major barriers.

Objective: This study aims to develop and evaluate the feasibility, usability, and acceptance of a cross-device platform for controlling robotics in nursing using a participatory, user-centered approach.

Methods: A convergent mixed methods feasibility study was conducted across 4 iterative workshops with 13 nurses from 2 German health care facilities. Quantitative measures included the System Usability Scale (SUS), Technology Usage Inventory (TUI), and Technology-based Experience of Need Satisfaction (TENS-Interface). Qualitative data were collected via think-aloud protocols and focus groups. Data integration supported iterative platform refinement and assessment of usability, acceptance, and satisfaction of psychological needs.

Results: Participants exhibited high curiosity, perceived usefulness, and strong intention to use robotics, with low skepticism. SUS scores indicated acceptable usability. TENS-Interface scores showed increased autonomy and competence following workflow simplification and stepwise interaction logic. Qualitative findings emphasized intuitive control, personalized interventions, centralized management of multiple technologies, integration with documentation systems, and structured training. Triangulation of quantitative and qualitative data confirmed that iterative, user-centered refinements enhanced usability, acceptance, and platform effectiveness.

Conclusions: Cross-device platforms integrating robotics can be successfully developed through participatory, user-centered methods. Technical usability, personalization, workflow integration, and structured training are key for adoption. The study demonstrates that technological barriers, rather than human resistance, are primary constraints to integrating robotics into nursing practice and can be mitigated through iterative co-creation aligned with real-world care contexts.

Trial Registration: Deutsches Register Klinischer Studien DRKS00034195; https://drks.de/search/de/trial/DRKS00034195QC

International Registered Report Identifier (IRRID): RR2-10.2196/63089

(JMIR Nursing 2026;9:e84118) doi:10.2196/84118

KEYWORDS

digital assistive technologies; digital transformation; human–technology interaction; customization; user-centered design; cocreation; intention to use; health care; long-term care; feasibility study

Introduction

Background

Worldwide demographic shifts, including population aging and the increasing prevalence of chronic illness, are driving a growing demand for nursing care, while persistent shortages of qualified nursing professionals threaten the quality and availability of services [1-4]. Robotic systems have therefore been proposed as a promising strategy to support nursing care; however, their implementation in routine practice remains challenging [5]. According to International Organization for Standardization (ISO) 8373:2021 [6], a robotic system is defined as a programmable, actuated mechanism capable of performing physical tasks autonomously or semi-autonomously under the control of a control system. Through the integration of sensors, actuators, and control components, robotic systems can interact

XSL•FO
RenderX

with their environment and execute predefined or adaptive actions. This broad definition includes service and assistive robots used in health care settings, such as telepresence robots and mobile manipulators.

In nursing contexts, robotic systems have been shown to support psychosocial and rehabilitative processes, with reported benefits for emotional well-being, cognition, social interaction, and quality of life among older adults [7]. Robots are also used in rehabilitation and physical therapy as well as in psychosocial interventions [8]. Beyond direct patient care, robotic technologies can support organizational processes, including telepresence, virtual consultations, and the training of nursing staff, and can assist with logistical tasks such as transporting medication and materials, disinfection, and procedural support [9]. Despite these potential benefits, most robotic systems remain at early stages of development or testing, and only a limited number have been implemented in routine care settings [10,11]. Moreover, the high financial investment required for acquisition and maintenance continues to be a major barrier to widespread adoption [11].

Previous research indicates that the successful implementation and integration of robotic systems in nursing care depends not only on technical performance but also on user acceptance, contextual implementation processes, and the systematic involvement of end users throughout development and evaluation phases. International reviews and empirical studies consistently demonstrate that sociocultural, organizational, and user-centered factors play a decisive role in facilitating practical adoption, beyond the technical design of robotic systems alone [10,12,13].

Nursing staff generally report a willingness to engage with robotic technologies, driven by perceived benefits such as workload reduction, increased efficiency, and improvements in care quality [14,15]. However, many existing robotic systems do not adequately meet users' needs with regard to usability, adaptability to individual patient requirements, and integration into established care workflows [7,8,11]. In particular, a lack of intuitive interaction concepts and standardized, cross-device solutions often results in complex and inconsistent operation [8,16,17]. Furthermore, the absence of common standards for the external control of robotic systems remains an unresolved challenge [16,18]. Participatory development approaches, such as cocreation [19] and design-based research [20], have been shown to address these challenges by actively involving end users throughout the design process. Early and continuous stakeholder involvement not only improves system usability and acceptance but may also positively influence care processes and outcomes by aligning technological solutions with user needs and care contexts [21-24]. Nevertheless, robotic systems that provide simple, cross-device, and adaptable interaction concepts and that are co-designed directly with nursing staff remain scarce [24].

Objective

Robotic systems have the potential to support nursing practice; however, their uptake is limited by insufficient usability and limited adaptability to individual care needs. The Educational Exploration Robot Application Platform (EduXBot) project addresses this gap by developing a simplified and unified interface that enables nursing staff to operate existing robotic systems. The accompanying mixed methods feasibility study pursued two objectives: (1) primary objective: to assess the usability, technology acceptance, and user experience of the EduXBot platform among nursing staff within the framework of an iterative, user-centered design process and (2) secondary objective: to develop a multilevel benefit concept for the platform through an in-depth qualitative evaluation using think-aloud and focus group methods to identify perceived value and areas for further improvement.

Together, these objectives aimed to evaluate the platform's effectiveness and its potential to support the sustainable integration of robotic systems into nursing practice.

Methods

Conceptual Framework

The EduXBot project aims to develop an interaction and control platform that simplifies the use of robotic systems in nursing care. The development process follows a participatory and user-centered approach, explicitly incorporating the expectations, experiences, and practical requirements of end users throughout all phases of the project. This process fosters the development of innovative, practical, and user-centered solutions that are more likely to be adopted in everyday practice [25].

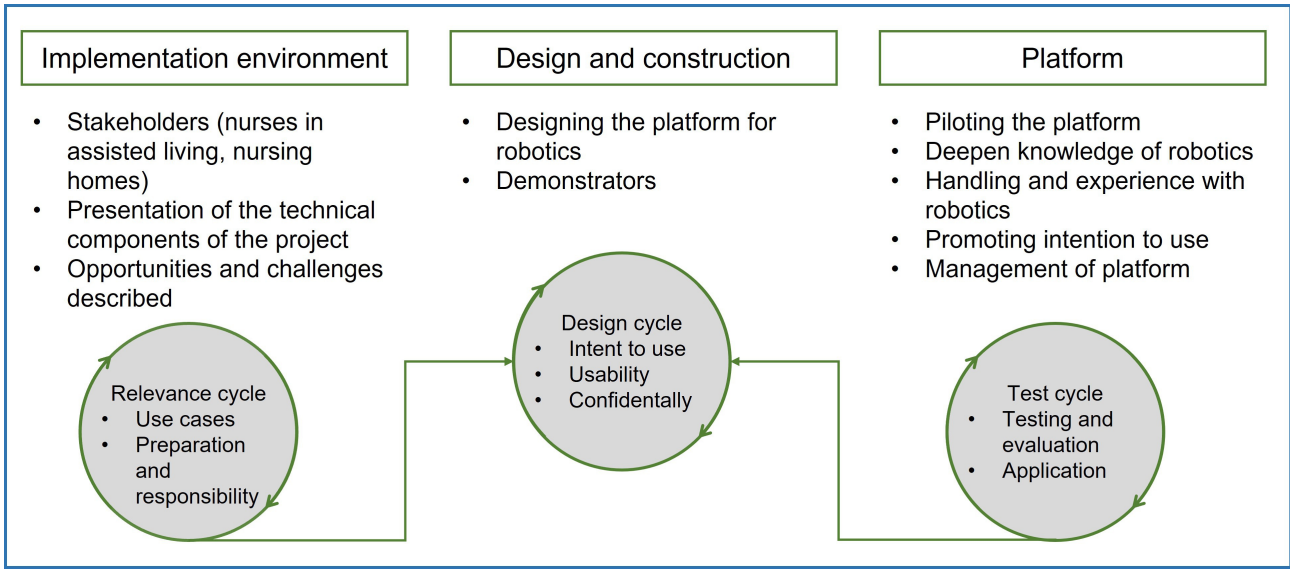
The development process was grounded in user-centered and cocreative principles that combine iterative engagement with end users and structured design cycles. Although Farao et al [26] described a user-centered framework that integrates iterative design and stakeholder engagement, our approach emphasizes active collaboration with end users throughout iterative prototyping, evaluation, and refinement phases consistent with cocreative design practices in participatory technology development.

This methodological approach supports the development of a tailored end product based on nurses' domain expertise while systematically considering potential implementation consequences [26-28]. Previous research has shown that cocreative and user-centered approaches can improve relevant outcomes and increase the usability and acceptance of technological systems in health care settings [29,30].

Figure 1 illustrates the integration of relevance, rigor, and design cycles to align nursing practice requirements, scientific evidence, and iterative platform development: the relevance cycle, which captures the realities and needs of end users; the rigor cycle, which draws on the existing scientific knowledge base; and the design cycle, which translates these inputs into the iterative development of the technical product [31-33]. Through this framework, requirements for the robotic systems addressed in the study can be systematically identified and practically operationalized. The close collaboration between researchers and end users further enables active user involvement in product evaluation, thereby supporting the functionality, usability, and overall success of the platform.

XSL•FO
RenderX

Figure 1. User-centered and cocreative design framework (adapted from Farao et al [26]).



Within this collaborative framework, EduXBot does not seek to develop new robotic hardware. Instead, it focuses on providing an extensible interaction and control platform for existing robotic systems. The platform is designed as a flexible and open system architecture that integrates diverse robotic, sensor, and interaction technologies into a unified interface. Its modular design allows nursing staff with varying levels of technological expertise to configure and use the system, thereby supporting a layperson-oriented approach to human-technology interaction. This design was developed in close collaboration with end users to ensure relevance and usability in nursing practice.

Study Design

An iterative feasibility study accompanied the development of the EduXBot platform. The study was conducted as a convergent mixed methods design in accordance with Creswell and Plano Clark [34], with quantitative and qualitative data collected in parallel and analyzed independently. Quantitative data were used to assess usability, technology acceptance, and the fulfillment of psychological needs, while qualitative data captured nursing staff's experiences, perceptions, and interaction processes. Following separate analyses, both data strands were integrated to enable a comprehensive interpretation of the findings and to inform the iterative development of the platform. This mixed methods approach aligns with a pragmatic research paradigm, in which methodological choices are driven by the research questions and support a triangulated understanding of the research problem.

Two health care facilities in Germany served as practice partners. Facility 1 was a long-term inpatient care facility. Facility 2 was a regional health center providing acute inpatient care, outpatient services, and long-term inpatient care. Both institutions had previously participated in technology-related research projects and agreed to collaborate in this study. Facility managers were informed about the project and asked to encourage nursing staff to participate.

Participants were recruited using purposive sampling with the aim of maximizing variation in age, gender, and professional experience, thereby capturing a broad range of perspectives relevant to the research topic [35]. The inclusion of nurses from diverse backgrounds enhanced the heterogeneity and informational value of the study sample. Participation was voluntary, and scheduling constraints or other commitments could limit participation or lead to withdrawal during the study.

Instruments

Quantitative Instruments

Surveys were administered at 4 measurement time points (T0-T3) during each development iteration and included the Technology Usage Inventory (TUI), System Usability Scale (SUS), and Technology-based Experience of Need Satisfaction (TENS-Interface). All instruments were administered as self-report questionnaires completed individually by each participant.

The TUI [36] assessed participants' intention to use robotic systems by measuring technology-related and psychological factors. For this study, the TUI was adapted to address robotic systems in general and was administered once per measurement point. We analyzed 7 subscales: curiosity, fear of technology, interest, skepticism, accessibility, usability, and usefulness. Higher scores indicate stronger expression of the corresponding construct, while lower scores indicate weaker expression. Scale sum scores were converted into stanine values ranging from 1 (well below average) to 9 (well above average).

The SUS [37] measured the usability of each technology integrated into the EduXBot platform. Collecting SUS scores for individual technologies allowed the identification of technology-specific usability issues, providing targeted feedback for iterative development. Medians and IQRs were aggregated to provide a generalizable estimate of the platform's overall usability, aligning the technology-level measurements with the study's primary focus on platform-level usability. Scores ≥68 were considered indicative of acceptable usability, supported by adjective-based interpretations from "outstanding" to "very poor."

XSL•FO
RenderX

The TENS-Interface [38] is a standardized questionnaire that assesses the extent to which a technology interface meets users' psychological needs (autonomy, competence, relatedness) in human-technology interaction. It was administered from T1 onward, when a prototype was available, and focused specifically on the interaction platform interface.

Qualitative Methods

Qualitative data were collected through think-aloud sessions and focus groups. During think-aloud sessions, participants performed predefined tasks for each technology, verbalizing their thoughts and experiences while interacting with the platform [39]. Standardized instructions ensured consistency, enabling systematic observation of interface interaction and usability. Sessions were audio-recorded and transcribed for analysis.

Focus groups were conducted after each workshop [40,41] to refine the platform iteratively. Participants discussed their experiences with the current prototype; evaluated its compatibility with daily care workflows; and identified unmet needs, usability barriers, and desired functionalities. Focus groups were guided by predefined questions, audio-recorded, and transcribed verbatim. Findings were systematically fed back into the design process and informed subsequent platform refinements.

Data Collection

Prior to the initial measurement (T0), a needs assessment workshop (Table 1) was conducted involving nurses, nursing scientists, and developers. Participants were introduced to existing robotic technologies and collaboratively generated potential application scenarios. Based on practical relevance and technical feasibility, 4 use cases were selected to guide the development of EduXBot prototypes:

1. Conversational partner: implemented via the Double3 telepresence robot, which recognizes residents by voice or face and can respond to personal preferences and topics
2. Telepresence: also using the Double3, enabling remote participants to communicate and engage in activities on site
3. Safe walking partner: implemented via the Go1 4-legged bionic robot, providing monitoring during walks and reporting falls
4. Record/playback of personal events: implemented using the PICO 4 virtual reality (VR) headset, allowing the recording and immersive playback of events such as family birthdays or facility celebrations, to support biographical work

Table . Overview of the 4 iterative workshops at the 4 measurement time points (T0-T3).

Measurement time point	Thematic focus	Use cases	Technologies	Instruments
T0	Define requirements	• Conversational partner • Telepresence • Record/playback of personal events	• Double3 • PICO 4	• TUI[a] original questionnaire (pre-post version) • SUS[b] • Think-aloud • Focus group
T1	Explore concept framework and interaction design	• Conversational partner • Telepresence • Record/playback of personal events • Safe walking partner	• Double3 • PICO 4 • Go1	• TUI II parallel questionnaire (complete version) • SUS • TENS-Interface[c] • Think-aloud • Focus group
T2	Testing the prototype for interaction design	• Conversational partner • Telepresence • Record/playback of personal events	• Double3 • PICO 4	• TUI II parallel questionnaire (complete version) • SUS • TENS-Interface • Think-aloud
T3	Test simplified step-by-step processes and integrated control	• Conversational partner • Record/playback of personal events • Safe walking partner	• Double3 • PICO 4 • Go1	• TUI II parallel questionnaire (complete version) • SUS • TENS-Interface • Think-aloud • Focus group

[a]Technology Usage Inventory.

[b]System Usability Scale.

[c]Technology-based Experience of Need Satisfaction.

XSL•FO
RenderX

Data were collected at 4 time points (T0-T3), each corresponding to one iterative development cycle. Workshops followed a standardized structure and integrated evaluation of both the interaction platform and selected technologies. Depending on the development stage, not all technologies nor use cases were presented at every workshop. Table 1 summarizes the technologies and use cases tested at each workshop, along with the applied quantitative and qualitative instruments and thematic focus.

A prototype, defined as a preliminary version of a product or system used to evaluate and refine its key characteristics, design, and functionality prior to final implementation, was used throughout the study [42,43]. Participants interacted in small groups (up to 3 individuals) at station-based setups, rotating between stations to test the prototype and technologies available for that workshop. Usability was assessed at the individual level using the SUS immediately after task completion at each station, ensuring that responses reflected individual perceptions rather than group-based consensus.

After station-based testing, participants completed the TUI and TENS-Interface questionnaires independently and individually, capturing overall intention to use and psychological need satisfaction related to the platform. To minimize potential social influence, participants were instructed to complete all questionnaires independently, and no collective rating nor group agreement was requested at any point during data collection.

Each workshop concluded with a focus group discussion in which participants collectively reflected on experiences, workflow compatibility, usability barriers, and suggested refinements. Insights from think-aloud sessions and focus groups were systematically fed back into the iterative development process and informed subsequent platform and technology adaptations.

Data Analysis

Quantitative data were analyzed using SPSS version 28 (IBM Corp) [44]. Continuous variables are reported as medians and IQRs, and categorical variables are reported as absolute and relative frequencies. Missing data were handled using pairwise deletion.

Group differences in cumulative SUS and TUI scores at T3 were examined using 1-way ANOVA considering factors such as gender, qualification, and facility. Assumptions of normality and homogeneity of variances were tested, with the Welch correction applied if homogeneity was violated. Post hoc Bonferroni tests were planned but not conducted due to nonsignificant effects. Pearson correlations were calculated to examine relationships between usability (SUS), intention to use (TUI subscales), and psychological need satisfaction (TENS-Interface) at T3. Hierarchical regression analyses assessed predictors of intention to use, with SUS, TENS-Interface, and TUI subscales entered in blocks. Model assumptions (linearity, independence of residuals, homoscedasticity, multicollinearity) were checked prior to interpretation.

Qualitative data were analyzed using inductive content analysis following the method by Kuckartz and Rädiker [45] with MAXQDA version 24.10.0 (VERBI Software) [46]. Transcripts from think-aloud sessions and focus groups were initially coded independently by two researchers (PM, RV). Both methods addressed the same overarching research question—platform design and adaptation of use cases. Coding of the two data sources was therefore integrated in a single schema, enabling triangulation, consistent identification of emerging themes, and comprehensive interpretation across data sources.

Codes were iteratively grouped into subcategories and overarching categories through discussion until consensus was reached. The coding process included systematic comparison across iterations to capture the evolution of requirements, usability challenges, and user perceptions. Intercoder reliability was assessed using Cohen kappa coefficient.

Qualitative data collection was structured according to predefined iterative development cycles rather than aiming for theoretical saturation. The number and timing of think-aloud sessions and focus groups were determined by the 4 development iterations, with each workshop representing 1 iteration of platform refinement. Data collection concluded upon completion of the final iteration, ensuring that all planned prototypes had been systematically evaluated. Emerging themes and usability issues were continuously reviewed after each iteration and directly informed subsequent platform adaptations. This approach was chosen because the study's objective was to iteratively refine the platform in close collaboration with end users, rather than to achieve saturation of qualitative data.

Mixed methods integration followed a convergent design [34], incorporating both formative and summative integration of quantitative and qualitative data. Formatively, findings from think-aloud sessions and focus groups were continuously fed back into the iterative development process to refine the platform. Summatively, quantitative and qualitative results were triangulated, comparing and integrating findings across data sources to provide a comprehensive understanding of platform usability, technology acceptance, and practical value in nursing workflows. This triangulation allowed complementary perspectives, convergence, and divergence to be identified and interpreted in context.

Ethical Considerations

The study received ethics approval from the ethics committee of the Medical Faculty of Martin Luther University Halle-Wittenberg (approval 2023 - 190, August 31, 2023) and was registered at the German Clinical Trials Register (DRKS00034195). Written informed consent was obtained from all participants, and no financial compensation was provided. Data were collected anonymously to ensure participant privacy and confidentiality.

The methodological rigor of the study was ensured by adhering to the Good Reporting of a Mixed Methods Study (GRAMMS) checklist [47] and the Mixed Methods Appraisal Tool (MMAT) [48]. The comprehensive study protocol has been published elsewhere [49].

XSL•FO
RenderX

Results

Participants

Between February 2024 and May 2025, 13 nurses from 2 health care facilities participated in the study. Some nurses participated in multiple iterations, yielding a total of 33 participation units across the 4 development iterations. Participants could not be clearly categorized into core or ad hoc groups due to variable attendance; however, each nurse participated in at least 2 iterations.

Participants' age, gender, professional qualifications, and experience are summarized in Table 2. Age and professional experience were calculated at the time of each participant's first iteration to account for staggered entry. Professional experience was defined as years since completion of formal nursing education or training. Median age was 46 (IQR 10.5) years, and the median length of professional experience was 15 (IQR 13.5) years. Most participants held managerial nursing roles, supplemented by support staff and occupational therapists, providing diverse perspectives for assessing the integration of the platform into daily care routines.

Table . Demographic characteristics of participating caregivers at the 4 measurement time points (T0-T3).

Characteristics	T0 (n=8)	T1 (n=9)	T2 (n=8)	T3 (n=8)
Age (years), mean (SD)	45.6 (9.1)	47.8 (9.3)	50.1 (7.8)	48.3 (9.5)
Gender, n (%)				
Male	1 (12.5)	1 (11)	1 (12.5)	1 (12.5)
Female	7 (87.5)	8 (89)	7 (87.5)	7 (87.5)
Qualification, n (%)				
3-year duration of training	1 (12.5)	1 (11)	0 (0)	1 (12.5)
3-year duration with professional training	2 (25)	1 (11)	1 (12.5)	1 (12.5)
Nursing service management	3 (37.5)	3 (33)	3 (37.5)	3 (37.5)
Studies	1 (12.5)	2 (22)	2 (25)	1 (12.5)
Other qualification	1 (12.5)	2 (22)	2 (25)	2 (25)
Professional experience (years), mean (SD)	13.5 (9)	16.3 (10.8)	16.9 (10.8)	18 (10.5)
Facility, n (%)				
Facility 1	6 (75)	7 (78)	7 (87.5)	6 (75)
Facility 2	2 (25)	2 (22)	1 (12.5)	2 (25)

Quantitative Evaluation

Participants exhibited higher openness to robotics than the general population, as indicated by above-average TUI stanine scores relative to the normed reference population [36]. Table 3 presents TUI subscale scores including curiosity, fear of technology, interest, usability, usefulness, skepticism, and accessibility. Scores are reported as medians and IQR, with stanine values indicating relative positioning versus the normed reference population. Median intention to use at T0 was 255 out of 300 and remained stable across iterations. High scores were observed for curiosity (median 23 from a possible total of 28) and perceived usefulness (median 22 from a possible total of 28), whereas skepticism was low (median 9 from a possible total of 28). These results indicate general interest in and acceptance of robotics.

XSL•FO
RenderX

Table . Technology Usage Inventory (TUI) subscale scores at the 4 measurement time points (T0-T3).

TUI subscale	Max score	T0 score		T1 score		T2 score		T3 score	
		Median (IQR)	Stanine	Median (IQR)	Stanine	Median (IQR)	Stanine	Median (IQR)	Stanine
Curiosity	28	22 (6.5)	7	22 (9.75)	7	20 (7.5)	7	23 (9.25)	8
Fear of technology	28	9 (13.5)	5	7 (11)	4	7 (3.75)	4	10 (7.5)	5
Interest	28	18 (12)	6	23 (2)	7	24 (5.5)	7	24 (6.75)	7
Usability	23	18 (4.75)	6	14 (5.5)	4	14 (2.75)	4	17 (5.5)	6
Usefulness	28	22 (4.5)	9	24 (3.5)	9	24 (9)	9	22 (5.5)	9
Skepticism	28	7 (5.75)	2	9 (7.5)	3	10 (4.5)	4	9 (5.75)	3
Accessibility	23	11 (9)	6	11 (4)	6	11 (3.75)	6	11 (3.75)	6
Intention to use	300	255 (69.5)	8	246 (58.5)	7	254 (40.25)	8	262 (46.5)	8

SUS scores were assessed per technology and aggregated to evaluate overall platform usability (Table 4). Median and IQR are reported. Overall usability decreased slightly from “excellent” at T0 to “good” at T3 but remained within the acceptable range. Double3 usability improved over time, whereas Go1 and PICO 4 showed declines, with Go1 falling below the usability threshold at T3. SUS data for Go1 were not available at all measurement points, limiting trend interpretation.

Table . System Usability Scale (SUS) scores per technology and overall at the 4 measurement time points (T0-T3).

Technology	T0 score, median (IQR)	T1 score, median (IQR)	T2 score, median (IQR)	T3 score, median (IQR)
Double3	68 (22.5)	71 (19.25)	80 (12.5)	75 (—[a])
PICO 4	85 (12)	85 (30)	69 (15)	70 (27.5)
Go1	—	70 (10)	—	64 (34.38)
Overall	74 (15.7)	75 (18.25)	74 (12.5)	72 (27.81)

[a]Not applicable.

Table 5 shows caregivers’ perceived satisfaction with autonomy, competence, and relatedness while interacting with the platform. Data were collected using the TENS-Interface from T1 onward. Autonomy and competence increased across iterations, indicating growing confidence and mastery, while relatedness showed minor fluctuations. Overall, the platform interface effectively supported caregivers’ needs, particularly autonomy and competence.

Table . Technology-based Experience of Need Satisfaction (TENS)–Interface subscale scores at 3 of the 4 measurement time points (T1-T3).

TENS-Interface	T1 score, median (IQR)	T2 score, median (IQR)	T3 score, median (IQR)
Autonomy	18 (3.5)	19 (3.75)	22 (5)
Competence	20 (3.5)	20 (3.5)	23 (3)
Relatedness	16 (5)	18 (1.75)	17 (3.75)

The 1-way ANOVA revealed no significant differences in cumulative SUS nor TUI scores at T3 based on gender, qualification, or facility. Pearson correlations indicated strong associations between SUS and TUI usability (r=0.809), usefulness (r=0.773), and intention to use (r=0.751). Intention to use was highly correlated with perceived usefulness (r=0.926). Higher skepticism correlated with lower usability (r=−0.856) and usefulness (r=−0.736). Greater technology interest correlated with higher curiosity (r=0.861). Higher qualification levels were associated with lower technology fear (r=−0.800), and more professional experience was associated with lower interest in technology (r=−0.766). Competence in handling the interface correlated strongly with autonomy (r=0.912) and positively with technology interest (r=0.716).

Hierarchical regression analyses indicated that perceived usefulness (β=0.908, P<0.001) and fear of technology (β=0.102, P=0.04) were significant predictors of intention to use robotics ($F_{1,6}$=1557.926; P<0.001; R^2=0.998; adjusted R^2=0.997). Multimedia Appendix 1 provides the complete results of the ANOVA, correlation matrices, and regression models.

Qualitative Evaluation

The qualitative evaluation comprised 7 think-aloud protocols and 3 focus group discussions conducted across 4 iterative workshops. The analysis yielded 36 codes, which were condensed into 13 subcategories and grouped into 4 overarching categories reflecting platform requirements: usability and interaction (category 1), intervention design and implementation

XSL•FO
RenderX

(category 2), training and skill development (category 3), and structural prerequisites (category 4). The full coding framework is provided in Multimedia Appendix 2.

Category 1 encompasses user-friendliness and interaction with the platform. Participants emphasized the need for simple and familiar operating concepts, such as gesture- and voice-based input, as well as a streamlined selection process with options for customization. As one caregiver put it, "What's important is that it's simple, easy to use, and works in just a few steps" (T1, Nurse 3). The platform should also enable centralized management of multiple technologies and provide interfaces to existing systems.

Category 2 focuses on intervention design and implementation. It emphasizes technological autonomy, flexible planning and booking systems, and coordination by nursing staff. Personalization emerged as a central requirement, necessitating adaptation of content to the biographies, mobility levels, and cognitive limitations of care recipients. Participants also highlighted the value of user-specific profiles to support personalization and efficiency: "You would need to create a profile so you can tap on the user who is supposed to use it, and everything that was previously stored for that person is already there" (T1, Nurse 3). Interventions were expected to be meaningful and to provide recognizable benefits, thereby supporting individualized care.

The platform was also expected to support training and skill development among care professionals (category 3). This included the provision of structured content, manuals, and step-by-step guides to facilitate independent learning, regardless of prior digital skills. As one participant noted, "If we were to work with it now, we would first need a written, step-by-step guide telling us exactly what we have to do" (T1, Nurse 2). In addition, participants emphasized the need for information on care processes and for safe practice environments in which new technologies could be tested.

Structural prerequisites (category 4) describe the framework conditions required for successful implementation. These included stable internet infrastructure, access to technical support, and compliance with privacy and data protection regulations. Financial considerations were also emphasized: "Usability and cost have to match. You need something that really makes sense" (T1, Nurse 13). Available resources and the cost-effectiveness of the technology were perceived as decisive factors influencing the platform's long-term adoption.

Initial intercoder agreement was low (16.28%, κ=0.16), likely due to inconsistent segment boundaries. After refining the coding procedures, including the use of sentence-bounded segments and consolidation of enumerations, agreement increased to 82.19% (κ=0.81), indicating reliable coding for analysis.

Iterative Development and Final Operating Concept

The platform was refined through 4 iterative workshops, during which qualitative findings were continuously triangulated with quantitative measures (SUS, TUI, and TENS-Interface). This iterative process enabled the identification and resolution of usability and workflow-related issues across successive development stages.

Iteration 1: Initial Requirements

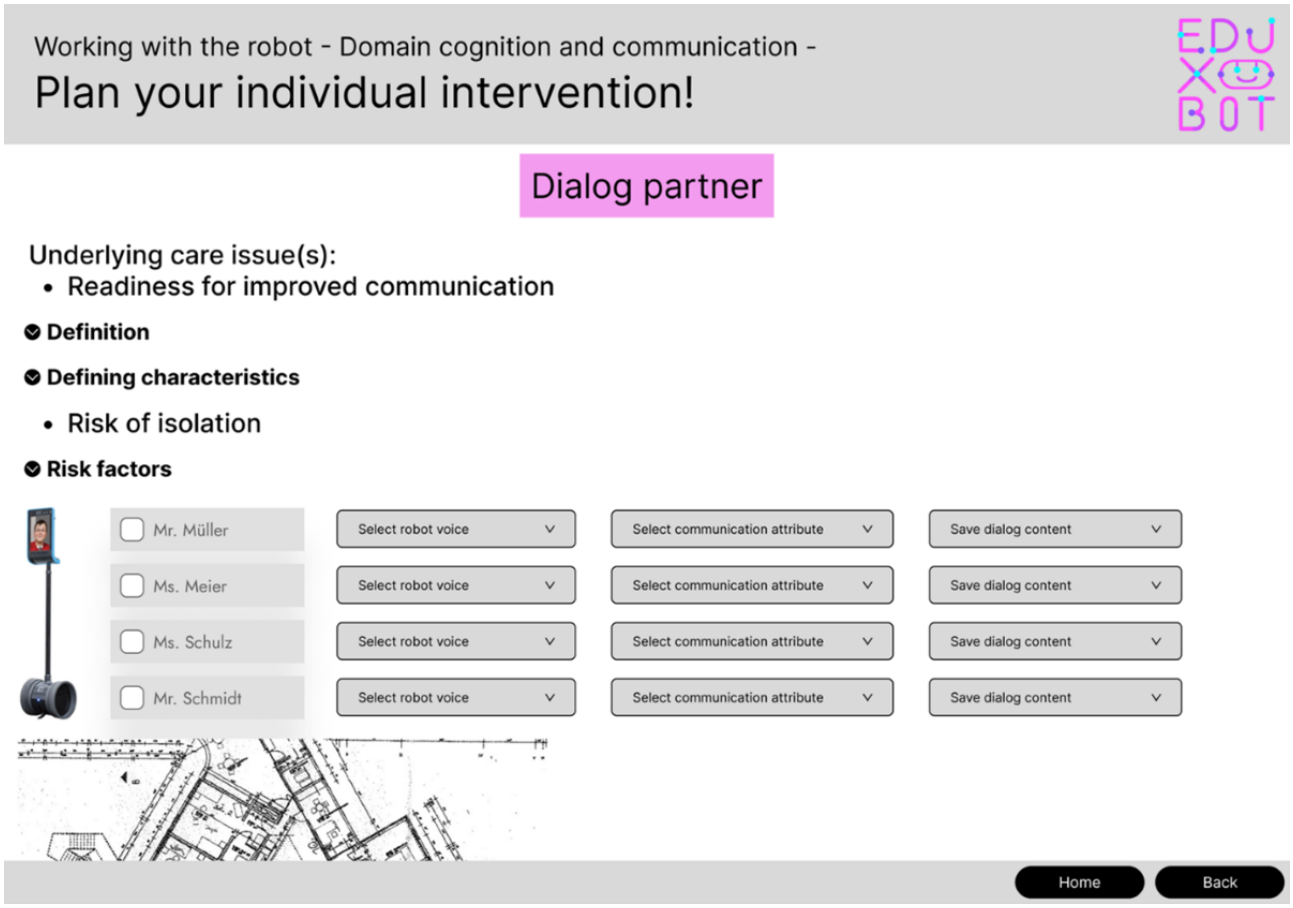
In the first iteration, participants defined core requirements for the platform. A tablet-based interface using familiar gestures (eg, tap, swipe, pinch-to-zoom) and voice input was preferred, enabling centralized control of multiple technologies. Integration with existing documentation systems was considered essential to avoid additional workload. Participants also favored concise, context-sensitive guidance: "Sure, absolutely. A brief description. So that the employee knows that I can use it for that purpose" (T0, Nurse 1). Quantitative baseline data indicated acceptable usability and a moderate intention to use robotics.

Iteration 2: Conceptual Prototype

The second iteration translated these requirements into a visual prototype. The prototype displayed all workflow steps on a single screen, including process information (care problem definitions, risk factors) presented in expandable drop-down menus, personalization options for technological interventions, and information on technology availability and location. Figure 2 shows the platform prototype.

XSL•FO
RenderX

Figure 2. Prototype representation of the Educational Exploration Robot Application Platform (EduXBot) platform (iteration 2).



Although participants valued the transparency of having all information readily accessible, extensive vertical scrolling was perceived as inefficient and contrary to the goal of minimizing interaction steps: “So it’s important that it’s simple, easy to use, and can be done in just a few steps” (T1, Nurse 3). The introduction of user profiles reflecting mobility, cognitive abilities, and biographical background to support personalization was positively received. Participants consistently emphasized that interventions should provide a discernible benefit and enable caregivers to actively support residents’ experiences: “What’s the goal? What’s the benefit?” (T1, Nurse 7).

At the same time, participants reiterated the need for written instructions (“So if we were to work with it now, we would first have to get it in writing, know step by step what we have to do” [T1, Nurse 2]) and brief training (“You have to practice a little bit, the handling, it takes some getting used to” [T1, Nurse 2]), particularly for more complex technologies.

Quantitative measures indicated stable usability and intention to use. The TENS-Interface showed moderate satisfaction with autonomy, competence, and relatedness. These metrics substantiate the qualitative findings, indicating that usability and acceptance were generally adequate but could benefit from workflow simplification and clearer guidance.

Iteration 3: Functional Prototype

In the third iteration, development shifted toward technical implementation. Full functional integration was available only for the PICO 4, while other technologies were simulated. This partial implementation revealed new usability challenges related to system feedback and transparency.

The management of VR content involved storing videos in a backend system, requiring several preparatory steps prior to use. System states were represented graphically using checkmark symbols: A single gray checkmark indicated assignment to a resident, two gray checkmarks indicated availability on the tablet, and two blue checkmarks indicated availability on the VR headset. Participants perceived this symbolic logic as lacking intuitive clarity: “Dots and hooks in blue, gray, and so on. It’s not intuitive” (T2, Nurse 7).

Additional confusion arose from synchronization issues between the tablet and the VR headset. Synchronization had to be actively initiated via an eye icon within the video control bar. Participants instead tapped the video preview on the tablet, expecting the VR headset to mirror the same content. Although participants relied on a paper-based manual containing step-by-step instructions, several emphasized that intuitive design should ultimately reduce reliance on such materials.

The absence of explicit safety alerts was identified as a critical shortcoming. During simulated testing of the walking partner use case, participants stressed the need for clear error or event notifications, particularly in the case of falls: “It runs in a specific area, detects falls, and alerts nursing staff via the nurse call system” (T2, Nurse 10).

SUS scores for the PICO 4 decreased slightly, indicating usability challenges during partial implementation. In contrast,

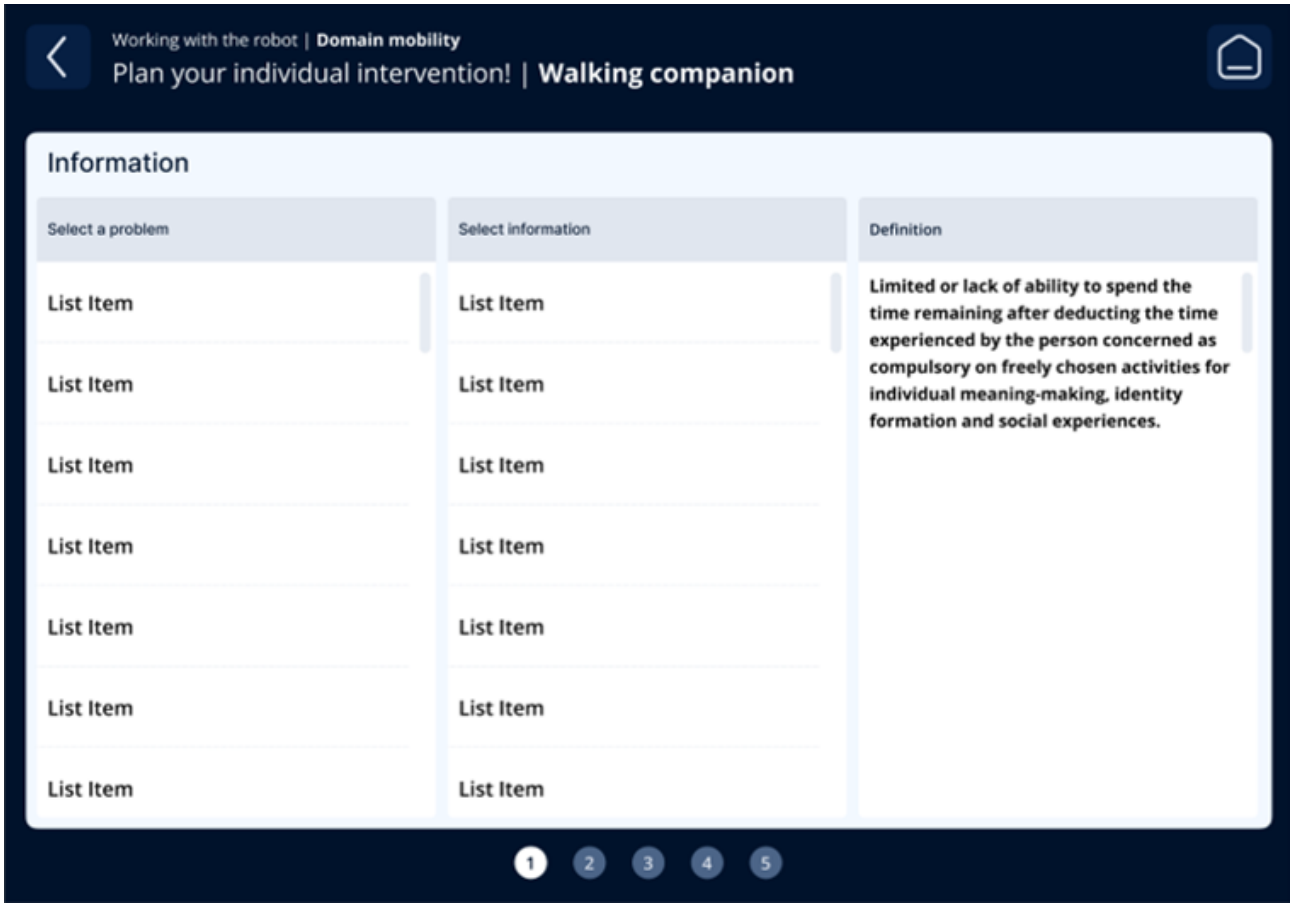
XSL•FO
RenderX

TENS-Interface scores showed modest improvements in competence and autonomy. This suggests that participants recognized the platform's potential despite ongoing functional limitations.

Iteration 4: Final Operating Concept

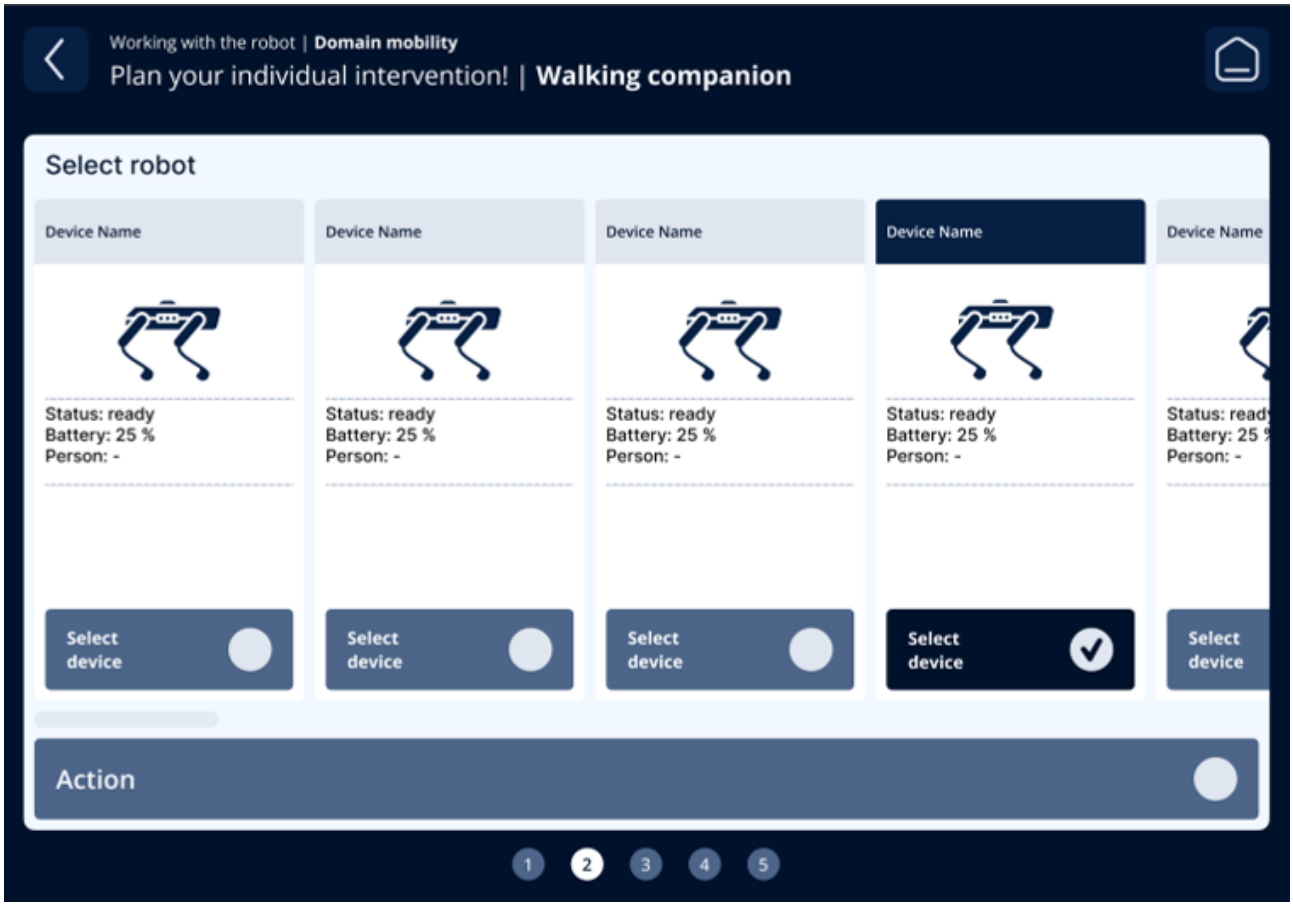
The fourth iteration resulted in a guided, step-by-step workflow aligned with nursing reasoning: (1) contextual information on the care problem clarifying the intended nursing objective of the intervention (Figure 3); (2) selection of the appropriate technology, including system feedback on availability, current usage, and battery status (Figure 4); (3) assignment of the intervention to one or more care recipients (Figure 5); (4) individualization of intervention parameters and scheduling via a booking system (Figure 6); and (5) a final overview summarizing all selected settings with an explicit confirmation step (Figure 7). These figures present the final operating concept.

Figure 3. Contextual information on the care problem.



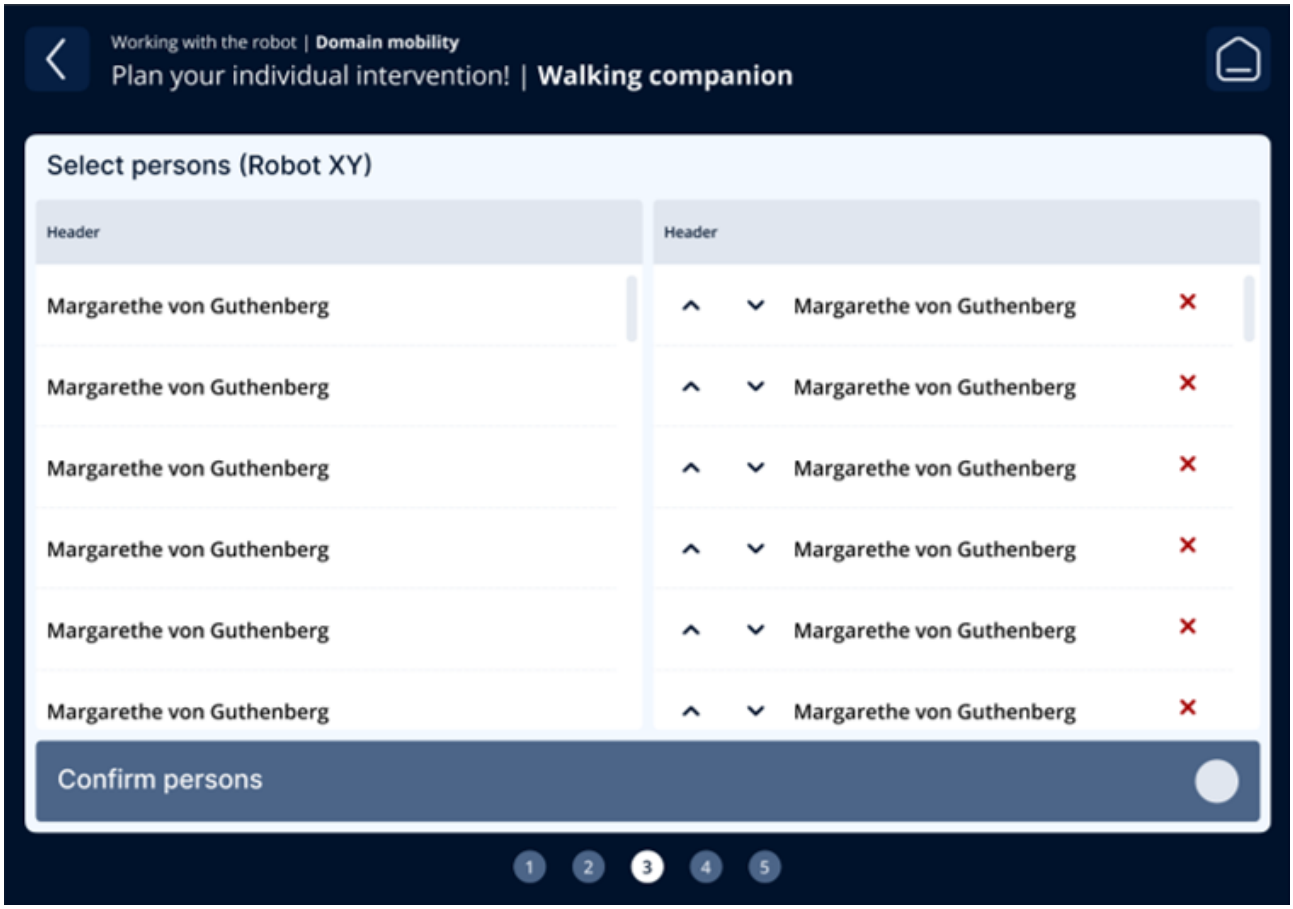
XSL•FO
RenderX

Figure 4. Selection of the appropriate technology.



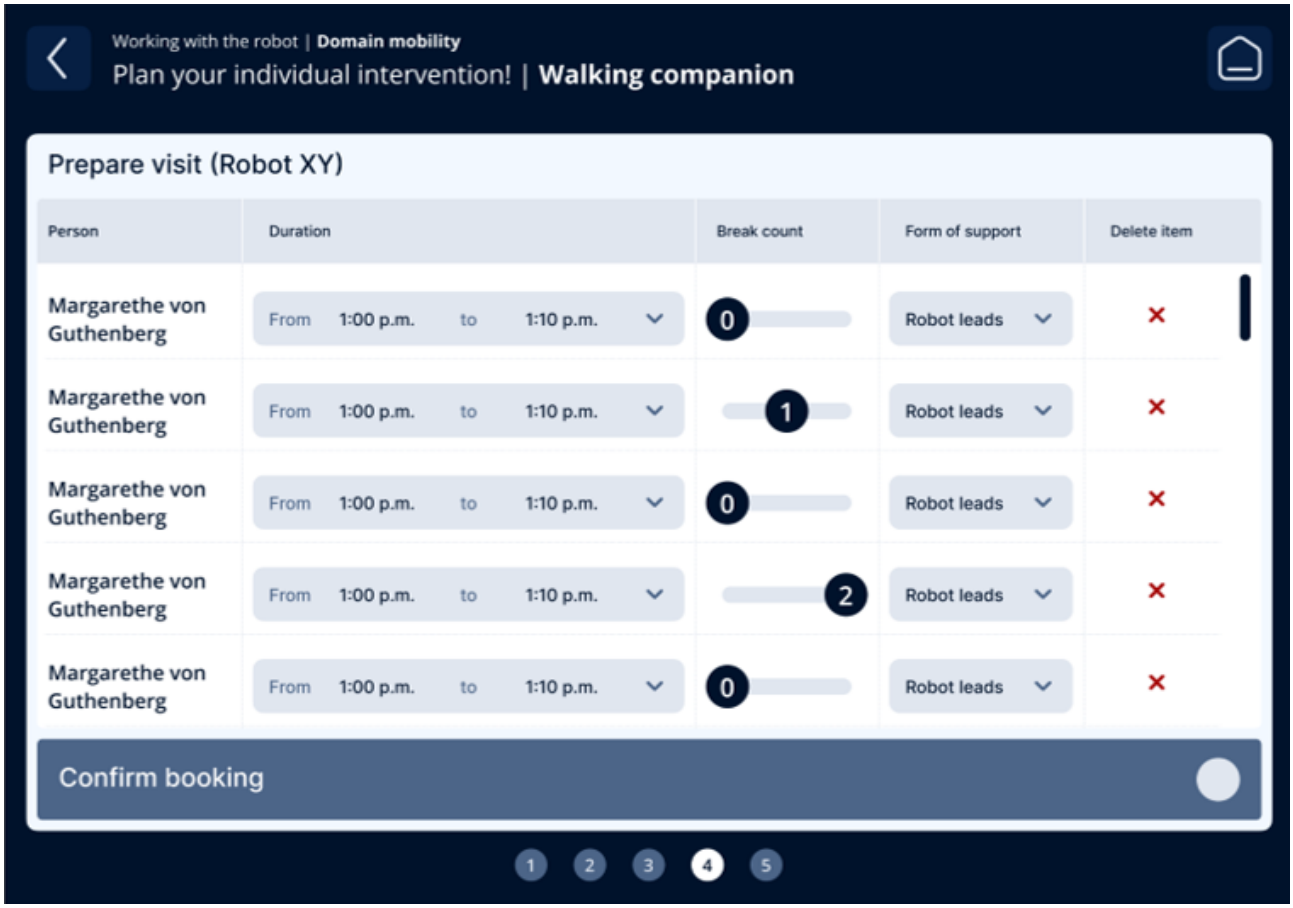
XSL•FO
RenderX

Figure 5. Assignment of the intervention.



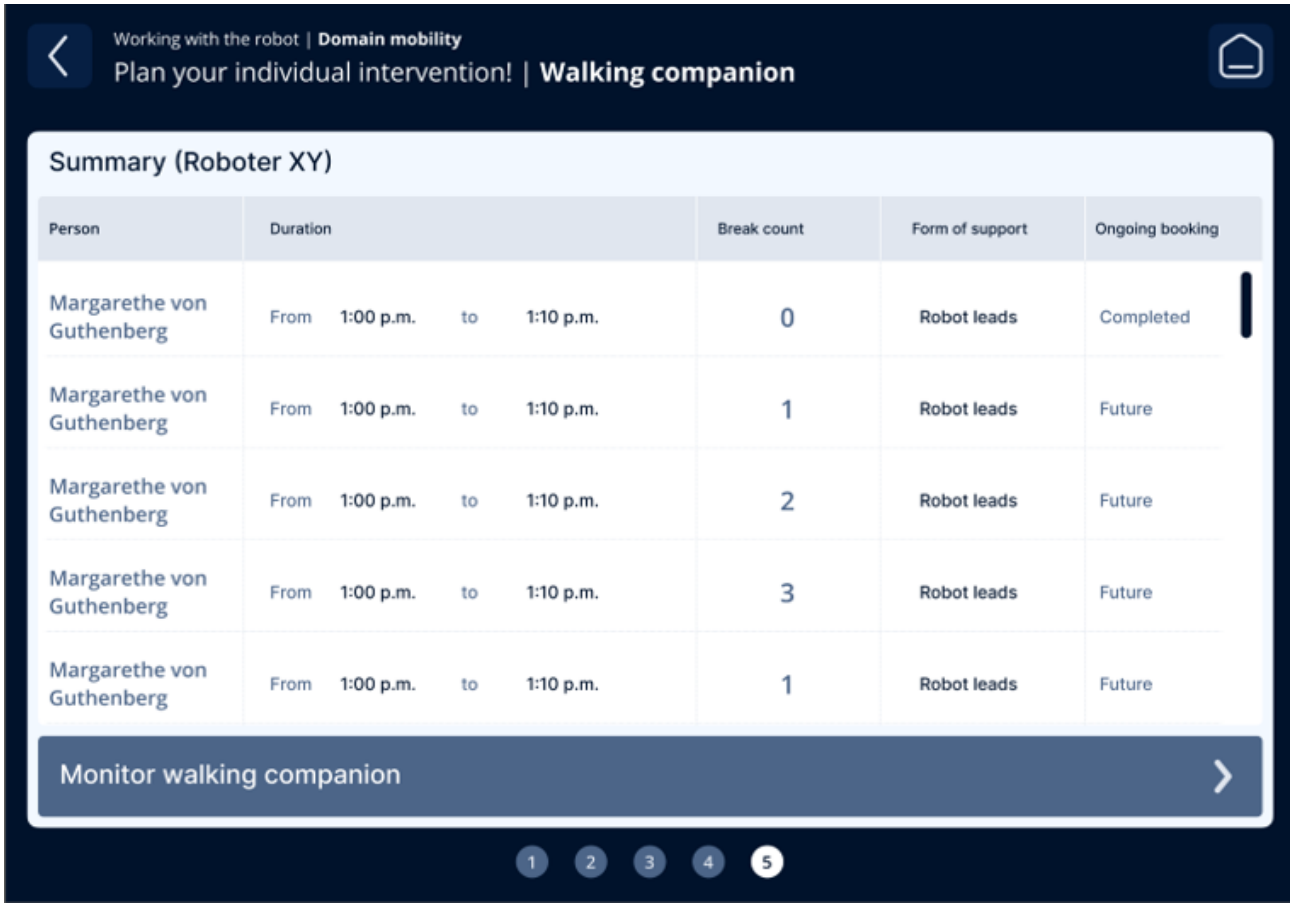
XSL•FO
RenderX

Figure 6. Individualization of intervention parameters and scheduling.



XSL•FO
RenderX

Figure 7. Final overview summarizing all selected settings.



The execution and monitoring of interventions were also revised. Care professionals retained the ability to enter an ongoing intervention at any time. Preparatory steps, such as file transfers to the tablet or synchronization with the PICO 4, were executed automatically and presented to users via a unified loading indicator. Safety-related events (eg, falls) were implemented as high-priority alerts using prominent pop-up notifications combined with audible warning signals. In addition, a conceptual interface to existing documentation systems was introduced to support routine care documentation, using standardized text modules to demonstrate the transfer of intervention-related information.

Quantitative data indicated stable or improved usability, a high intention to use, and increased satisfaction with autonomy and competence, suggesting that the revised workflow successfully addressed previously identified barriers.

Discussion

Integrative Overview

This study examined the feasibility, acceptance, and usage requirements of a cross-device interaction platform for controlling robotic systems in nursing care. Quantitative measures (SUS, TUI, TENS-Interface) and qualitative methods (think-aloud sessions and focus groups) were first analyzed separately and subsequently integrated in a mixed methods approach. The primary research question regarding overall platform feasibility was addressed through usability and interaction measures, while secondary questions concerning workflow integration and personalization were examined through combined TENS-Interface and qualitative feedback. The data were integrated both formatively, to inform iterative platform refinement, and summatively, to evaluate the platform's overall feasibility and value. This comprehensive approach allowed convergent and divergent findings to be interpreted in context.

Usability and Interaction

Caregivers demonstrated high levels of curiosity, perceived usefulness, and intention to use robotic technologies, accompanied by low levels of skepticism. Perceived usefulness emerged as the strongest predictor of intention to use robotic systems. These findings indicate that resistance among nursing staff is not the primary barrier to technology adoption. Rather, successful integration depends on whether technological solutions adequately address users' practical needs.

Although SUS scores revealed usability challenges—particularly for the Go1 robot and the PICO 4 headset—these issues largely reflect the constraints of iterative prototype development rather than fundamental rejection by users. These findings align with the qualitative findings. Early usability challenges were primarily related to complex interaction sequences, unclear system feedback, and inconsistent symbolic representations. Participants emphasized the need for intuitive, gesture- and voice-based input, streamlined workflows, and consistent symbolic representations. Iterative refinements, such as simplified step-by-step procedures and automated preparatory steps, directly addressed these usability challenges. Notably,

XSL•FO
RenderX

acceptance and intention to use remained high even when usability temporarily declined.

The triangulated evidence demonstrates that high perceived usability and meaningful interaction design are critical enablers of platform adoption, supporting the primary research question regarding overall platform feasibility. This pattern is consistent with the Technology Acceptance Model, which emphasizes perceived usefulness as a stronger determinant of adoption than ease of use alone [50].

Personalization and Workflow Integration

Participants consistently emphasized the importance of tangible benefits, including automation of routine processes, integration into existing documentation systems, and the ability to tailor interventions to individual care recipients. Technology skepticism was mitigated when systems were intuitive, workflows were transparent, and benefits were clearly recognizable. These findings align with previous research demonstrating that technologies lacking practical relevance or workflow integration are unlikely to be adopted in nursing care [7,8,11]. The strong emphasis on personalization further supports existing evidence that user-centered and cocreative design approaches enhance acceptance, satisfaction, and care-related outcomes [21-24].

Training and Implementation Support

Training and implementation support were identified as critical enablers of successful adoption. Participants' feedback on instructional materials, step-by-step guidance, and protected environments for testing technologies was reflected in increasing TENS-Interface competence scores. Nursing staff reported growing confidence in their ability to use the platform safely and independently. These findings are consistent with Basic Psychological Need Theory, which posits that satisfaction with competence and autonomy fosters motivation and engagement [51]. In line with the Motivation, Engagement, and Thriving in User Experience model, improvements in autonomy and competence were associated with greater willingness to engage with robotic systems [38].

Structural Preconditions and Organizational Perspective

Although quantitative analyses did not directly capture organizational factors, qualitative findings underscored critical preconditions for successful implementation: stable technical infrastructure, access to technical support, compliance with privacy and data protection regulations, and cost-effectiveness. Participants with managerial responsibilities provided insights into resource allocation and long-term sustainability, emphasizing that organizational and structural factors must be addressed alongside usability and personalization to ensure successful adoption.

Limitations and Strengths

A major strength of this study is its iterative, user-centered development process, which systematically involved nursing staff through cocreative workshops and repeated prototype evaluations. The integration of qualitative and quantitative data enhanced the depth, credibility, and interpretability of the findings. Adherence to established methodological standards (GRAMMS, MMAT) and formal ethical approval further strengthen the study's rigor.

Several limitations must be acknowledged. The small and relatively homogeneous sample of 13 participants—many of whom held managerial roles and reported positive attitudes toward technology—limits generalizability and may reflect selection bias. Continuous involvement across multiple iterations may have influenced participants' evaluations, as familiarity with successive prototypes could lead to more favorable perceptions of improvement. Social interaction with the research team may also have increased the likelihood of socially desirable responses.

The absence of a control group restricts the ability to obtain independent, neutral assessments of usability and perceived benefit. In addition, participation was limited to two facilities with prior experience in research projects, potentially underrepresenting settings or individuals who are more skeptical of or burdened by technology. Finally, evaluations were conducted in simulated environments rather than routine clinical practice, limiting conclusions about long-term usability, scalability, and impact on care outcomes. The study also focused on a limited set of technologies and did not include perspectives of care recipients. Accordingly, the findings should be interpreted as exploratory rather than confirmatory.

Implications for Practice and Future Research

This study demonstrates that a unified, cross-device platform for robotic systems can be successfully developed through sustained end user involvement. Although the platform was initially designed to operate robotic technologies, needs articulated during the early requirements assessment led to the integration of VR content via the PICO 4 headset. This development illustrates the platform's extensibility beyond robotics to other digital assistive technologies. Such flexibility is essential for future-oriented nursing environments, as it enables the centralized control of multiple technologies through a single interface while preserving usability and opportunities for personalization.

Nursing staff expressed willingness and motivation to use robotic technologies; however, successful implementation depends on intuitive usability, meaningful personalization, seamless workflow integration, and structured training. Quantitative improvements in autonomy and competence were reflected in design features that support familiar interaction patterns, centralized control of multiple technologies, and adaptation to individual care needs. Qualitative findings further emphasized the importance of personalization, transparent workflows, and integration with existing documentation systems for both acceptance and actual use. Collectively, these results suggest that technological rather than human factors represent the primary barriers to the integration of robotics into nursing care. Such barriers can be mitigated through collaborative design, iterative refinement, and close alignment with everyday care practices.

Future research should complement participatory development approaches with evaluations conducted by independent user

XSL•FO
RenderX

groups who were not involved in the design process. This would allow assessment from a more neutral perspective and reduce potential biases associated with repeated exposure or social desirability. Expanding participant diversity, including staff with lower digital affinity, will further strengthen generalizability. Testing the platform in real-world care settings with larger and more heterogeneous samples and incorporating the perspectives of care recipients will be essential to evaluate long-term acceptance, effectiveness, and impact. Given its open architecture, the platform provides a promising foundation for integrating additional assistive technologies and supporting broader digital transformation in nursing care.

Acknowledgments

The authors declare the use of generative artificial intelligence (GAI) in the manuscript preparation process. In accordance with the GAIDeT taxonomy [52], the following task was delegated to a GAI tool under full human supervision: translation. The GAI tool used was DeepL. Full responsibility for the content of the manuscript rests with the authors. GAI tools are not listed as authors and do not bear responsibility for the final manuscript or its conclusions.

Funding

This research was funded by the Federal Ministry of Research, Technology and Space as part of the WIR!—Wandel durch Innovation in der Region funding scheme (grant ID: 03WIR3118B). The project was funded for a 3-year period starting in January 2023, with a total funding volume of €129,805 (US $149,709). The funding body had no role in the study design, data collection, analysis, interpretation of data, or the writing of the manuscript.The authors acknowledge the financial support of the Open Access Publication Fund of Martin Luther University Halle-Wittenberg.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Statistical analysis.
[XLSX File, 22 KB - nursing_v9i1e84118_app1.xlsx]

Multimedia Appendix 2
Qualitative evaluation.
[XLSX File, 15 KB - nursing_v9i1e84118_app2.xlsx]

References

1. Boniol M, Kunjumen T, Nair TS, Siyam A, Campbell J, Diallo K. The global health workforce stock and distribution in 2020 and 2030: a threat to equity and “universal” health coverage? BMJ Glob Health 2022 Jun;7(6):e009316. [doi: 10.1136/bmjgh-2022-009316] [Medline: 35760437]
2. Chowdhury SR, Chandra Das D, Sunna TC, Beyene J, Hossain A. Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis. EClinicalMedicine 2023 Mar;57:101860. [doi: 10.1016/j.eclinm.2023.101860] [Medline: 36864977]
3. Health at a glance 2023: OECD indicators. OECD. 2023. URL: https://www.oecd.org/en/publications/health-at-a-glance-2023_7a7afb35-en.html [accessed 2026-03-13]
4. Prados-Torres A, Poblador-Plou B, Gimeno-Miguel A, et al. Cohort profile: the epidemiology of chronic diseases and multimorbidity. The EpiChron Cohort study. Int J Epidemiol 2018 Apr 1;47(2):382-384f. [doi: 10.1093/ije/dyx259] [Medline: 29346556]
5. Schlicht L, Wendsche J, Melzer M, Tschetsche L, Rösler U. Digital technologies in nursing: an umbrella review. Int J Nurs Stud 2025 Jan;161:104950. [doi: 10.1016/j.ijnurstu.2024.104950] [Medline: 39603090]
6. ISO 8373:2021 Robotics - Vocabulary, 3rd edition. International Organization for Standardization. 2021. URL: https://www.iso.org/standard/75539.html [accessed 2026-03-13]
7. Trainum K, Tunis R, Xie B, Hauser E. Robots in assisted living facilities: scoping review. JMIR Aging 2023 Mar 6;6:e42652. [doi: 10.2196/42652] [Medline: 36877560]
8. Silvera-Tawil D. Robotics in healthcare: a survey. SN COMPUT SCI 2024;5(1). [doi: 10.1007/s42979-023-02551-0]
9. Morgan AA, Abdi J, Syed MAQ, Kohen GE, Barlow P, Vizcaychipi MP. Robots in healthcare: a scoping review. Curr Robot Rep 2022;3(4):271-280. [doi: 10.1007/s43154-022-00095-4] [Medline: 36311256]

XSL•FO
RenderX

10. Ohneberg C, Stöbich N, Warmbein A, et al. Assistive robotic systems in nursing care: a scoping review. BMC Nurs 2023 Mar 18;22(1):72. [doi: 10.1186/s12912-023-01230-y] [Medline: 36934280]
11. Kyrarini M, Lygerakis F, Rajavenkatanarayanan A, et al. A survey of robots in healthcare. Technologies (Basel) 2021;9(1):8. [doi: 10.3390/technologies9010008]
12. Madi M, Nielsen S, Schweitzer M, et al. Acceptance of a robotic system for nursing care: a cross-sectional survey with professional nurses, care recipients and relatives. BMC Nurs 2024 Mar 14;23(1):179. [doi: 10.1186/s12912-024-01849-5] [Medline: 38486244]
13. Servaty R, Kersten A, Brukamp K, Möhler R, Mueller M. Implementation of robotic devices in nursing care. Barriers and facilitators: an integrative review. BMJ Open 2020 Sep 21;10(9):e038650. [doi: 10.1136/bmjopen-2020-038650] [Medline: 32958491]
14. Persson M, Redmalm D, Iversen C. Caregivers' use of robots and their effect on work environment – a scoping review. J Technol Hum Serv 2022 Jul 3;40(3):251-277. [doi: 10.1080/15228835.2021.2000554]
15. Toghian Chaharsoughi N, Ahmadifaraz M, Kahangi LS. The impact of digital technologies in nursing care and their application: a narrative review. J Multidiscip Care 2022;11(3):149-156. [doi: 10.34172/jmdc.2022.1127]
16. Walzer S, Armbruster C, Mahler S, Farin-Glattacker E, Kunze C. Factors influencing the implementation and adoption of digital nursing technologies: systematic umbrella review. J Med Internet Res 2025 Jul 31;27:e64616. [doi: 10.2196/64616] [Medline: 40743516]
17. Krick T, Huter K, Seibert K, Domhoff D, Wolf-Ostermann K. Measuring the effectiveness of digital nursing technologies: development of a comprehensive digital nursing technology outcome framework based on a scoping review. BMC Health Serv Res 2020 Mar 24;20(1):243. [doi: 10.1186/s12913-020-05106-8] [Medline: 32209099]
18. Ragno L, Borboni A, Vannetti F, Amici C, Cusano N. Application of social robots in healthcare: review on characteristics, requirements, technical solutions. Sensors (Basel) 2023 Jul 31;23(15):6820. [doi: 10.3390/s23156820] [Medline: 37571603]
19. Vargas C, Whelan J, Brimblecombe J, Allender S. Co-creation, co-design, co-production for public health - a perspective on definition and distinctions. Public Health Res Pract 2022 Jun 15;32(2):3222211. [doi: 10.17061/phrp3222211] [Medline: 35702744]
20. Armstrong M, Dopp C, Welsh J. Design-based research. In: Kimmons R, Caskurlu S, editors. The Students' Guide to Learning Design and Research: EdTech Books; 2020:70-78 URL: https://edtechbooks.org/studentguide [doi: 10.59668/10]
21. Dopp AR, Parisi KE, Munson SA, Lyon AR. Aligning implementation and user-centered design strategies to enhance the impact of health services: results from a concept mapping study. Implement Sci Commun 2020;1:17. [doi: 10.1186/s43058-020-00020-w] [Medline: 32885179]
22. Göttgens I, Oertelt-Prigione S. The application of human-centered design approaches in health research and innovation: a narrative review of current practices. JMIR Mhealth Uhealth 2021 Dec 6;9(12):e28102. [doi: 10.2196/28102] [Medline: 34874893]
23. Lyon AR, Munson SA, Renn BN, et al. Use of human-centered design to improve implementation of evidence-based psychotherapies in low-resource communities: protocol for studies applying a framework to assess usability. JMIR Res Protoc 2019 Oct 9;8(10):e14990. [doi: 10.2196/14990] [Medline: 31599736]
24. Müller P, Jahn P. Cocreative development of robotic interaction systems for health care: scoping review. JMIR Hum Factors 2024 Sep 12;11:e58046. [doi: 10.2196/58046] [Medline: 39264334]
25. Hensley J. What is co-design? A primer on participatory design. Mural. 2023. URL: https://www.mural.co/blog/co-design-method [accessed 2026-03-13]
26. Farao J, Malila B, Conrad N, Mutsvangwa T, Rangaka MX, Douglas TS. A user-centred design framework for mHealth. PLoS ONE 2020;15(8):e0237910. [doi: 10.1371/journal.pone.0237910] [Medline: 32813711]
27. McCurdie T, Taneva S, Casselman M, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
28. Risling TL, Risling DE. Advancing nursing participation in user-centred design. J Res Nurs 2020 May;25(3):226-238. [doi: 10.1177/1744987120913590] [Medline: 34394630]
29. Altman M, Huang TTK, Breland JY. Design thinking in health care. Prev Chronic Dis 2018 Sep 27;15:E117. [doi: 10.5888/pcd15.180128] [Medline: 30264690]
30. Cronholm S, Göbel H. Evaluation of the Information Systems Research Framework: empirical evidence from a design science research project. Electron J Inf Syst Eval 2016;19(3):158-166 [FREE Full text]
31. Hevner AR. A three cycle view of design science research. Scand J Inf Syst 2007;19(2) [FREE Full text]
32. Gallupe RB. The tyranny of methodologies in information systems research. SIGMIS Database 2007 Jul 28;38(3):20-28. [doi: 10.1145/1278253.1278258]
33. Schnall R, Rojas M, Bakken S, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251. [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
34. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research, 3rd edition: SAGE Publications, Inc; 2017.
35. Renjith V, Yesodharan R, Noronha JA, Ladd E, George A. Qualitative methods in health care research. Int J Prev Med 2021;12(20):20. [doi: 10.4103/ijpvm.IJPVM_321_19] [Medline: 34084317]

XSL•FO
RenderX

36. Kothgassner OD, Felnhofer A, Hauk N, Kastenhofer E, Gomm J, Kryspin-Exner I. TUI: Technology Usage Inventory manual. : FFG Forschungsförderungsgesellschaft; 2012 URL: https://ucrisportal.univie.ac.at/de/publications/tui-technology-usage-inventory-manual/
37. Brooke J. SUS: a retrospective. J Usability Stud 2013;8(2):29-40 [FREE Full text]
38. Peters D, Calvo RA, Ryan RM. Designing for motivation, engagement and wellbeing in digital experience. Front Psychol 2018;9(797):797. [doi: 10.3389/fpsyg.2018.00797] [Medline: 29892246]
39. Ericsson KA, Simon HA. Verbal reports as data. Psychol Rev 1980;87(3):215-251. [doi: 10.1037/0033-295X.87.3.215]
40. Holloway I, Galvin K. Qualitative Research in Nursing and Healthcare, 4th edition: John Wiley & Sons Inc; 2017.
41. Krueger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research, 5th edition: SAGE; 2015.
42. Preece J, Rogers Y, Sharp H. Interaction Design: Beyond Human-Computer Interaction, 4th edition: Wiley; 2015.
43. Snyder C. Paper Prototyping: The Fast and Easy Way to Design and Refine User Interfaces: Elsevier Science; 2012.
44. IBM SPSS software. IBM Corp. 2021. URL: https://www.ibm.com/products/spss [accessed 2026-03-13]
45. Kuckartz U, Rädiker S. Qualitative Inhaltsanalyse Methoden, Praxis, Umsetzung Mit Software Und Künstlicher Intelligenz, 6th edition: Beltz Juventa; 2024.
46. MAXQDA. VERBI GmbH. 2025. URL: https://www.maxqda.com [accessed 2026-03-13]
47. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
48. Hong QN, Fàbregues S, Bartlett G, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. EFI 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
49. Müller P, Hofstetter S, Jahn P. Sustainable implementation of digital assistive technologies in health care through a simplified interaction and control platform: protocol for a cocreative feasibility study. JMIR Res Protoc 2025 Mar 18;14:e63089. [doi: 10.2196/63089] [Medline: 40101746]
50. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manage Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
51. Ryan RM, Deci EL. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness: The Guilford Press; 2017.
52. Suchikova Y, Tsybuliak N, Teixeira da Silva JA, Nazarovets S. GAIDeT (Generative AI Delegation Taxonomy): a taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing. Account Res 2026 Apr;33(3):2544331. [doi: 10.1080/08989621.2025.2544331] [Medline: 40781729]

Abbreviations

EduXBot: Educational Exploration Robot Application Platform
GRAMMS: Good Reporting of a Mixed Methods Study checklist
ISO: International Organization for Standardization
MMAT: Mixed Methods Appraisal Tool
SUS: System Usability Scale
TENS: Technology-based Experience of Need Satisfaction
TUI: Technology Usage Inventory
VR: virtual reality

Edited by E Borycki; submitted 15.Sep.2025; peer-reviewed by C Buhtz, TN Wen, WK Bong; revised version received 14.Jan.2026; accepted 03.Feb.2026; published 19.Mar.2026.

Please cite as:
Müller P, Veit R, Hofstetter S, Jahn P
Developing a Cross-Device Platform for Robotic Systems in Nursing Care: Mixed Methods Feasibility Study
JMIR Nursing 2026;9:e84118
URL: https://nursing.jmir.org/2026/1/e84118
doi:10.2196/84118

© Pascal Müller, Ruven Veit, Sebastian Hofstetter, Patrick Jahn. Originally published in JMIR Nursing (https://nursing.jmir.org), 19.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perception of AI Symptom Models in Oncology Nursing: Mixed Methods Evaluation Study

Bridget Nicholson[1*], RN, PhD; Elizabeth A Sloss[1*], RN, MBA, PhD; Aref Smiley[2], PhD; Joseph Finkelstein[2], MD, PhD; Kathi Mooney[3*], RN, PhD

[1]College of Nursing, University of Utah, 10 2000 E, Salt Lake City, UT, United States

[2]College of Medicine, University of Arizona, 1401 University Boulevard, Tucson, AZ, United States

[3]College of Nursing, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, United States

[*]these authors contributed equally

Corresponding Author:
Bridget Nicholson, RN, PhD
College of Nursing, University of Utah, 10 2000 E, Salt Lake City, UT, United States

Abstract

Background: Patients undergoing cancer treatment experience a significant symptom burden. The standard process of symptom management includes patient reporting and clinical response following symptom escalation. Emerging predictive symptom models use artificial intelligence (AI) components of machine learning and deep learning to identify the risk of symptom deterioration, facilitating earlier intervention to prevent downstream effects. However, integrating predictive symptom models into clinical practice will require oncology nurses to adopt innovative approaches.

Objective: This study aims to explore oncology nurses' perceptions of the use of predictive symptom models in cancer care and the factors influencing the adoption of this symptom care innovation.

Methods: The evaluation was guided by the Rogers Diffusion of Innovation Theory, which describes the process of how individuals adopt new technologies. The investigators developed an interview guide that asked oncology nurses to rate their perceptions of AI symptom models on a Likert scale. Participants were also asked to provide qualitative comments to support their ratings for each question, in order to better understand the key factors that would influence AI predictive model adoption. Investigators analyzed demographic data and Likert ratings with descriptive statistics. Qualitative analysis of participant comments included content analysis and inductive coding to identify themes. Nurses' perception of factors that would influence the adoption of AI symptom models, based on the Rogers theory, included relative advantage, compatibility, complexity, trialability, and observability.

Results: Responses of 15 oncology nurses with more than 1 year of experience in oncology were analyzed. There was high agreement among nurse participants that an AI model could improve symptom management for patients with cancer (n=10, 67%) and increase early intervention to prevent the escalation of symptoms (n=12, 86%). All participants (N=15) agreed that receiving symptom information would be helpful. Nearly three-quarters of participants (n=11, 73%) endorsed that the information would save time. Most (n=12, 80%) recommended that clinicians receive information about the predicted symptom deterioration of their patients. Among open-ended responses, key themes were consistent with factors identified in the Diffusion of Innovation theory including: (1) perceptions related to the AI model (compatibility or complexity), (2) nurses' perception of patients' benefit (observability), (3) improved clinical processes (relative advantage or observability), (4) apprehension over model accuracy and impact (compatibility or trialability or observability), and (5) implementation or adoption (trialability or complexity or observability).

Conclusions: Oncology nurses agree that predictive symptom models could help improve symptom management for patients undergoing cancer treatment. However, nurses noted that transparency in the factors included in the AI model was essential, that nurses should be involved in the development and testing of models, and that the observability of the benefit in symptom care would need to be evident for ultimate adoption.

(JMIR Nursing 2026;9:e82283) doi:10.2196/82283

KEYWORDS

artificial intelligence; AI; oncology; nursing; symptom management; implementation

XSL•FO
RenderX

Introduction

Patients undergoing cancer treatment experience a wide range of symptoms that impact functional status, quality of life, and health care use [1-3]. Currently, symptom reporting is a reactive process based on patient reporting, followed by a response for poorly controlled symptoms. Oncology nurses have historically screened patients both at clinical visits and via phone triage when patients report increasing symptom burden [3]. Increasingly, but slowly, clinical workflows are implementing symptom monitoring with electronic patient-reported outcome (ePRO) systems, which enable patients to report symptoms electronically and allow oncology clinicians to respond accordingly [4]. However, the ePRO-based symptom management decreases care escalations, which is notable given that worsening of symptoms is a primary driver of health care use among patients with cancer [3]. However, oncology symptoms can change rapidly, and some, such as fever, require prompt evaluation and clinical action [5,6]. While responsive ePRO reporting systems have improved patient symptom burden, high levels of symptoms and health care use persist. ePRO symptom models remain reactive, with detection following a patient reporting a change and lacking the ability to anticipate symptom escalations.

One application of artificial intelligence (AI) is the use of computer-based models to analyze large quantities of data, in this case, symptom data. Predictive symptom models attempt to evaluate data and detect a change prior to patient symptom escalation. AI models are being tested using retrospective and prospective data [7]. AI models, paired with ePRO collection, are being developed to enable predictive and anticipatory warnings that may help categorize patients at an increased risk of symptom escalation [8]. AI models use patient-generated data to predict the likelihood of a specific outcome or a set of outcomes [8]. Models to diagnose both diseases and symptoms, as well as health care use, are being integrated into oncology use cases [7,9]. Predictive symptom models, which inform the identification of symptom patterns, show promise as a mechanism to enhance the accuracy of symptom detection before escalation [7,10]. AI-derived alerting models, using machine learning (ML) or deep learning methods, have the potential to predict emerging symptom escalations. These models seek to prioritize patients at increased risk for changes before symptom escalation. Detecting symptoms, such as an impending fever, before the patient experiences it can facilitate earlier intervention and better outcomes [11]. Alternative predictive approaches are necessary to detect dynamic symptom changes while reducing the burden of symptom reporting.

Transitioning to proactive care models requires a complete shift, both cognitively and operationally, for both patients and clinicians. Moving from a reactive reporting structure to a predictive symptom management model requires adoption by the oncology team. Notably, the implementation of this shift will require the engagement of oncology nurses, who will be the clinicians responsible for responding to AI-based alerts.

Few studies have examined nurses' perceptions of implementing AI-based symptom models [12]. A recent study that assessed nurse perspectives on ML-based clinical decision support systems broadly found that previous experience with technology and nurse perceived engagement in the development process, among other factors, influenced perceived use of ML clinical decision support systems [13]. The use of AI in the clinical setting is expanding, and a key theme consistently identified by nurses, nurse informaticists, and nurse leaders regarding the development, implementation, and adoption of AI-based tools is the importance of engaging nurse end-users at the beginning of the development process [14-16]. Thus, the purpose of this evaluation was to examine nurses' willingness to adopt AI-derived alert notifications about impending symptom escalations. In anticipation of implementing these AI-based symptom management systems, this exploration addresses an existing gap in the literature regarding oncology nurses' perceptions, including usefulness and anticipated efficiency, of AI-derived symptom prediction models for cancer symptom management.

The Rogers Diffusion of Innovation Theory describes the process of how users decide to participate in the adoption of new technologies [17] and framed our work to nurses' consideration to adopt AI. Using AI-based models in clinical practice will require a significant transition from current symptom evaluation processes, and oncology nurses, who are largely responsible for symptom triage, will need to adopt and use this innovation in care management workflows. Perception of the innovation, rather than the innovation itself, is key to adoption. The Diffusion of Innovation theory identifies 5 perceived attributes that influence adoption, including relative advantage, for this study whether the AI predictive models are perceived as improving current symptom monitoring and would benefit patients; compatibility—whether the AI predictive models are consistent with symptom treatment values, past experiences, and the needs of nurses providing symptom care; complexity—whether AI predictive models are seen as easy to understand and use; trialability—whether the AI predictive models can be piloted and tried out; and observability—whether the symptom management benefits of the AI predictive models can be seen by the nurses. According to the theory, adoption occurs at varying speeds based on individual characteristics and perceptions, such that a small percentage of the population will be innovators and early adopters, and others are more likely to adopt later after others have accepted the innovation. The focus of this study is on these factors and how they may influence nurses' perceptions and decision-making about the adoption of AI predictive symptom models.

Methods

Design, Setting, and Participants

We conducted a mixed methods exploration of oncology nurses' perspectives regarding the use of AI-based symptom predictive models to detect symptom changes in patients with cancer. The use of both structured questionnaire (eg, Likert-scale questions) and interview questions allowed for a more in-depth analysis of perspectives regarding the adoption of AI predictive symptom models and is well-suited for implementation research [18] Specifically, we conducted interviews with participants using

XSL•FO
RenderX

both structured, Likert-scale–based questions and open-ended questions.

A convenience sample of registered nurses with at least 1 year of experience in oncology from across the United States was recruited to participate in this project. Participants were excluded if they lacked fluency in spoken or written English, lacked access to Zoom (Zoom Communications, Inc) web-conferencing technology or were unable to meet in person, or if they had less than 1 year of experience as a nurse in oncology. Recruitment methods included direct professional referrals, social media (such as LinkedIn and Facebook), and snowball sampling. Investigators contacted participants via email to schedule interviews. Interviews, the duration of which ranged from 20 to 30 minutes, were conducted in December 2024 and January 2025 via web teleconferencing platform (Zoom) and in-person by 2 investigators (BN and EAS). Interviews were not recorded or transcribed, though detailed notes were kept by the investigators who conducted the interviews and included capturing verbatim quotes from participants.

Individual Interviews

The team developed the interview guide (Multimedia Appendix 1) to gather information on the acceptability of implementing AI predictive symptom monitoring and management. The interview guide was initially drafted by 2 investigators (BN and EAS) and feedback was obtained from other members of the team before being finalized (Multimedia Appendix 1). Before starting the interview, as outlined in the interview guide, the concept of an AI-based symptom model was presented to the participants. The description was broad in that it included general model types but emphasized the predictive capability of AI algorithms in the identification of symptom deterioration. In addition to demographic questions, the final interview guide consisted of 6 total Likert-scale questions, in which participants responded to statements about the hypothetical clinical usefulness and efficiency of a symptom prediction model, indicating their agreement or disagreement using the Likert scale (1="Strongly Disagree" to 5="Strongly Agree"). Following each Likert-scale question, participants were asked to provide open-ended comments in response to the Likert-scale question that they had previously answered. Three additional open-ended questions were meant to elicit additional information, for example, "If you received a notification that your patient is at high risk for experiencing worsening symptoms in the next 24 hours, what would you do?"

Saturation was assessed on an ongoing basis. No new information was elicited, and subsequently, no new codes were identified over the final 5 interviews, indicating that we achieved content-level saturation.

Data Analysis

Quantitative Analysis

Descriptive analyses, including means and SDs, were calculated using demographic data to describe the sample. Due to the small sample size, we rounded frequencies (percentages) to the whole number. Additionally, investigators evaluated the frequency of Likert ratings by participants through descriptive statistics. The Likert-scale ratings were on a 1 to 5 rating with responses initially coded based on a 1 to 5 rating (1="Strongly disagree," 2="Disagree," 3="Neutral," 4="Agree," and 5="Strongly agree"). However, in further analyses, we combined ratings of 1 to 2 and 3 to 5 to create categorical ratings of "Disagree," "Neutral," and "Agree." Methodologically, this approach is used to improve interpretability in smaller sample sizes, which have limited responses in multiple categories [19]. Our quantitative analysis of Likert-scale responses ultimately provided a clearer picture of the reportable trends within the sample.

Qualitative Analysis

For qualitative analysis, the team members (BN and EAS) used open coding and initially coded qualitative responses independently. After resolving disagreements and reaching consensus on codes, the investigators recoded each qualitative interview. Data were then analyzed using thematic analysis which involved several steps: data familiarization, keyword selection, identification of initial themes, and comparison of the investigators' initial themes.

Triangulation

In keeping with a mixed methods approach, the investigators synthesized quantitative and qualitative data and identified findings that converged, complemented, or diverged across data modalities [20,21]. Quantitative data from Likert-scale responses were triangulated concurrently with qualitative, open-ended responses to the questions and/or follow-up prompts. Finally, the investigators compared the codes for the factors within the Diffusion of Innovation Model. Participant quotes were used to represent themes. We used the GRAMMS (Good Reporting of a Mixed Methods Study) guidelines to aid clarity of reporting (Checklist 1) [22].

Ethical Considerations

The University of Utah Institutional Review Board reviewed the project protocol and deemed it a quality improvement project and not human participants research (00166873). Each participant was informed of the purpose of the project, including that participation was and could be discontinued at any time for any reason. Verbal consent was obtained prior to proceeding with the interview. No compensation was provided to the participants. In accordance with the rigor of human participants research, the study team followed procedures to protect the participants' privacy and confidentiality, including deidentifying participant data, not sharing data outside of the study team, and storing data securely on password-encrypted computers.

Results

User Statistics

Sample characteristics are summarized in Table 1. The sample included 15 nurses who self-identified as working in oncology for more than 1 year. Participants were all female (15/15, 100%) with a mean age of 44.6 (SD 11.44) years. The cohort consisted of an experienced group of nurses, with an average of 18.33 (SD 9.82) years of nursing experience. Most of this experience (mean 14.10, SD 9.92 y) was in oncology. Participants reported working in diverse oncology settings, including inpatient

XSL•FO
RenderX

oncology and outpatient infusion, as well as in roles related to quality improvement and patient navigation (Table 1). Furthermore, the sample was highly educated, with 8 out of 15 (53%) having completed a master's degree (Table 1).

Table . Participant demographics (N=15).

Characteristics	Participants
Age (y), mean (SD)	44.6 (11.44)
Years of experience, mean (SD)	18.33 (9.82)
Years of experience in oncology	14.10 (9.92)
Highest level of education, n (%)	
Diploma	1 (7)
Bachelor's degree	4 (27)
Master's degree	8 (53)
Doctoral	2 (13)
Practice environment, n (%)	
Inpatient oncology	4 (26)
Outpatient oncology	8 (53)
Quality	1 (7)
Navigation	2 (13)

Quantitative Evaluation

Results are presented in categorical (agree or disagree) percentages for the 6 Likert-scale questions (Table 2). All nurse participants overwhelmingly agreed that receiving the symptom information would be helpful, signaling compatibility with existing values. Furthermore, 12 out of 15 (86%) nurses believed that an AI model would enable early intervention to prevent the escalation of symptoms, aligning with this view. Most nurses (12/15, 86%) also thought that an AI model would allow the relative advantage of early intervention to prevent the escalation of symptoms. A smaller majority, or 10 out of 15 (67%) nurses, agreed that an AI model could improve symptom management for patients with cancer. The remaining one-third or 5 out of 15 (33%) participants were neutral about whether the symptom prediction model could help improve symptom management related to the disease. There was similar agreement on the expectation that a symptom prediction model would enhance a patient's quality of life, with 10 out of 15 (67%) nurses agreeing. From an efficiency perspective, 11 out of 15 (73%) nurses felt that the information may save time. Despite nurses obtaining significant volumes of clinical information during a clinical day, 12 out of 15 (80%) nurses recommended that clinicians receive information about the predicted deterioration of their patients.

XSL•FO
RenderX

Table . Perceptions of artificial intelligence (AI) predictive ratings' value in oncology symptom management (N=15).

Topic	Disagree, n (%)	Neutral, n (%)	Agree, n (%)
Knowing that a patient is at risk of symptom deterioration earlier is helpful information for me to have as an oncology clinician	0 (0)	0 (0)	15 (100)
I expect that information from an AI model would allow me to intervene earlier, preventing an escalation of patient symptoms.	0 (0)	2 (13)	12 (86)
I would recommend oncology clinicians receive information about predicted deterioration from an AI algorithm when caring for patients with cancer.	1 (7)	2 (13)	12 (80)
Having this information might save me time and/or help improve my efficiency in helping my patients reduce their symptom burden	1 (7)	3 (20)	11 (73)
I expect information from an AI model would add to reducing symptom burden and improve my patients' quality of life.	2 (13)	3 (20)	10 (67)
I expect information from an AI model would help me better manage symptoms related to cancer treatment or disease.	0 (0)	5 (33)	10 (67)

Only 1 out of 15 (7%) respondents indicated that they believed AI-based symptom models would not improve efficiency or would not recommend that oncology clinicians receive information regarding patient deterioration from an AI symptom model. A small number, 2 out of 15 (13%) nurses indicated that AI-based symptom models would not reduce symptom burden or improve quality of life. These concerns reflect fears of complexity, given the complete shift in operational paradigm. A higher percentage of respondents, ranging from 13% to 33%, were neutral in their responses, indicating that they were still considering the information on the innovation.

Qualitative Evaluation

Themes

Participants' comments further explained their perceptions about the development of symptom prediction models, including: (1) factors related to the AI model (compatibility or complexity), (2) nurse perception of patient benefit (relative advantage or observability), (3) improved clinical processes (relative advantage), (4) apprehension over model accuracy and impact (compatibility or trialability), and (5) implementation or adoption (trialability or complexity). Table 3 displays the identified themes, perceptions influencing adoption, and select quotes. Figure 1 highlights the development of themes from codes and the attributes of the Diffusion of Innovation Theory.

XSL•FO
RenderX

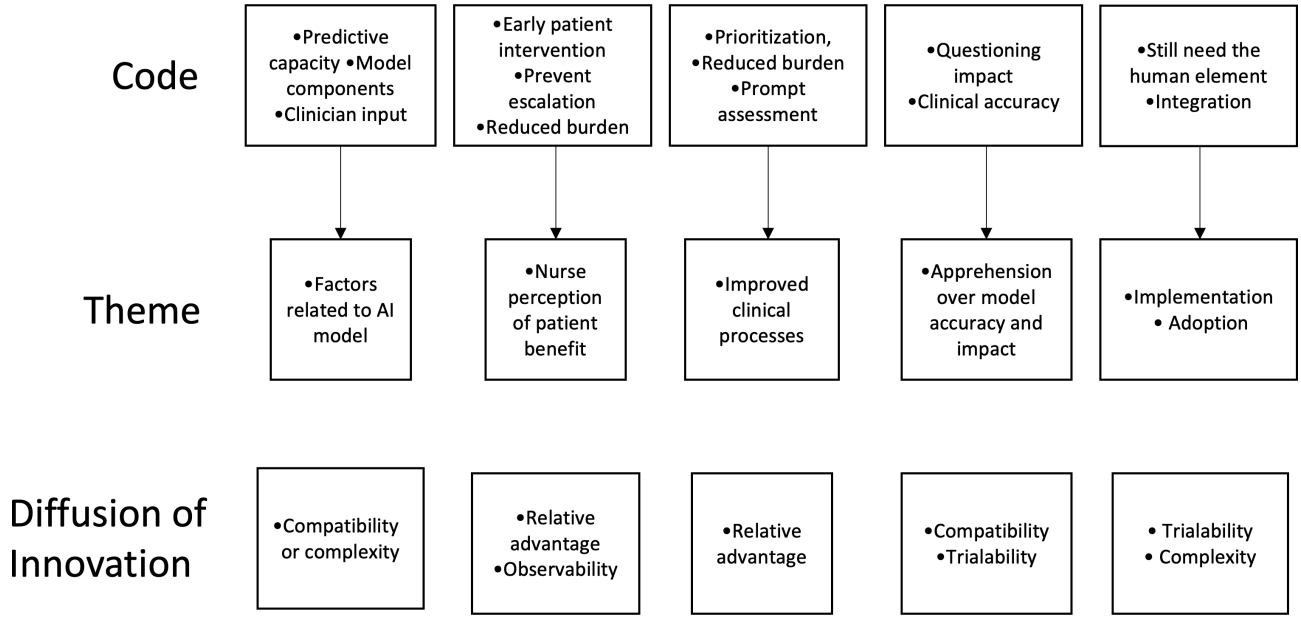
Table . Themes, codes, participant number, and quotes.

Themes	Perception attributes	Codes	Exemplar quote(s)
Factors related to the AI[a] model	• Compatibility • Complexity	• Predictive capacity • Model components • Clinician input	*I think it might help if the algorithm says the patient is likely to develop a fever. What would I do with that information if there were no infectious symptoms? It might be helpful if managing a lot of patients helps someone to rise to the top for check-ins.* [Participant 11] *Depends on where they get their data from. We already know when chemotherapy induced nausea will occur... so if AI is using appropriate standardized resources, I would be fine with that.* [Participant 8]
Nurse perception of patient benefit	• Relative advantage	• Early patient intervention • Prevent escalation • Reduced burden	*...a lot of time patients wait until symptoms are worse to call. We can intervene sooner.* [Participant 14] *Oncology patients can deteriorate quickly, this could help get symptoms before out of control to avoid ER/ hospital visit.* [Participant 1] *If a [model] did exist I would highly recommend [it] to prevent or mitigate life events to help prevent life or death.* [Participant 6]
Improved clinical processes	• Relative advantage	• Prioritization • Reduced burden • Prompt assessment	*Prioritization should occur digitally, rather than me doing it, adding more to my work burden.* [*Participant 12*] *Delays in care, due [to patient] burden of calling in. The RN reaching out directly could increase patient satisfaction and response.* [Participant 12]
Apprehension over model accuracy and impact	• Compatibility • Trialability • Observability	• Questioning impact • Clinical accuracy	*Based on clinical practice, you can usually pinpoint those patients anyway.* [Participant 12] *I don't really know if I would fully trust every time until it proves itself.* [Participant 3]
Implementation or adoption	• Complexity • Trialability • Observability	• Still need the human element • Integration • Workflow • Communication of model output (eg, notifications, text) • Ease of use • Alert fatigue	*AI can't supersede one-on-one contact.* [Participant 10] *[I] would want notifications if they are interruptions… I would want [them] to be relevant.* [Participant 12]

[a]AI: artificial intelligence.

XSL•FO
RenderX

Figure 1. Codes, themes, and diffusion of innovation category. AI: artificial intelligence.



Factors in the AI Model

Themes focused heavily on nurses' ability to understand the factors within the model and test it to assess its predictive capacity, components, and provide input into its development. Predictive capacity refers to the model's ability to make accurate assessments of future behavior. For example, participants emphasized the importance of the model being accurate and relevant to the patient's clinical presentation, reflecting the need for compatibility with existing systems. For example, one participant stated:

> *[I would want to know] … what led to notification, reason behind alert… what was their trend before… algorithm that shows patients who exhibit X also show Y, in the context of what's going on with the specific patient.* [Participant 15]

Participants also commented on specific factors necessary as model components, such as temperature and respiration. Nurses strongly emphasized the importance of involving oncology clinicians in the development of the AI model, highlighting their need to understand its compatibility with current systems.

Patient Benefit

Another theme identified was the benefits to patients, which included codes for early patient intervention, prevention of escalation, and reduced patient burden, all of which are compatible with current systems. Early patient intervention, as noted by many participants, was identified as a benefit of AI-based symptom management and is defined as having contact with the patient in a manner that occurs earlier than standard care as a relative advantage. For example, participants noted that a model could allow them "to intervene earlier before symptoms progress into dangerous situations" (Participant 2) or "prevent hospitalization and improve quality of life by managing symptoms at home" (Participant 8). Early intervention is the mechanism by which ePRO alerting systems have effectively decreased escalations of care from a current setting to a higher level of care, such as an emergency room. Other participants disagreed that patients would benefit more than they already do, with one participant stating:

> *I don't think the AI model will provide much additional information...Nurses already watch for specific symptoms.* [Participant 13]

Nurse participants also focused heavily on reducing the burden of cancer care delivery on patients. They highlighted the fact that the combination of early intervention, for example, early symptom detection, can prevent later-stage symptoms and care escalations, thereby improving the experience of cancer care, which aligns with the goals of current systems, but may also represent perceived advantages over the current system.

Enhanced Clinical Processes

Participants could envision that an AI-based symptom model may enhance the process of clinical care by improving prioritization and response times, thereby facilitating the prompt assessment of clinically significant changes in a manner superior to current systems. Additionally, participants felt that the process of care could decrease clinical burden, for example, stating "being able to streamline information would be helpful" (Participant 7). Nurses also reported wanting to reduce the patient's need for reporting and the burden of care escalations to clinicians. However, some participants also expressed concerns that the model could increase clinical burden and highlighted concerns about complexity, noting:

> *The nurse will have to contact the patient. Just because they have an alert doesn't mean they will have the symptoms.* [Participant 10]

Model Accuracy

Participants also reported it would be imperative to test the model to verify its accuracy, noting that nurses would be more likely to use a model they could participate in testing. Participants cautioned that patient engagement may influence the clinical accuracy of the tool. Many nurses have progressed beyond the initial knowledge stage and are now considering not

XSL•FO
RenderX

only whether, but also how, systems should adopt AI symptom models. Nurses have experience in integrating new technologies into clinical practice; as such, they understand the importance of accepting innovation to facilitate its diffusion and optimal use. Nurse participants also reported some apprehension about the use of AI models. Specifically, participants voiced concerns about the effectiveness and clinical use. Participants noted that training the model with the correct inputs would be crucial in confirming the model's accuracy.

Implementation Processes

Nurses' comments emphasized that decision-making also depends on the practical implementation of AI-based models. Evaluation and trialability of escalation alerts would be necessary for both initial and long-term adoption. Participants reported that AI-based predictive alerts for symptom management will not replace human nurse assessment and response. Participants also noted that the integration into the workflow needs to be seamless. There were many comments related to the importance of ensuring that communication of model output to nurses and other clinical staff does not increase the time burden, though many thought it would. For example, one participant noted, "I don't know it will save time, [it] may add time, but that is the sacrifice to catch something early" (Participant 12). Furthermore, most participants expressed a firm belief that a model could be easy to use and would not contribute to alert fatigue.

Discussion

Principal Findings

The majority of our sample of nurses agree with statements that support the use of AI-based symptom models, reflecting nurses' belief that these models may represent a relative advantage to current practice. The themes that nurse participants identified as essential to the adoption of AI symptom models aligned with the Diffusion of Innovation Theory. Nurses have recognized that the compatibility of AI-based symptom models holds promise for predicting, detecting, and enabling a response to changes in patient symptoms. Nurses' strong agreement to receive symptom information via new models revealed an overall favorable view of this type of model and alignment with existing values. These models align with nurses' strong commitment to providing patients with the best possible care, and by fostering the potential for AI-based symptom management models to improve patient care. Specific benefits identified by participants include improving clinician response by increasing the information clinicians receive and reducing patient burden through the elimination of unnecessary reporting or care escalations. This type of agreement indicates that nurses have progressed beyond the knowledge stage in the innovation process, toward identifying the necessary information to adopt the use of models. Overall, oncology nurses have positive views regarding the perceived advantages for patients and the compatibility with current care. This study demonstrates that many nurses have positive perceptions of the advantages and usability of AI-based symptom models and are now considering the implementation and use beyond the potential value.

Despite support for adoption, nurses urged caution in the development and implementation of these models. In particular, nurses emphasized the importance of involving end-users in the development, pilot testing, and implementation of these models, as this will help determine their value and appropriate integration into clinical workflow, thereby facilitating their adoption. Nurses have experience in integrating new technologies into clinical practice; as such, they understand the importance of accepting innovation to facilitate its diffusion and optimal use. This aligns with a framework developed for designing and implementing AI models from a systematic review, which recommends the inclusion of health care providers in development and implementation [8].

Nurse participants recognized the importance of trialability through accurately training and testing models, as well as ensuring that the data sources used are adequate to positively impact patient outcomes. Nurses strongly felt that model development requires the careful selection of clinically appropriate inputs, such as the inclusion of temperature and laboratory values, to support clinically accurate results. Confirming models that are appropriate for the input data and the desired outcomes is necessary for accuracy. As frontline users, nurses who currently assess patient symptoms should be included in model factor selection and testing. Often, these models are developed in collaboration with other clinical providers, and yet nurses will be the ones to receive the alerts and need to triage them. Trialability and observability for nurses, not just physicians, are keys to adoption. These themes are consistent with the recommendations for transparency in the development of AI-based clinical models, ensuring that both clinicians and patients understand and agree on the inputs to the model [23]. Creating transparent and explainable models is a step toward combating the perpetuation of healthcare bias in AI models and will facilitate long-term adoption [24,25].

While participants identified the need to understand model inputs and testing, they also reported a need to see the model's impact on outcome to feel confident in making clinical decisions based on the model, again underscoring the importance of observability. Model outcome achievement depends on the implementation of models as designed. For this to occur, there must be transparency and trialability of both inputs and clinical outcomes. For example, our early work developing a predictive model demonstrated the ability to predict symptom escalation more accurately in short intervals than at longer intervals [10]. Transparency will enable clinical teams to implement models for the purpose they were developed, thereby supporting accuracy. Efforts to transform and train models for additional uses will need to follow proper rigor to ensure the models are adapted and updated effectively. Transparency and inclusion in development will enable oncology nurses to effectively use AI-based models.

Experienced oncology nurses in our sample reported both a strong interest in using and some reluctance to immediately trust AI-based symptom models. While involving nurses in the development and implementation will facilitate trust, oncology nurses may lack the education and training to understand how these models work. An extensive national survey of nurses revealed that only 30% of nurses are aware of how AI is used

XSL•FO
RenderX

in nursing practice [26]. Although information regarding the use and daily applications has increased in the last several years, this highlights the need to provide AI education to both students and to disseminate it to nurses at the point of care delivery. Future work should specifically evaluate the education needs of oncology nurses regarding AI-based symptom models.

With many clinical symptom escalation models still in development, gaining a clear understanding of nurse perceptions regarding the use, decision to adopt, and maintenance of these models is essential. Our examination revealed that oncology nurses share similar concerns to those documented in the literature regarding the use of clinical predictive models, including alert fatigue and increased time burden, which represent a source of complexity [27]. Additional barriers to adoption of AI technology in healthcare include ethics, technological considerations such as data access and infrastructure, and liability and regulatory issues [28]. However, as evidenced in our results, nurses also hold favorable perceptions that these models have advantages and align with current treatment values, prioritizing the reduction of cancer symptom burden. Implementation strategies that could be used to overcome adoption barriers include, but are not limited to, identifying implementation champions as well as ensuring adequate interpretability of the model [29,30]. AI-based symptom models have the potential to improve patient outcomes and enhance clinical processes when implemented thoughtfully into the clinical workflow. As the end users of AI-based symptom management models, nurses should be involved as content experts, beginning with model development and continuing through the design, integration, and evaluation of the model into workflows, to maximize both short-term implementation and long-term adoption. However, additional research is needed to identify which implementation strategies are effective at promoting the adoption and sustained use of AI-based symptom management models.

Limitations

We sought to elicit oncology nurses' initial thoughts on AI-based symptom prediction models. We should continue to inductively evaluate nurses' adoption of AI. Our exploration was limited by a small sample size and a homogenous population that was skewed by age (older) and education levels (high) that impacts the generalizability of our findings. This may be attributable to our convenience sampling approach and the fact that participants recommended other individuals who were recruited to participate, possibly introducing selection bias. Both the sample skew and homogeneity may have influenced the overall positive perceptions of an AI model for use in symptom management. Additionally, the use of unrecorded and note-based qualitative data analysis may have limited our ability to accurately assess content-level saturation; however, we believe that the detailed notes taken by interviewers permitted accurate assessment of content-level saturation. Finally, while we used the Rogers Diffusion Theory of Innovation to improve the descriptive analysis of the qualitative themes, it may limit our understanding of the responses and future work. Further work should survey a larger sample of nurses to understand oncology nurses' perceptions of AI symptom models and consider the impact of education levels on their views regarding AI. Additionally, future work should highlight the gaps in nurses' understanding of the application of AI in clinical care. The inclusion of end users in the design and testing of AI-based models facilitates adoption, and additional work should concentrate on and focus on implementation processes, which include user-centered design testing of best practices, such as alerting, alert visualization, and responses to care.

Conclusions

Overall, nurses showed a positive attitude toward the adoption of AI-based symptom models, particularly highlighting the perceived advantages of such models and their compatibility with nurses' goals of enhancing the patient experience. Proper use of AI symptom prediction models creates the opportunity to decrease the burden of patient reporting of cancer symptoms, improve clinician responsiveness, and enable prompt intervention to reduce unnecessary care and escalations. To facilitate the seamless integration of AI-based symptom models, thoughtful inclusive design strategies must include end users to test and modify transparent clinical models for long-term adoption.

Acknowledgments

There was no use of artificial intelligence, including generative artificial intelligence, in the development of the study or in the generation of text, tables, figures, or informational content of this manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: BN, EAS, AS, JF, KM
Data curation: BN, EAS
Formal analysis: BN, EAS
Methodology: BN, EAS
Software: BN, EAS
Validation: AS, JF, KM
Visualization: EAS

XSL•FO
RenderX

Writing–original draft: BN, EAS, KM
Writing–review and editing: BN, EAS, AS, JF, KM

Bridget Nicholson is a consultant for Daymark Health. All other authors have nothing to disclose.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Symptom Care at Home (SCH) artificial intelligence (AI) model interview script.
[DOCX File, 33 KB - nursing_v9i1e82283_app1.docx]

Checklist 1
GRAMMS Checklist.
[DOCX File, 8 KB - nursing_v9i1e82283_app2.docx]

References

1. Mooney KH, Beck SL, Wong B, et al. Automated home monitoring and management of patient-reported symptoms during chemotherapy: results of the symptom care at home RCT. Cancer Med 2017 Mar;6(3):537-546. [doi: 10.1002/cam4.1002] [Medline: 28135050]
2. Basch E, Deal AM, Dueck AC, et al. Overall survival results of a trial assessing patient-reported outcomes for symptom monitoring during routine cancer treatment. JAMA 2017 Jul 11;318(2):197-198. [doi: 10.1001/jama.2017.7156] [Medline: 28586821]
3. Rocque GB, Caston NE, Hildreth K, et al. Engagement among diverse patient backgrounds in a remote symptom monitoring program. JCO Oncol Pract 2024 Oct;20(10):1426-1435. [doi: 10.1200/OP.24.00066] [Medline: 38917385]
4. Yamaguchi K, Higashiyama N, Umemiya M, et al. Electronic patient-reported outcomes as digital therapeutics for patients with cancer: a narrative review of current practices and future directions. Int J Clin Oncol 2025 Jan;30(1):1-16. [doi: 10.1007/s10147-024-02651-8] [Medline: 39549219]
5. Daly B, Nicholas K, Flynn J, et al. Analysis of a remote monitoring program for symptoms among adults with cancer receiving antineoplastic therapy. JAMA Netw Open 2022 Mar 1;5(3):e221078. [doi: 10.1001/jamanetworkopen.2022.1078] [Medline: 35244701]
6. Innominato PF, Komarzynski S, Dallmann R, et al. Impact of assessment frequency of patient-reported outcomes: an observational study using an eHealth platform in cancer patients. Support Care Cancer 2021 Nov;29(11):6167-6170. [doi: 10.1007/s00520-021-06262-1] [Medline: 33963910]
7. Zeinali N, Youn N, Albashayreh A, Fan W, Gilbertson White S. Machine learning approaches to predict symptoms in people with cancer: systematic review. JMIR Cancer 2024 Mar 19;10:e52322. [doi: 10.2196/52322] [Medline: 38502171]
8. de Hond AAH, Leeuwenberg AM, Hooft L, et al. Guidelines and quality criteria for artificial intelligence–based prediction models in healthcare: a scoping review. NPJ Digit Med 2022 Jan 10;5(1):2. [doi: 10.1038/s41746-021-00549-7] [Medline: 35013569]
9. Meijerink LM, Dunias ZS, Leeuwenberg AM, et al. Updating methods for artificial intelligence–based clinical prediction models: a scoping review. J Clin Epidemiol 2025 Feb;178:111636. [doi: 10.1016/j.jclinepi.2024.111636] [Medline: 39662644]
10. Finkelstein J, Smiley A, Echeverria C, Mooney K. AI-driven prediction of symptom trajectories in cancer care: a deep learning approach for chemotherapy management. Bioengineering (Basel) 2024 Nov 20;11(11):1172. [doi: 10.3390/bioengineering11111172] [Medline: 39593830]
11. Flora C, Tyler J, Mayer C, et al. High-frequency temperature monitoring for early detection of febrile adverse events in patients with cancer. Cancer Cell 2021 Sep 13;39(9):1167-1168. [doi: 10.1016/j.ccell.2021.07.019] [Medline: 34388378]
12. Miller M, McCann L, Lewis L, et al. Patients' and clinicians' perceptions of the clinical utility of predictive risk models for chemotherapy-related symptom management: qualitative exploration using focus groups and interviews. J Med Internet Res 2024 Jun 20;26:e49309. [doi: 10.2196/49309] [Medline: 38901021]
13. Wieben AM, Alreshidi BG, Douthit BJ, et al. Nurses' perceptions of the design, implementation, and adoption of machine learning clinical decision support: a descriptive qualitative study. J Nurs Scholarsh 2024 Jun;57(1):82-94. [doi: 10.1111/jnu.13001] [Medline: 38898636]
14. Burford JS, Booth RG, McIntyre A. Exploring the intersection of nursing leadership and artificial intelligence: scoping review. JMIR Nurs 2025 Nov 14;8(1):e80085. [doi: 10.2196/80085] [Medline: 41237291]
15. Reading Turchioe M, Austin R, Lytle K. Opportunities and challenges for digital health and artificial intelligence to support nurses: results of a survey of nursing informaticists. Comput Inform Nurs 2025 Jul 1;43(7):e01279. [doi: 10.1097/CIN.0000000000001279] [Medline: 40052895]

XSL•FO
RenderX

16. Reading Turchioe M, Pepingco C, Ronquillo C, et al. Education, empowerment, and elevating nursing voices: nursing informatics leaders' perspectives on the path forward with artificial intelligence in nursing. Nurs Outlook 2025;73(5):102484. [doi: 10.1016/j.outlook.2025.102484] [Medline: 40652598]
17. Rogers EM. Diffusion of Innovations, 5th edition: Free Press; 2003.
18. Green CA, Duan N, Gibbons RD, Hoagwood KE, Palinkas LA, Wisdom JP. Approaches to mixed methods dissemination and implementation research: methods, strengths, caveats, and opportunities. Adm Policy Ment Health 2015 Sep;42(5):508-523. [doi: 10.1007/s10488-014-0552-6] [Medline: 24722814]
19. Koo M, Yang SW. Likert-type scale. Encyclopedia 2025 Feb;5(1):18. [doi: 10.3390/encyclopedia5010018]
20. Heale R, Forbes D. Understanding triangulation in research. Evid Based Nurs 2013 Oct;16(4):98-98. [doi: 10.1136/eb-2013-101494] [Medline: 23943076]
21. Teddlie C, Tashakkori A. Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences: Sage Publications, Inc; 2009. [doi: 10.1016/J.AENJ.2009.07.004]
22. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
23. Hassan N, Slight R, Morgan G, et al. Road map for clinicians to develop and evaluate AI predictive models to inform clinical decision-making. BMJ Health Care Inform 2023 Aug;30(1):e100784. [doi: 10.1136/bmjhci-2023-100784] [Medline: 37558245]
24. Benda NC, Novak LL, Reale C, Ancker JS. Trust in AI: why we should be designing for APPROPRIATE reliance. J Am Med Inform Assoc 2021 Dec 28;29(1):207-212. [doi: 10.1093/jamia/ocab238] [Medline: 34725693]
25. Rosenbacke R, Melhus Å, McKee M, Stuckler D. How explainable artificial intelligence can increase or decrease clinicians' trust in AI applications in health care: systematic review. JMIR AI 2024 Oct 30;3:e53207. [doi: 10.2196/53207] [Medline: 39476365]
26. Swan BA. Assessing the knowledge and attitudes of registered nurses about artificial intelligence in nursing and health care. Nurs Econ 2021;39(3):139. [doi: 10.62116/NEC.2021.39.3.139]
27. Lee TC, Shah NU, Haack A, Baxter SL. Clinical implementation of predictive models embedded within electronic health record systems: a systematic review. Informatics (MDPI) 2020 Sep;7(3):25. [doi: 10.3390/informatics7030025] [Medline: 33274178]
28. Ahmed MI, Spooner B, Isherwood J, Lane M, Orrock E, Dennison A. A systematic review of the barriers to the implementation of artificial intelligence in healthcare. Cureus 2023 Oct;15(10):e46454. [doi: 10.7759/cureus.46454] [Medline: 37927664]
29. Nair M, Svedberg P, Larsson I, Nygren JM. A comprehensive overview of barriers and strategies for AI implementation in healthcare: mixed-method design. PLoS One 2024;19(8):e0305949. [doi: 10.1371/journal.pone.0305949] [Medline: 39121051]
30. Staes CJ, Beck AC, Chalkidis G, et al. Design of an interface to communicate artificial intelligence-based prognosis for patients with advanced solid tumors: a user-centered approach. J Am Med Inform Assoc 2023 Dec 22;31(1):174-187. [doi: 10.1093/jamia/ocad201] [Medline: 37847666]

Abbreviations

AI: artificial intelligence
ePRO: electronic patient-reported outcome
GRAMMS: Good Reporting of a Mixed Methods Study
ML: machine learning

Edited by J Sarvestan; submitted 12.Aug.2025; peer-reviewed by C Bhagat, JCL Chow, Q Wang; revised version received 31.Dec.2025; accepted 01.Jan.2026; published 04.Feb.2026.

Please cite as:
Nicholson B, Sloss EA, Smiley A, Finkelstein J, Mooney K
Perception of AI Symptom Models in Oncology Nursing: Mixed Methods Evaluation Study
JMIR Nursing 2026;9:e82283
URL: https://nursing.jmir.org/2026/1/e82283
doi:10.2196/82283

© Bridget Nicholson, Elizabeth A Sloss, Aref Smiley, Joseph Finkelstein, Kathi Mooney. Originally published in JMIR Nursing (https://nursing.jmir.org), 4.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction

XSL•FO
RenderX

in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Recommendations From a Health Care Provider's Perspective for Better Implementation of Telemedicine and Telenursing: Descriptive Qualitative Approach

Rr Tutik Sri Hariyati[1,2*], SKp, MARS, DN; Min-Huei Hsu[3*], MD, PhD; Hanny Handiyani[1,2], SKp, MKep, DN; Muhammad Hanif Amiruddin[4], SE, MBA; Tuti Afriani[1], SKp, MKep, DN; Dara Ayu Wardhani[2], SKep, Ns; Satria Dwi Pranoto[2], SKep, Ns; Andi Amalia Wildani[1,2], SKep, Ns, MKep; Arieny Rizafni[1], SKep, Ns, MKep; Citra Hafilah Shabrina[1], SKep, Ns, Mkep

[1]Faculty of Nursing, Universitas Indonesia, Health Cluster, Building E7, Depok, West Java, Indonesia

[2]Hospital of Universitas Indonesia, Depok, West Java, Indonesia

[3]Graduate Institute of Data Science, College of Management, Taipei Medical University, Taipei, Taiwan

[4]International P.hD. Program in Biotech and Healthcare Management, College of Management, Taipei Medical University, Taipei, Taiwan

[*]these authors contributed equally

Corresponding Author:
Rr Tutik Sri Hariyati, SKp, MARS, DN
Faculty of Nursing, Universitas Indonesia, Health Cluster, Building E7, Depok, West Java, Indonesia

Abstract

Background: Digital technology-based health care services like telemedicine and telenursing are expanding quickly globally. Although the use of telemedicine and telenursing has shown promise, there are a number of obstacles to overcome.

Objective: This study aimed to determine the experiences of stakeholders with using telenursing and telemedicine during and after the COVID-19 pandemic.

Methods: This study combined descriptive qualitative analysis using a focus group discussion (FGD) and validation using a review of the literature. The literature search used a convenience approach with the keywords benefits, advantages, disadvantages, obstacles, and constraints of telenursing or telemedicine to find relevant journal articles. Two FGDs were conducted during the COVID-19 period, namely 2020 - 2022, with the first FGD held in 2022. The second FGD was conducted in 2023 - 2024 to determine the implementation of telenursing and telemedicine during the recovery period, with data collection in 2024. The FGDs were conducted via the online Zoom platform, and a total of 23 participants were selected based on purposive sampling. People involved in the development and implementation of telemedicine or telenursing were included. Data analysis was conducted in 2 stages: the first using an inductive method, which was followed by triangulation using a deductive method. The results were verified and validated based on the reviewed literature.

Results: We identified three themes: (1) health care providers have experience using telemedicine/telenursing; (2) telemedicine and telenursing have many benefits; and (3) telemedicine and telenursing have 5M (method, man, money, material, and machine) factors that must be anticipated. Key statements were identified and validated, then matched with the literature review.

Conclusions: The benefits and barriers of telemedicine and telenursing were identified. For telemedicine to run smoothly, preparation of 5M elements is necessary. Implications for health care services necessitate the creation of operational guidelines for the implementation of telemedicine and telenursing.

(JMIR Nursing 2026;9:e90610) doi:10.2196/90610

KEYWORDS

human resources; 5M; method, man, money, material, and machine; telemedicine; telenursing; regulation

Introduction

The role of digital health technology in all countries grew rapidly during the COVID-19 pandemic. Telemedicine and telenursing, which provide remote service facilities and can reduce physical contact that causes transmission of the COVID-19 virus, were very important during the COVID-19 pandemic [1,2]. The use of telemedicine when treating patients with COVID-19 allows consultations and administering drugs without going to health facilities, strengthening the health system. This is beneficial for the handling of COVID-19 because it can reduce the number of COVID-19 exposures [3]. Some telemedicine and telenursing technologies were used in the

XSL•FO
RenderX

COVID-19 era for treatment consultations, monitoring, follow-up after hospitalization, providing counseling during isolation, and providing health education [4].

In Indonesia, telemedicine and telenursing are needed due to its archipelagic geography and numerous seas. Digital systems are essential for delivering services, as traveling to health care facilities can take more than 1 hour. Several studies provide evidence that telemedicine and telenursing are beneficial for increasing access and connecting urban and urban areas [5,6]. Telenursing can be used to direct when to take medications, to describe how to treat wounds, to help patient movement, and as a medium for health consultation for patients [5,7-9]. Facilitating the reach to more areas can save transportation and direct health care costs [10-14].

In line with the results of previous studies, telenursing can be used to monitor patients and provide coverage of services while reducing exposure to infection [10]. Several other studies also conveyed that telemedicine and telenursing can improve quality of life, reduce service costs, and be time efficient because patients do not have to go to health services; in addition, it can reduce the costs of consultation between health workers [11].

Since telenursing's readiness frequently affects success, a 5M (man, money, method, material, and machine) preparation strategy is crucial. This 5M strategy was developed to prepare for telenursing and telemedicine. Poor community literacy competency and inadequate infrastructure limit the use of telemedicine [12,14,15]. There are many obstacles to the use of telenursing and telemedicine, such as limited patient digital literacy [14,16], and a lack of facilities can decrease the enthusiasm for telemedicine after the COVID-19 era [13,14]. The absence of full regulations, such as a credentialing process for the telemedicine service provider industry, is an obstacle in Indonesia. A previous study stated that regulation was needed for telemedicine implementation. In addition, certification and credentialing for human resources who perform telemedicine practices are not available. The issue of patient data security and confidentiality is also present in Indonesian telehealth practices [12,13].

In some areas, the use of telehealth, including telemedicine and telenursing, decreased after the COVID-19 era due to patient preferences [14,17]. This online consultation model boomed in the COVID-19 era, but has gradually decreased since. Many consumers returned to direct services because they could meet directly with care providers such as doctors and nurses. According to prior research, a more accurate disease diagnosis, more accurate and better patient examination by a physician, and more accurate and better disease treatment are the primary reasons for selecting in-person appointments [17,18]. This study aimed to determine the experiences of stakeholders with using telenursing and telemedicine during and after the COVID-19 pandemic. The study also aimed to identify the readiness of 5M aspects to support telenursing implementation. The research results were expected to provide recommendations for the ideal implementation of telenursing and telemedicine.

Methods

Study Design

This study used a descriptive qualitative technique to investigate the perceptions of health stakeholders for managing the advantages of telemedicine and telenursing and to determine the weaknesses of 5M factors. In this study, we excluded market aspects and focused our evaluations on how workflows (methods), people, money, machines, and materials were executed.

The study results will provide recommendations for better telemedicine and telenursing implementation by conducting a synthesis that combines a literature review and descriptive qualitative analysis of focus group discussions (FGDs). We searched for journal articles using a convenience approach with the following keywords: benefits, advantages, disadvantages, 5M obstacles, and constraints of telenursing or telemedicine. We posed 4 questions regarding the use of telenursing and telemedicine during 2 data collection periods, namely during the final period of the COVID-19 pandemic (2022 - 2023) and after the recovery period (2024).

We conducted 2 FGDs. The first was held in 2022, to represent the COVID-19 period (2020 - 2022). The second FGD was conducted in 2024 to determine the implementation of telenursing and telemedicine during the recovery period (2023-2024). The FGDs were conducted via the online Zoom platform. Data collection was conducted twice to examine perspectives during and after the COVID-19 pandemic. After the COVID-19 pandemic, there was a tendency for patients to prefer direct care at health facilities.

The questions asked during the FGDs were almost the same, with the aim of understanding the perspectives, benefits, and potential challenges that arise during the implementation of telehealth. These questions were as follows: (1) Describe the telemedicine or telenursing services you have utilized. (2) What are the prospects for telemedicine and telenursing implementation? (3) Could you explain the shortcomings experienced and 5M factors when using telemedicine and telenursing? and (4) What do you hope to achieve by utilizing telemedicine or telenursing?

Participants and Approach to Their Selection

Data were gathered through semistructured FGD with policymakers, hospital organizations, medical doctors, nurses, and health care providers. Purposeful sampling was used to ensure representation across different occupations and demographic groups. Participants were invited by sending a letter of application to become a participant to hospitals, clinics, health centers and health worker associations. The organization then appointed the sections that would participate in the FGD according to the requested criteria, namely having experience in implementing digital health, including telemedicine or telenursing.

We selected 12 FGD participants in 2022 and 11 FGD participants in 2024 via purposive sampling. Ideally, the FGD participants would have been the same individuals, but this study had a weakness in that participants in the first FGD could

XSL•FO
RenderX

not participate in the second FGD. Participants in 2022 and 2024 were different people because they were sent by the target organizations according to the research objectives. Some differences in the participants were due to changes in personnel and tasks of the organization during the COVID-19 period. The professions of the participants were diverse, namely nurses, midwives, doctors, lecturers, medical recorders, lawyers, and others. Additionally, involving a variety of stakeholders was expected to help identify potential and 5M barriers to telemedicine and to develop targeted recommendations to address them effectively.

Data Collection

Response bias was possible because the data were gathered through focus group conversations, where individuals might alter their answers in reaction to what other group members think. To reduce prejudice, the facilitator gave each participant information at the outset so they may voice their opinions independently of other participants' experiences. At the conclusion of the conversation, each member had the opportunity to share personal experiences that had not been discussed. Saturation was reached when no new topics nor information surfaced from the focus groups. One way to reach data saturation is to ask whether there are any unresolved issues from the focus group.

The researcher's online Zoom account was used to conduct the FGD. Participants who were not speaking were muted to reduce noise while someone was speaking. Participants were invited to repeat what was said when the signal was clear again if signal interference was present. The atmosphere could be managed during the FGDs, and there were no issues with internet connectivity.

Analysis and Data Validity

Data analysis was conducted in 2 stages: (1) inductive method and (2) triangulation with the deductive method, which was verified and validated based on the reviewed literature.

In the first stage, the research team used an inductive content analysis on each communication from the FGD, which they verified without looking at the existing literature. Field comments were added to the transcribed audio data. To identify any writing errors and validate the comments made by the participants, the transcriptions were read twice. This methodical approach ensured that every facet of the data was carefully considered and comprehended, producing thorough and trustworthy results. The transcription stage was carried out immediately after the FGD was completed. The researcher re-examined all statements from participants in both FGDs. The transcription, transcribing, and coding processes in stages 1 and 2 were carried out in the same steps. The facilitator retrieved words by attaching parallel keywords, then grouped and compared all codes. From these grouped keywords, the principal researcher then discussed them with a research team consisting of 3 participants in the FGD who acted as field observers and research assistants. The keyword data were compared with the final field notes to gain consensus in order to obtain themes that ensured trustworthiness and transparency and were not biased toward only one researcher.

The researchers used a thematic approach to analyze qualitative data in 2022 and 2024. After obtaining initial codes, terms, and themes from the FGDs, we evaluated and reviewed these themes by comparing them with the entire dataset to ensure they accurately reflected the views of health care providers. During this stage, we analyzed the themes as they evolved across time periods, specifically comparing participants' telemedicine implementation during the peak of the COVID-19 pandemic in 2020 - 2022 with the subsequent postpandemic situation in 2023 - 2024. We then clearly defined and named each theme to establish its specific meaning and scope in relation to the research questions.

The information was then classified, extracted to identify significant statements, and analyzed to create meanings. After categorization, themes were developed from the data in each FGD. These themes were eventually reduced to codes and incorporated into a comprehensive analysis of the studied situation. The researchers counted the communication from participants in each phase of the FGD, grouped them by the same theme, and described them in the 2022 and 2024 FGD stages. In the end, the data are presented as study findings in the form of tables and discussion points.

In order to assure the rigor of our telemedicine FGD in 2022 and 2024, we undertook member checks to ensure credibility. In this study, we clarified remarks made by participants over the phone on the decrease in telemedicine use following the COVID-19 pandemic. Additionally, we performed peer debriefing by triangulating data from articles and national policies and addressing all themes. Additionally, we maintained confirmability by exercising researcher reflexivity and closely following the FGD guidelines, and we established dependability by guaranteeing the lack of researcher bias. Our FGD topics were cross-referenced against reputable peer-reviewed articles as part of our second strategy to ensure rigor. The triangulation stage involved analysis using a deductive approach, validating the existing themes with several previously conducted studies. This method aimed to compare existing findings with those in other studies to provide a comprehensive picture both abroad and domestically. This method was expected to provide better recommendations and a more comprehensive implementation of telenursing and telemedicine.

In the assessment of the overall readiness for telemedicine implementation, this study used triangulation by combining the results of the FGDs and theoretical foundations from journal articles synthesized through the keywords readiness, telenursing, or telemedicine. All relevant articles available as full text were read and analyzed using the telemedicine readiness approach of the 5M elements. The 5M dimensions identified in this literature review were aligned with qualitative themes extracted from the FGDs. The articles were found from the year 2005 to 2025 and from the Ministry of Health regulations in Indonesia. The articles and national regulations provided basic indicators and best practices for the 5M components; meanwhile, the FGD data offered a localized, real-world context regarding stakeholder perceptions and challenges. By cross-verifying the 2 datasets, we were able to comprehensively conclude the strengths, gaps, and readiness levels of facilities for adopting telemedicine or telenursing.

XSL•FO
RenderX

The weakness in this study's literature review is that it was not a systematic review but only used data from articles retrieved by searching research databases with the keywords telemedicine, telenursing, benefit, barrier, and COVID-19. We conducted a convenience search of journals through ProQuest and Google Scholar for articles published from 2005 through 2025. A total of 46 journal articles were analyzed and combined with the FGD results to obtain themes. Researchers reviewed the literature using unstructured (qualitative/narrative) and incidental data, involving the application of thematic and content analysis frameworks to synthesize diverse, nontraditional, and serendipitous sources. The steps included identifying core concepts through keyword extraction, underlying arguments, and key narratives. Coding was then performed, grouping insights into themes. Synthesis was then performed, connecting, comparing, and building themes. The Quality Rating System (QRS) consists of 21 explicit items covering study design, data collection, and analysis. It helps authors write better papers, assists reviewers with evaluation, and guides readers when appraising evidence, ensuring full clarity and trust in the study. Finally, the Standards for Reporting Qualitative Research (SRQR) checklist was used to improve the transparency of all aspects of the qualitative research [19].

Ethical Considerations

All participants provided written informed consent, and the participants were aware of the research goals, advantages, and methods. The team members maintained anonymity throughout the completion of the FGD, guaranteed confidentiality, and discussed their experiences with digital implementation. The authors adhere to ethical norms when publishing the improvement study's data by protecting both personal and organizational privacy and data confidentiality when implementing telemedicine and telenursing. This study was approved by the Ethical Board of the Faculty of Nursing, Universitas Indonesia: KET-138/UN2.F12.D1.2.1/PPM.00.02/2022 and KET 180/UN2.F12.D.1.2.1/PPM.00.02/2024.

Results

Attendance at focus groups has been an issue of concern, which was also identified in our research. FGDs were attended by 12 participants in 2022 and 11 participants in 2024 (Table 1). Participants included nurses, nurse practitioners, medical doctors, lecturers, and a lawyer. Verbatim responses were compiled per question and combined in participant groups in 2022 (code -22), and 2024 (code -24). After being delivered verbatim, at the end of the analysis, theme mapping was conducted for each question.

Participants' statements in the 2022 and 2024 FGDs were then grouped, resulting in three themes: (1) health care providers have experience using telemedicine/telenursing, (2) telemedicine and telenursing have many benefits, and (3) telemedicine and telenursing have 5M factors that must be anticipated. Complete participant statements and keywords were grouped and grouped into themes, as described in Table 2.

XSL•FO
RenderX

Table . Focus group discussion participant characteristics in 2022 and 2024.

Year and participant code	Age (years)	Sex	Education	Job role
2022				
P1	33	Male	Nurse specialist	Lecturer
P2	51	Male	S1+ profession	Nurse practitioner
P3	33	Female	Master	Nurse
P4	49	Female	Medical doctor	Doctor
P5	43	Male	Master	Nurse
P6	70	Male	Professor	Lecturer
P7	51	Female	Master	Nurse practitioner
P8	29	Male	Master	NGO[a]
P9	44	Female	S1+ profession	Nurse
P10	36	Male	Master	Lawyer
P11	54	Male	Master	Nurse manager
P12	40	Male	S1+ profession	Nurse
2024				
P1	57	Male	Medical doctor	Medical doctor
P2	37	Female	S1+ profession	Nurse
P3	48	Female	Medical doctor	Medical doctor
P4	35	Male	S1+ profession	Nurse
P5	29	Female	S1+ profession	Nurse
P6	37	Male	S1+ profession	Nurse
P7	30	Female	S1+ profession	Nurse
P8	36	Female	S1+ profession	Midwife
P9	51	Female	S2+ profession	Nurse practitioner
P10	32	Female	S1+ profession	Nurse
P11	47	Female	Medical doctor	Medical doctor

[a]NGO: nongovernmental organization.

XSL•FO
RenderX

Table . Telenursing and telemedicine themes during and after the COVID-19 pandemic.

Themes and years	Supporting input
Health care providers have experience using telemedicine/telenursing.	
2022	• Terminology telenursing and telemedicine (P6) • Complement of digital transformation (P4; P6)
2024	• Telenursing part of telemedicine (P1) • Telemedicine and insurance (P10) • As digital transformation (P1)
Telemedicine and telenursing have benefits.	
2022	• Education for wound care (P2; P4) • Support for education and consultation (P1; P6; P7; P9) • Facilities for religious and spiritual studies (P1; P6; P9; P7) • Decrease direct cost (P9; P7) • Facilitating access and speeding up the consultation (P1; P6; P9; P7) • Consultation and education (P6; P7; P8) • Send medication to a patient in isolation (P2) • Decreasing virus transmission (P1; P2; P6; P7; P8) • Consultation during the pandemic (P2; P6; P7; P8) • Monitoring and follow-up (P4; P11)
2024	• Complement digital transformation (P1; P2; P3) • Consultation and education (P2; P3; P4; P8) • Facilitating access (P1; P8; P9) • Decrease cost for direct treatment (P9) • Referral from public health services to hospital (P4; P6) • Support for education, consultation, and disaster conditions (P4; P8; P9) • Education for wound care (P2:P4:P9)
Telemedicine and telenursing have 5M[a] factors that must be anticipated.	
Method factors	
2022	• Method of delivery (P1; P5; P11; P12) • Payment guide (P7; P10-11) • Security and privacy (P6; P10; P11; P3) • Ethical conduct guide (P6; P11) • Health services credentialing (P10; P5) • Satisfaction and comfort (P9) • Not for all, combination, education, follow-up care (P6; P10; P8) • Method: need caring (P1; P12; P5: P11)
2024	• Regulation (P1; P4; P9; P10) • Payment national guide (P1; P4) • Security and privacy (P1; P4) • Kind of services • Not for all services, education, follow-up care (P1; P8; P9; P11)
Man factors	
2022	• Competencies (P1; P5; P9; P7) • Credentialing for health provider (P10; P5) • Capabilities (P1; P3; P5; P6) • Training and socialization (P5; P1; P3; P6)
2024	• Competencies, training, and socialization (P1; P2; P3; P4; P9) • Capabilities (P1; P2; P3; P4; P9)
Money factors	
2022	• Training cost (P5; P1; P3; P6) • Infrastructure and investment (P8)
2024	• Training cost (P1; P2; P3; P9) • Infrastructure and investment (P4; P1; P2; P5; P7)

XSL•FO
RenderX

Themes and years	Supporting input
Material and machine factors	
2022	• Infrastructure for integrated video calls and Zoom (P8; P10) • Access (P4; P7) • Network and teledevice (P8; P10; P11) • Health facility credentialing (P10; P5) • Integrated with EMR[b] (P10; P3)
2024	• Infrastructure: chat, AI[c], device, application, EMR (P4; P1; P2; P5; P7) • Network and teledevice (P7) • Integrated with EMR (P1)

[a]5M: people/man, methods, money, materials, and machines.

[b]EMR: electronic medical record.

[c]AI: artificial intelligence.

The first theme was "health care providers have experience using telemedicine/telenursing." Participants discussed that telenursing was part of telemedicine. The existence of telemedicine and telenursing is already largely known by stakeholders. FGD participants in both 2022 and 2024 familiar with telemedicine and telenursing that help with patient care used the technology. Some implemented it in hospitals, intensive care units, and health centers, as described in Textbox 1.

Textbox 1. Participant input supporting the "health care providers have experience using telemedicine/telenursing" theme.

- Terminology in the health regulation; telemedicine is a clinical service provided by a health service provider through or using information technology.; telenursing as a part of telemedicine (P1-24; P6-22).
- Previously, a telenursing and telemedicine for intensive care units were built in collaboration with the Japan International Cooperation Agency. It is no longer running (P8-24).
- Telemedicine in primary health care following Badan Penyelenggara Jaminan Sosial Kesehatan assurance (P10-24)
- Telemedicine or telenursing will complement the digital transformation in Indonesia (P4-22; P6-22; P1-24).
- It was explained that, in the future, there will be development of telenursing and telemedicine (P4-24).

The second theme was "telemedicine and telenursing have many benefits." The benefits of telemedicine and telenursing include increasing access to services. During the pandemic, it could help reduce the transmission of COVID-19. In addition, it can help with educational media, consultations, wound care, follow-up care, and spiritual care reinforcement. Participant inputs that represent this theme are shown in Textbox 2.

Textbox 2. Participant input supporting the "telemedicine and telenursing have many benefits" theme.

- Speed up consultations, especially during the COVID-19 pandemic (P1-22; P2-22; P6-22; P7-22; P8-22)
- Decreasing virus transmission during the COVID-19 pandemic; send drugs (P6-22; P7-22; P8-22; P2-22)
- Consultation, education (P6-22; P7-22; P8-22; P2-22; P2-24; P3-24; P4-24; P8-24)
- Telenursing will bring patients closer to nurses, facilitating access. The service is without walls (P1-22; P6-22; P9-22; P7-22; P1-24; P8-24; P9-24).
- Telenursing will decrease costs for direct treatment (P9-22; P7-22; P9-24).
- Make referrals easier from public health services to hospital (P4-24; P6-24)
- Telenursing can be used for long-distance wound care consultations. Consultations from the disaster location are also useful during disasters (P2-22; P7-22; P8-24).
- In addition to education, this application also facilities religious and spiritual studies (P1-22; P6-22; P9-22; P7-22).
- This is to support education for patients with wounds care (P2-24; P4-24; P9-24).
- Telenursing is very useful for monitoring and following up patients after being treated at the hospital (P4-22; P11-22).
- During COVID-19 virtual consultations, nurses first access their patients, then after being assessed, they are given education related to their nursing; they then consult a medical doctor. The doctor makes a diagnosis, then they are given a new prescription order to the pharmacy, and the medicine will be sent via application parties (P8-22).

XSL•FO
RenderX

The third theme was “telemedicine and telenursing have 5M factors that must be anticipated.” The weaknesses of telemedicine and telenursing were identified by participants. The main problem was the incomplete national regulations on service rules through technology, including payment. Human resource competence, connection, and infrastructure readiness were still an issue. Several participants said that it, especially telenursing, cannot be implemented entirely based on technology. Nursing care must still have contact with the patient as a manifestation of caring from nurses to patients.

The expectations of participants toward the implementation of telemedicine and telenursing were synthesized using the 5M approach, namely method, man, money, material, and machine. In the method category, management of telemedicine and telenursing is needed. This includes selecting services that can use tele/digital services, including education, monitoring, follow-up care, providing motivation, and spiritual aspects. Some services that require direct contact and diagnosis enforcement still require a dual-method model, namely direct and technology-based services. Moreover, telemedicine and telenursing services must also prioritize patient and family satisfaction and comfort. There is a need for health facility credentials to permit the delivery of a telemedicine or telenursing program. Human resources must also have the authority for technology-based services and to receive training to improve digital literacy. Money should be set aside specifically for infrastructure upgrades, for training to use the equipment, and to increase digital literacy. Participant input that represents this theme is described in Textbox 3.

Textbox 3. Participant input supporting the “telemedicine and telenursing have 5M (man, methods, money, materials, and machines) factors that must be anticipated” theme.

- I hope this telenursing will make it easier for patients to access it (P4-22; P7-22).
- The competencies that must be possessed by this telenursing; are specialist nurses allowed to do tele-nursing nursing or nurses with a minimum of S1 (Bachelor nursing) or D3 (diploma)? (P4-24; P9-22)
- Need credentials for human resources and health facilities that organize telemedicine (P10-22; P5-22)
- It is necessary to improve the human resource capabilities of both doctors and nurses; perhaps it needs to be socialized if there is an application…training cost (P5-22; P1-22; P3-22; P6-22; P1-24; P2-24; P3-24; P9-24).
- For telenursing method of delivery, not all care can be done by tele. Nurses need direct touch; caring becomes an issue when only via tele (P1-22; P12-22; P5-22, P11-22).
- What platform will be used in telemedicine or telenursing; must prepare it? Cost for infrastructure and investment (P4-24; P1-24; P2-24; P5-34; P7-24; P8-22)
- Network problems are not all good; not all areas have teledevices (P8-22; P10-22; P12-22; P11-22; P7-24).
- To data security, it must be integrated with electronic medical record (EMR; P10-22; P3-22; P1-24).
- Regulation must clear; payment must be managed by national regulation (P7-22; P11-22; P1-24; P4-24).
- We need to push for national guidelines for the implementation of telemedicine and telenursing (P10-22).
- Need professional ethical guidelines in using telenursing-based services (P6-22; P11-22)
- It’s not a clear condition. I've never seen the exact regulation, like payment regulations. But, as far as I know, already a regulation. I just haven't read it (P9-24; P10-24).
- Regulations regarding telenursing or telemedicine services and payments are not yet clear (P1-24; P4-24).
- Even though virtual is used, it must guarantee patient satisfaction and comfort (P9-22).
- Suggestions for telenursing should choose a type of service that can use digital technology. Monitoring, follow-up care, or education can be done with telenursing. This activity is still combined with direct meetings between nurses and patients (P6-22; P10-22; P8-22; P1-24; P8-24; P9-24; P11-24).
- Not all treatments can be done with telenursing such as wound care. Consultation and education can be done with telenursing or telemedicine (P1-22; P5-22; P11-22; P12-22).
- But it’s necessary to follow up on the patient’s condition, such as using video calls, Zoom to be able to carry out a more in-depth assessment (P8-22; P12-22).

Complete participant statements and keywords were grouped into themes, as described in Table 2.

Table 2 describes the keywords used in 2022 and 2024 that were compiled into themes derived from FGDs related to telenursing and telemedicine. There are differences in emphasis: in 2022, heavy focus was placed on the benefits of telenursing and telemedicine reducing the transmission of COVID-19, whereas in 2024, the benefits of telemedicine expanded to include access, education, and consultations. The 5M factors were described almost identically by the participants in 2022 and 2024. In this study, researchers also calculated keywords used to identify trends or provide support for themes and validated using incidental studies found in the literature (Table 3).

XSL•FO
RenderX

Table . Validation between the focus group discussion (FGD) themes and the literature review.

Themes and supporting FGD participant statements	Supporting participant statements in 2022 (n=23), n (%)	Supporting participant statements in 2024 (n=23), n (%)	Supporting literature
Telemedicine and telenursing have many benefits.			Song et al [2], Almahroos et al [20], Tort-Nasarre et al [1], Smith et al [3], Nejadshafiee et al [4], Hariyati et al [8], Moriyama et al [9], Ariyanto and Rosa [11], Hjelm [13], Li et al [12], Haimi [14], Costantini et al [21], Main et al [22], Zhang et al [23]
Complement digital transformation	3 (13)	3 (13)	
Speeding up and facilitating access to consultations	5 (22)	4 (17)	
Decreasing virus transmission, send drugs, consultation, and education	4 (17)	1 (4)	
Bringing patients closer to nurses, facilitating access to hospital referrals	2 (9)	4 (17)	
Decreasing indirect and direct costs	2 (9)	1 (4)	
Wound care and disaster consultation	2 (9)	3 (13)	
Facilities for religious studies	4 (17)	0 (0)	
5M[a] factors			Hjelm [13], Hariyati et al [6], Kennedy and Yaldren [24], Boro and Hariyati [7], Reed et al [25], Hwei and Octavius [15], Alexander et al [26], Moulaei et al [18], Hariyati et al [8], Wang and Cao [17], Liu et al [27]; Lee and Tak [28], Haimi [14], Tort-Nasarre et al [1], Zhang et al [23]; Mattisson et al [29], Sri Hariyati et al [30]; Vidal-Silva et al [31], Li et al [12], Hariyati et al [16]
Man			
Competencies: training, socialization	4 out of 17 (13%)	3 (13)	
Credentialing for health care providers	2 (9)	0 (0)	
Method			
Regulation and national guidance, payment guide	4 (17)	5 (22)	
Security and privacy	4 (17)	2 (9)	
Ethical guide	2 (9)	0 (0)	

XSL•FO
RenderX

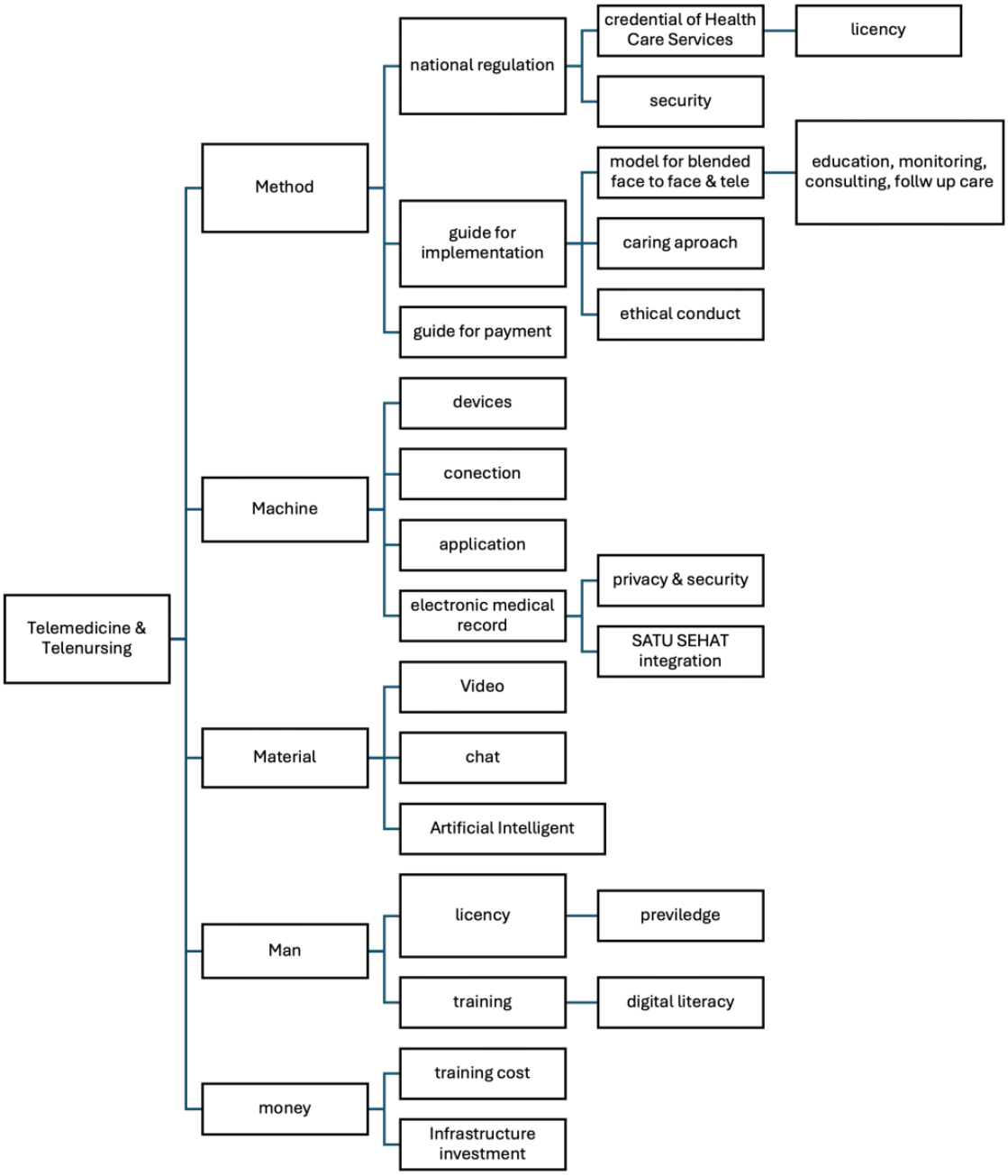
Themes and supporting FGD participant statements	Supporting participant statements in 2022 (n=23), n (%)	Supporting participant statements in 2024 (n=23), n (%)	Supporting literature
Health care services credential	2 (9)	0 (0)	
Not all can be done/combination with telenursing	4 (17)	4 (17)	
Money			
Training cost	4 (17)	4 (17)	
Infrastructure investment	1 (4)	4 (17)	
Material and machine			
Network problems, devices, application	3 (13)	3 (13)	
Data security, data confidentiality	3 (13)	3 (13)	

[a]5M: people/man, methods, money, materials, and machines.

Table 3 gives details of the analysis comparing the journal articles and the description of subthemes and themes. We also provide the percentage of participants who conveyed similar meanings of the sentences. All topics are represented in the theme table, which was built by combining the synthesis of the journal articles and participant statements. After understanding the advantages and disadvantages of telenursing, recommendations were made for the implementation of telemedicine. The recommendations were developed using the 5M approach (Figure 1).

XSL•FO
RenderX

Figure 1. Recommendation for better implementation of telemedicine and telenursing.



As shown in Figure 1, the method includes telemedicine and telenursing management, as well as the selection of tele/digital services that can be used, such as education, monitoring, follow-up care, motivation, and spiritual aspects. This represents a dual direct and indirect model using a technology-based service that still maintains direct contact for certain services such as diagnostic confirmation. Credentials for health care facilities authorized to conduct telemedicine or telenursing must be obtained from the health facility side. Patient service documentation must be arranged so that it securely uses electronic medical records (EMRs), which must be integrated with national data from SATU SEHAT Indonesia in the future. The availability of facilities including telemedicine equipment and internet connections must continue to be developed so that they can reach all Indonesian islands. Additionally, human resources must be in charge of technology-based services and undergo training to increase their digital literacy level.

Discussion

Principal Findings

This research explored the experiences of health care stakeholders using telemedicine and telenursing. We identified the following three themes: (1) health care providers have experience using telemedicine/telenursing, (2) telemedicine and telenursing have many benefits, and (3) telemedicine and telenursing have 5M factors that must be anticipated. The research results indicate that telemedicine has been implemented in Indonesia, and its use increased during the COVID-19 pandemic. Although some participants mentioned that the hospitals where they work had not yet implemented telemedicine, all participants were aware of the implementation of telemedicine. Telenursing is also one type of telemedicine

XSL•FO
RenderX

that is widely used by nurses for follow-up care, monitoring, wound care consultations, and providing motivation.

The results of the research analysis describe many benefits of telemedicine, especially considering Indonesia's archipelagic condition. Telemedicine itself is a health care service for remote examinations, diagnosis determination, and therapy administration remotely, aided by technology such as phones, video conferences, and other methods [2,7]. Differences in emphasis were observed: in 2022, the advantages of telenursing and telemedicine centered primarily on decreasing the transmission of COVID-19, whereas by 2024, the benefits of telemedicine had expanded to encompass access, education, and consultations. With the help of technology, access to services will be improved, and patients will be brought closer to health care services without having to go to hospitals or community health centers. This study supports the statement that telemedicine helps rural areas access health care services, in line with previous research [14,15,20,32].

The ability of technology to reach various rural areas means that telemedicine and telenursing can reduce both direct and indirect costs. This has also been conveyed by previous research, which states that, if consultations are conducted online, travel costs to the hospital can be saved; additionally, it will save indirect costs such as those caused by time inefficiency [14,21-23]. Another previous study also stated telestroke was a highly cost-effective strategy, which may enhance stroke care and improve favorable outcomes in patients in rural China [33].

In addition to the benefits, telemedicine also faces challenges in its implementation. The participants provided many recommendations to anticipate the 5M barriers to implementation. The obstacle of implementation failure must be addressed by creating national regulations related to telemedicine and telenursing [14]. In 2023, Indonesia started requiring the use of EMRs, which aims for data integration and interoperability toward one data called SATU SEHAT (ONE HEALTH) Indonesia [34]. Telemedicine and telenursing must also begin to be integrated into SATU SEHAT and become part of the digital transformation in the health sector. Policies ensuring the security and confidentiality of EMR data must be enforced. This is in accordance with the synthesis of studies stating that medical records must document telemedicine-based service results and that data security must be maintained [34,35].

Another recommendation regarding policy is that credentialing requirements must be established for telemedicine providers. Only institutions with permits are allowed to conduct telemedicine or telenursing. In Indonesia, a regulatory sandbox was required, and all electronic-based health care service providers should be registered and assessed for feasibility before implementing telemedicine [35]. So this policy can be implemented, the government must be encouraged to apply it according to the regulations [14,17,35]. In addition to the organizing bodies that must have permits, the human resources providing care must be credentialed and licensed to practice telemedicine or telenursing. Nurses and doctors who use technology must receive training to improve digital literacy [8,16,36]. This digital literacy is essential for optimal telenursing performance.

Barriers that need to be anticipated are also related to telemedicine facilities. Text, video, telephone, and AI-based consultation media must be developed thoroughly [21,37,38]. The use of video facilitates the interaction between nurses and patients. The use of video also allows health care workers to monitor patients' conditions, and decisions will be faster with the help of AI [5,39,40]. Unfortunately, sophisticated technology often becomes a barrier for patients who lack digital literacy. Many older adult patients and vulnerable individuals do not have the tools for telemedicine, which can hinder the development of telemedicine and telenursing [10,27,41]. In addition to the tools, network connections are not evenly distributed in every region. Therefore, previous research recommends that the government strengthen the network connections to ensure the successful implementation of telemedicine [12-14,42].

Not all care can be perfectly delivered through telemedicine. Some participants stated that they had to choose the type of service that could be provided through telenursing but still used a dual approach, namely face-to-face consultation followed by telenursing. Nurses, whose core responsibility is caring, must continue to prioritize care that consistently enhances satisfaction and comfort [8,29,43].

Patient and family preferences also significantly influence the success of telenursing. The paradox of telenursing and telemedicine, where there is no direct contact with health care personnel, means there is no physical touch, and the fear of misassessment often leads to hesitation in choosing the telemedicine method [18,26]. Therefore, it is recommended to maintain deep interactions [29] and enhance the sophistication of tools that can monitor patient conditions such as vital signs, for example, with heart rate monitors that automatically connect to servers [33,44]. Digital literacy education for the community must also be continuously strengthened so that the community increasingly understands the benefits of telenursing or telemedicine [24,28]. Improving digital literacy for the community can help them choose telemedicine and telenursing as an alternative support option for health care services [25]. Enhancing social impact is another recommendation for telemedicine success, as it can lessen apprehension about using new technology, particularly if the patient and health care practitioner have had favorable experiences with telemedicine [30]. The role of manager in encouraging the use of telemedicine is also very important, such as how managers controlled infections during COVID-19 by modifying services using technology. They must also be innovative and be role models in implementing telemedicine.

Limitations

This study used a mixed approach to analyze the qualitative descriptive results from the FGDs and journal articles. The weakness of this study is that the journal search used a convenience approach and was not part of a systematic review. The FGD participants in 2022 and 2024 were not the same people, so follow-up on the development of telenursing and telemedicine was not possible. Although it has weaknesses, the results of this study can be used as initial findings to improve the implementation of telemedicine in Indonesia and other

XSL•FO
RenderX

countries that have facilities, resources, and management that are almost the same as Indonesia. In future research, a systematic review and an in-depth interview method should be conducted.

Conclusion

These findings highlight the advantages and disadvantages of telemedicine and telenursing. We identified the following three themes from the study results: (1) health care providers have experience using telemedicine/telenursing, (2) telemedicine and telenursing have many benefits, and (3) telemedicine and telenursing have 5M factors that must be anticipated. The study's findings suggest that 5M factors should be prepared to implement telemedicine. These include procedures like payment plans, human resource credentials, and licensing rules for organizers. Enhancing digital literacy for the community and health professionals is also recommended, as well as creating a budget that focuses on infrastructure improvements and training for technology use and digital literacy. For telemedicine to function well, upgrading facilities and adding internet access are required. A service type must be selected to generate a service recommendation, which may involve a combination of telemedicine and in-person care, such as an initial in-person visit to a health care facility followed by monitoring via telenursing. Telenursing and telemedicine can help with activity monitoring, follow-up treatment, patient and family education and consultation, and incentive building. Enhancing patient satisfaction and care must be a priority when providing nursing care via telenursing. There are implications for governance; operational rules for telemedicine and telenursing must be developed, and health care providers must increase their digital literacy.

Acknowledgments

Thank you to all the participants, facilitators, and teams at the Working Group Ministry of Health, Faculty of Nursing Universitas Indonesia, and Digital Transformation Office (DTO) Ministry of Health who supported this study.

Funding

This work was supported by Grant for Policy Brief NKB-003/UN2.F12.D/HKP.01.03/2022 and a PUTI Q1-2024 NKB09/UN2.RST/HKP.05.00/ grant from Universitas Indonesia.

Conflicts of Interest

None declared.

References

1. Tort-Nasarre G, Espart A, Galbany-Estragués P, Álvarez B, Subias-Miquel M, Romeu-Labayen M. Experiences of telenursing in overcoming challenges and applaying strategies by COVID-19 patients in home isolation: qualitative study in primary care. Health Care (Don Mills) 2023 Jul 1;11(14):2093. [doi: 10.3390/healthcare11142093]
2. Song X, Liu X, Wang C. The role of telemedicine during the COVID-19 epidemic in China-experience from Shandong province. Crit Care 2020 Apr 28;24(1):178. [doi: 10.1186/s13054-020-02884-9] [Medline: 32345359]
3. Smith AC, Thomas E, Snoswell CL, et al. Telehealth for global emergencies: implications for coronavirus disease 2019 (COVID-19). J Telemed Telecare 2020 Jun;26(5):309-313. [doi: 10.1177/1357633X20916567] [Medline: 32196391]
4. Nejadshafiee M, Nekoei-Moghadam M, Bahaadinbeigy K, Khankeh H, Sheikhbardsiri H. Providing telenursing care for victims: a simulated study for introducing of possibility nursing interventions in disasters. BMC Med Inform Decis Mak 2022 Mar 2;22(1):54. [doi: 10.1186/s12911-022-01792-y] [Medline: 35236344]
5. Rintyarna BS, Insantuan OD, Widiawati I, Purwoko RY. Telehealth in remote areas: a new artificial intelligence-based model. IJSOC 2023;5(4):243-254 [FREE Full text] [doi: 10.54783/ijsoc.v5i4.782]
6. Hariyati RTS, Sahar J. Perceptions of nursing care for cardiovascular cases, knowledge on the telehealth and telecardiology in Indonesia. Int J Collab Res Intern Med Public Health 2012;4(2):116-128 [FREE Full text]
7. Boro MFV, Hariyati R. The use of telenursing through the nursing care. Jendela Nursing Journal 2019;3(2):114-121. [doi: 10.31983/jnj.v3i2.5411]
8. Hariyati RTS, Handiyani H, Zaharany TA, et al. Digital nursing promotion has increased ethics digital literacy: program improvement. F1000Res 2024;13(366). [doi: 10.12688/f1000research.139211.1]
9. Moriyama M, Kazawa K, Jahan Y, et al. The effectiveness of telenursing for self-management education on cardiometabolic conditions: a pilot project on a remote island of Ōsakikamijima, Japan. J Prim Care Community Health 2021;12:21501327211030817. [doi: 10.1177/21501327211030817] [Medline: 34238053]
10. Yuan Y, Tao C, Yu P, et al. Demand analysis of telenursing among empty-nest elderly individuals with chronic diseases based on the Kano model. Front Public Health 2022;10:990295. [doi: 10.3389/fpubh.2022.990295] [Medline: 36249233]
11. Ariyanto H, Rosa EM. Effectiveness of telenursing in improving quality of life in patients with heart failure: a systematic review and meta-analysis. J Taibah Univ Med Sci 2024 Jun;19(3):664-676. [doi: 10.1016/j.jtumed.2024.04.009] [Medline: 38807966]

XSL•FO
RenderX

12. Li X, Huang L, Zhang H, Liang Z. Enabling telemedicine from the system-level perspective: scoping review. J Med Internet Res 2025 Mar 5;27:e65932. [doi: 10.2196/65932]
13. Hjelm NM. Benefits and drawbacks of telemedicine. J Telemed Telecare 2005;11(2):60-70. [doi: 10.1258/1357633053499886] [Medline: 15829049]
14. Haimi M. The tragic paradoxical effect of telemedicine on healthcare disparities- a time for redemption: a narrative review. BMC Med Inform Decis Mak 2023 May 16;23(1):95. [doi: 10.1186/s12911-023-02194-4] [Medline: 37193960]
15. Hwei LRY, Octavius GS. Potential advantages and disadvantages of telemedicine: a literature review from the perspectives of patients, medical personnel, and hospitals. J Community Empowerment for Health 2021;4(3):228. [doi: 10.22146/jcoemph.64247]
16. Hariyati RTS, Handiyani H, Wildani AA, Afriani T, Nuraini T, Amiruddin MH. Disparate digital literacy levels of nursing manager and staff, specifically in nursing informatics competencies and their causes: a cross-sectional study. J Healthc Leadersh 2024;16:415-425. [doi: 10.2147/JHL.S470456] [Medline: 39473455]
17. Wang J, Cao Y. Factors influencing continuous intention to use telemedicine after the COVID-19 pandemic in China: an extended technology acceptance model. JSS 2022;10(12):344-359. [doi: 10.4236/jss.2022.1012023]
18. Moulaei K, Sheikhtaheri A, Fatehi F, Shanbehzadeh M, Bahaadinbeigy K. Patients' perspectives and preferences toward telemedicine versus in-person visits: a mixed-methods study on 1226 patients. BMC Med Inform Decis Mak 2023 Nov 15;23(1):261. [doi: 10.1186/s12911-023-02348-4] [Medline: 37968639]
19. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
20. Almahroos AA, Al-Harbi HYN, Jafar Alali A, et al. The impact of telemedicine on healthcare access and outcomes. JPTCP 2022;29(4):2332-2340. [doi: 10.53555/jptcp.v29i04.5361]
21. Costantini M, Sleeman KE, Peruselli C, Higginson IJ. Response and role of palliative care during the COVID-19 pandemic: a national telephone survey of hospices in Italy. Palliat Med 2020 Jul;34(7):889-895. [doi: 10.1177/0269216320920780] [Medline: 32348711]
22. Main F, Zubala A, Gorman J, et al. Technology-enabled remote management of diabetes foot disease and potential for reduction in associated health costs: a pilot study. J Foot Ankle Res 2021 Jan 13;14(1):7. [doi: 10.1186/s13047-020-00444-6] [Medline: 33436028]
23. Zhang Y, Leuk JSP, Teo WP. Domains, feasibility, effectiveness, cost, and acceptability of telehealth in aging care: scoping review of systematic reviews. JMIR Aging 2023 Apr 18;6:e40460. [doi: 10.2196/40460] [Medline: 37071459]
24. Kennedy S, Yaldren J. A look at digital literacy in health and social care. British Journal of Cardiac Nursing 2017 Sep 2;12(9):428-432. [doi: 10.12968/bjca.2017.12.9.428]
25. Reed ME, Huang J, Graetz I, et al. Patient characteristics associated with choosing a telemedicine visit vs office visit with the same primary care clinicians. JAMA Netw Open 2020 Jun 1;3(6):e205873. [doi: 10.1001/jamanetworkopen.2020.5873] [Medline: 32585018]
26. Alexander KE, Ogle T, Hoberg H, Linley L, Bradford N. Patient preferences for using technology in communication about symptoms post hospital discharge. BMC Health Serv Res 2021 Feb 15;21(1):141. [doi: 10.1186/s12913-021-06119-7] [Medline: 33588829]
27. Liu S, Zhao H, Fu J, et al. Current status and influencing factors of digital health literacy among community-dwelling older adults in Southwest China: a cross-sectional study. BMC Public Health 2022 May 17;22(1):996. [doi: 10.1186/s12889-022-13378-4] [Medline: 35581565]
28. Lee J, Tak SH. Factors associated with eHealth literacy focusing on digital literacy components: a cross-sectional study of middle-aged adults in South Korea. Digit Health 2022;8:20552076221102765. [doi: 10.1177/20552076221102765] [Medline: 35615270]
29. Mattisson M, Börjeson S, Lindberg M, Årestedt K. Psychometric evaluation of the Telenursing Interaction and Satisfaction Scale. Scand J Caring Sci 2023 Sep;37(3):687-697. [doi: 10.1111/scs.13149] [Medline: 36718016]
30. Sri Hariyati RT, Handiyani H, Erwin E, et al. Managing the nursing shortage during the COVID-19 pandemic: a qualitative study. Iran J Nurs Midwifery Res 2023;28(1):118-121. [doi: 10.4103/ijnmr.ijnmr_127_21] [Medline: 37250950]
31. Vidal-Silva C, Sánchez-Ortiz A, Serrano-Malebrán J, et al. Social influence, performance expectancy, and price value as determinants of telemedicine services acceptance in Chile. Heliyon 2024 Mar 15;10(5):e27067. [doi: 10.1016/j.heliyon.2024.e27067] [Medline: 38562504]
32. McConnell KA, Krisher LK, Lenssen M, Bunik M, Bunge Montes S, Domek GJ. Telehealth to expand community health nurse education in rural Guatemala: a pilot feasibility and acceptability evaluation. Front Public Health 2017;5:60. [doi: 10.3389/fpubh.2017.00060] [Medline: 28405582]
33. Tupsakhare P. The role of electronic medical records (EMR) in telemedicine and remote patient monitoring. IJSR 2020 Dec 5;9(12):1862-1865. [doi: 10.21275/SR24914154715]
34. Alfiyyah A, Ayuningtyas D, Rahmanto A. Telemedicine and electronic health record implementation in rural area: a literature review. J INDO HEALTH POLICY ADM 2022;7(2):221. [doi: 10.7454/ihpa.v7i2.4116]

XSL•FO
RenderX

35. Fuad A, Tiara A, Kusumasari RA, Rimawati R, Murhandarwati EEH. Introducing a regulatory sandbox into the Indonesian health system using e-Malaria as a use case: participatory action study. J Med Internet Res 2023 Dec 5;25:e47706. [doi: 10.2196/47706] [Medline: 38051555]
36. Sevean P, Dampier S, Spadoni M, Strickland S, Pilatzke S. Patients and families experiences with video telehealth in rural/remote communities in Northern Canada. J Clin Nurs 2009 Sep;18(18):2573-2579. [doi: 10.1111/j.1365-2702.2008.02427.x] [Medline: 19694885]
37. Bokolo Anthony Jnr B. Use of telemedicine and virtual care for remote treatment in response to COVID-19 pandemic. J Med Syst 2020 Jun 15;44(7):132. [doi: 10.1007/s10916-020-01596-5] [Medline: 32542571]
38. Kamei T. Telenursing and artificial intelligence for oncology nursing. Asia Pac J Oncol Nurs 2022 Dec;9(12):100119. [doi: 10.1016/j.apjon.2022.100119] [Medline: 36276880]
39. Kurniawan MH, Handiyani H, Nuraini T, Hariyati RTS, Sutrisno S. A systematic review of artificial intelligence-powered (AI-powered) chatbot intervention for managing chronic illness. Ann Med 2024 Dec;56(1):2302980. [doi: 10.1080/07853890.2024.2302980] [Medline: 38466897]
40. Weber S, Günther E, Hahnel S, Nitschke I, Rauch A. Utilization of dental services and health literacy by older seniors during the COVID-19 pandemic. BMC Geriatr 2022 Jan 31;22(1):84. [doi: 10.1186/s12877-022-02758-5] [Medline: 35100994]
41. Tan E, Gao L, Tran HN, Cadilhac D, Bladin C, Moodie M. Telestroke for acute ischaemic stroke: a systematic review of economic evaluations and a de novo cost-utility analysis for a middle income country. J Telemed Telecare 2024 Jan;30(1):18-30. [doi: 10.1177/1357633X211032407] [Medline: 34292801]
42. So KH, Ting CW, Lee CP, Lam TTN, Chiang SC, Cheung YT. Medication management service for old age homes in Hong Kong using information technology, automation technology, and the Internet of Things: pre-post Interventional study. JMIR Med Inform 2021 Feb 10;9(2):e24280. [doi: 10.2196/24280] [Medline: 33565993]
43. Marco-Franco JE, Reis-Santos M, Barrachina-Martinez I, Jurewicz A, Camaño-Puig R. Telenursing: the view of care professionals in selected EU countries. A pilot study. Heliyon 2023 Jun;9(6):e16760. [doi: 10.1016/j.heliyon.2023.e16760] [Medline: 37313150]
44. Grøndahl VA, Helgesen AK, Holm E, Magnussen J, Leonardsen AC. Remote monitoring of cancer patients during the COVID-19 pandemic - an interview study of nurses' and physicians' experiences. BMC Nurs 2022 Jun 28;21(1):169. [doi: 10.1186/s12912-022-00953-8] [Medline: 35765023]

Abbreviations

5M: man, money, method, material, and machine
EMR: electronic medical record
FGD: focus group discussion
QRS: Quality Rating System
SRQR: Standards for Reporting Qualitative Research

Edited by L Donelle; submitted 01.Jan.2026; peer-reviewed by I Rizany, M Paquet; revised version received 18.Jun.2026; accepted 20.Jul.2026; published 09.Sep.2026.

Please cite as:
Hariyati RTS, Hsu MH, Handiyani H, Amiruddin MH, Afriani T, Wardhani DA, Pranoto SD, Wildani AA, Rizafni A, Shabrina CH Recommendations From a Health Care Provider's Perspective for Better Implementation of Telemedicine and Telenursing: Descriptive Qualitative Approach
JMIR Nursing 2026;9:e90610
URL: https://nursing.jmir.org/2026/1/e90610
doi:10.2196/90610

© Rr Tutik Sri Hariyati, Min-Huei Hsu, Hanny Handiyani, Muhammad Hanif Amiruddin, Tuti Afriani, Dara Ayu Wardhani, Satria Dwi Pranoto, Andi Amalia Wildani, Arieny Rizafni, Citra Hafilah Shabrina. Originally published in JMIR Nursing (https://nursing.jmir.org), 9.Sep.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Influencing Nursing Internship Students' Readiness to Use AI: Cross-Sectional Study Using Neural Network Analysis

Sameer A Alkubati[1], PhD; Wesam T Almagharbeh[2], PhD; Talal A Alqalah[1], PhD; Basma Salameh[3], Prof Dr; Hamdan Albaqawi[1], Prof Dr; Awatif M Alrasheeday[4], PhD; Abdulhafith Alharbi[5], PhD; Layla Alshammari[1], PhD; Bushra Alshammari[1], PhD; Eddieson Pasay-an[6], PhD; Anwar Abdulkareem[7], BSc; Shimmaa Mohammed Elsayed[8], PhD

[1]Department of Medical Surgical Nursing, College of Nursing, University of Ha'il, Hail University, Hail, Ha'il Region, Saudi Arabia

[2]Medical Surgical Nursing Department, University of Tabuk, Tabuk, Tabuk Region, Saudi Arabia

[3]Department of Nursing, Arab American University, Jenin, Palestinian Territory

[4]Nursing Administration Department, Faculty of Nursing, University of Hail, Hail, Saudi Arabia

[5]Department of Psychiatric and Mental Health Nursing, College of Nursing, University of Hail, Hail, Saudi Arabia

[6]Fundamental of Nursing Department, College of Nursing, King Khalid University, Abha, Saudi Arabia

[7]Artificial Intelligence and Data Science Department, College of Computer Science and Engineering, College of Computer Science and Engineering, University of Hail, Hail, Saudi Arabia

[8]Critical Care and Emergency Nursing Department, Faculty of Nursing, Damanhour University, El Beheira, Egypt

Corresponding Author:
Sameer A Alkubati, PhD
Department of Medical Surgical Nursing, College of Nursing, University of Ha'il, Hail University, Hail, Ha'il Region, Saudi Arabia

Abstract

Background: Enhancing nursing students' awareness, attitudes, beliefs, and preparedness toward AI may help improve their health care knowledge and practice.

Objective: This study aimed to assess nursing students' attitudes, perceptions, self-efficacy, barriers, and anxiety, which influence their readiness to adopt AI in nursing practice.

Methods: This study used a cross-sectional, correlational design. Data were collected from 307 nursing internship students using an 8-part, self-administered questionnaire.

Results: Increased self-efficacy with computers was correlated with decreased barriers to accessing AI technology, lower computer anxiety scale scores (r=−0.27, P<.001 and r=−0.57, P<.001, respectively), and higher perceptions of using AI (r=0.27, P<.001). Meanwhile, nursing students' readiness to adopt AI in nursing practice was negatively associated with barriers to accessing AI technology (r=−0.20, P<.001) and positively associated with attitudes toward and perceptions of using AI (r=0.32, P<.001 and r=0.14, P=.01, respectively). Increased barriers to accessing AI technology were associated with negative attitudes toward AI and nursing students' perceptions of using AI (r=−0.34, P<.001 and r=−0.39, P<.001, respectively). A multilayer neural network model identified barriers (relative importance=0.27), attitudes (relative importance=0.16), and perceptions (relative importance=0.15) as the most significant predictors, while self-efficacy (relative importance=0.11) and anxiety (relative importance=0.07) showed smaller contributions, despite nonsignificant bivariate associations with nursing students' AI readiness. The model demonstrated strong predictive performance, achieving a low relative error of 0.62 in the training set. The stability and generalization ability of the model were supported by the training and testing set results, which yielded a training sum of squares error of 65.93 and a testing sum of squares error of 35.49, showing no signs of overfitting.

Conclusions: Several contributing factors influenced nursing students' readiness to embrace AI, with barriers, attitudes, and perceptions emerging as the most consistent, whereas self-efficacy and anxiety may play indirect roles. To improve the adoption of AI among nursing students, such factors should be dealt with in such educational programs; an interrelated adoption of AI in nursing practice is expounded as a more favorable environment.

(JMIR Nursing 2026;9:e92533) doi:10.2196/92533

KEYWORDS

nursing internship students; attitude; perception; self-efficacy; barriers; anxiety; readiness; AI

XSL•FO
RenderX

Introduction

AI refers to a broad category of technology that empowers the hardware and software of computers to emulate the intelligent behavior of humans and perform at a level comparable to that of humans [1,2]. AI significantly enhances nursing care procedures, improves patient outcomes, and increases the effectiveness of care delivery [3,4]. It can enhance clinical decision-making by offering nurses insightful data and recommendations supported by evidence to inform patient care decisions [5]. AI can help nurses diagnose and treat patients more accurately by analyzing large amounts of data [6]. It can also support predictive analytics, improve patient monitoring, and help nurses identify the early warning signs of deterioration [3]. It is becoming more important for nursing education programs to prepare student nurses with the skills they need to use AI technology and to embrace its use in nursing practice as it develops [7-9]. The health care system requires this level of training in light of technological improvements and guarantees that nurses can use AI for patient benefits [8,10]. Enhancing nursing students' awareness, attitudes, beliefs, and preparedness toward AI may help improve their care [11-13]. To determine whether nursing students are prepared to incorporate AI-based technology into their academic work, it is important to examine various factors including their attitudes, perceptions, anxiety, self-efficacy, and barriers to the value of AI. Thus, it is critical to train nursing students to efficiently use AI technology and provide innovative health care delivery. In addition, nursing students' competencies need to assess their self-efficacy in using computers and their AI literacy to optimize the use of AI sources [14].

Although there has been an increase in the involvement of AI in nursing practice and education, researchers have focused on investigating nursing students' interest and readiness to adopt AI in the future [15,16]. One study revealed that nursing students' intention to use AI in the health care system can be affected by their positive attitudes and self-efficacy [17]. According to a study by Labrague et al [16], self-rated technological ability, knowledge of AI-powered technology, and perceived usage of AI in nursing practice all influenced Philippine nursing students' modest preparedness to incorporate AI into their practice. Accessibility issues with technology may also influence intentions and future behaviors [18,19]. These obstacles may be caused by a lack of computer proficiency, lack of understanding of AI, lack of money, or time constraints [20]. This study aimed to assess nursing students' attitudes, perceptions, self-efficacy, barriers, and anxiety that influence their readiness to adopt AI in nursing practice.

Methods

Design

A cross-sectional correlation design was used in this study.

Setting

This study was conducted among nursing students at the College of Nursing, Damanhour University, from November 2024 to January 2025. This faculty serves nursing students in the Beheira Governorate, the first accredited faculty member at Damanhour University. It consists of 4 years, with 1 year of internship training in hospitals. Students experience technology when dealing with online lectures, quizzes, and assignments. Nursing students are obligated to complete and pass both required and elective courses within a specific time frame to join the internship program. Students in the nursing program spend a year working under the supervision of faculty members in a variety of hospital settings as part of the internship program's primary goal of preparing them for real-world nursing.

Study Population and Sampling

The total number of nursing students engaged in the internship program was 1400. The required sample size of 302 internship students was calculated using the OpenEpi online calculator, version 3.01 [21], according to the criteria of 5% absolute precision and a 95% confidence level. A total of 350 internship students were invited to participate, and 307 (response rate: 87.7%) completed the survey. Students completing nursing internships with a minimum of 2 months of clinical experience were eligible to participate.

Tools of the Study

Data were collected using 8 sections of a self-administered questionnaire. The first section included the participants' sociodemographic details, including age, sex, level of education, previous year's grade, duration of internet use, awareness of AI in nursing practice and health care, previous AI knowledge, and sources of information. The second section assessed the technological abilities of nursing students, as adopted from Labrague et al [16], using a single question: "In general, how do you assess your technological proficiency?" to gauge the technological abilities of nursing students. On a 5-point Likert scale, which goes from "1=basic knowledge" to "5=expert," participants scored their level of technology competency; higher scores denote greater technological skill. As this is a single-item measure, the original study established the reliability of the instrument using test-retest reliability (r=0.910) [16]. Likewise, this study demonstrated excellent test-retest reliability (r=0.902), consistent with recommendations that test-retest reliability is an appropriate method for evaluating the reliability of single-item measures [22,23].

The third section, developed by Knezek [24], assessed nursing students' self-efficacy using computers. It consists of 8 self-reported questions, with a 4-point Likert scale ranging from 0 (strongly disagree) to 3 (strongly agree), and 6 items that are reverse-scored. The total rating ranges from 0 to 24, with higher scores indicating better computer self-efficacy. Although this scale was originally developed prior to the emergence of contemporary AI technologies, it assesses general computer self-efficacy, which is considered a foundational determinant of an individual's ability to engage with and adapt to new digital tools including AI-based systems. The questionnaire was tested for reliability with a Cronbach α of 0.801 [24]. In this study, the reliability of the scale was confirmed by Cronbach α of 0.875.

The fourth part assessed nursing students' readiness to adopt AI using a 5-item scale originally developed by Ayanwale et

XSL•FO
RenderX

al [25]. In this study, the scale was adopted based on its prior modification and application among nursing students by Labrague et al [16]. To enhance its relevance to the nursing education context, Labrague et al [16] modified the original teacher-oriented items by replacing the term "teach" with "learn," "class" with "nursing practice," and "teaching in class" with "nursing curriculum." These contextual modifications ensured that the scale more accurately reflected AI readiness among nursing students, while preserving the conceptual framework of the original instrument [16]. It consists of 5 self-reported questions to assess nursing students' readiness to adopt AI in nursing practice. Each question is scored on a 5-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (5). Higher scores indicate greater willingness to use AI-based technologies. We divided this number by the number of items and added the average score for each item to determine the final score. Internal consistency was assessed using Cronbach α, which yielded a value of 0.920. The reliability of the scale was confirmed using Cronbach α, which was 0.937.

The fifth part assessed students' perceptions of the use of AI using an 11-item scale developed by Swan [26]. Participants used a 5-point Likert scale, with 1 denoting "strongly disagree" and 5 denoting "strongly agree," to score their level of agreement. The implementation and use of AI technology in nursing practice were perceived more favorably by those who scored higher on this scale. The reliability of the scale was established with a Cronbach α of 0.900 [26]. In this study, the reliability of the scale was confirmed by Cronbach α, which was 0.836.

The sixth part, adopted from Ayanwale et al [25], was used to assess nursing students' perceived barriers to using AI technology. Each item is scored on a 5-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (5). Higher ratings suggested that participants felt there were more obstacles to using the AI-powered solutions. The internal consistency of the scale, as determined by Cronbach α, was 0.880 [25]. In this study, the reliability of the scale was confirmed by Cronbach α, which was 0.894.

The seventh part assessed nursing students' attitudes toward AI, which was developed by Schepman and Rodway [27]. The questionnaire consisted of 20 self-report questions on a 5-point Likert scale, with a higher score indicating a positive attitude toward AI [27]. The reliability of the scale was established with a Cronbach α of 0.930 [28]. In this study, the reliability of the scale was confirmed by Cronbach α, which was 0.899.

The eighth part was the Short Computer Anxiety Scale, which was adopted from Lester et al [29]. A 6-item measure was developed that included questions on computer confidence. Of these, 2 items showed computing comfort, whereas the other 4 items showed shortcomings. Internal reliability was confirmed by Cronbach α of 0.760 [29]. In this study, the reliability of the scale was confirmed by Cronbach α, which was 0.844. All study instruments were administered in their original English versions, since English is the official language of instruction for internship nursing students. Therefore, no translation procedure was required.

Ethical Considerations

This study was approved by the Damanhour Research Ethics Committee (Approval 2024 - 030685). Following an explanation of the study's purpose, all the respondents provided written informed consent. The participants were made aware of their freedom to leave the study at any moment and without any explanation. Confidentiality of the respondents' data and anonymity were guaranteed. During their breaks, all intended nursing students received an invitation to participate in the study, as well as informed consent forms and questionnaires.

Data Analysis

SPSS statistics (version 26, IBM Corp) was used to evaluate the data, with a significance level of less than .05. While the mean and SD are used to summarize regularly distributed continuous data, frequencies and percentages were used to show categorical data. Histogram and statistical tests (Shapiro-Wilk and Kolmogorov-Smirnov tests) were used to determine whether the data were normal. The data were deemed to be normally distributed because the *P* values of these tests were higher than .05. One-way ANOVA (for more than 2 groups) and 2-tailed independent-sample *t* tests (for 2 groups) were used to examine the association between nursing students' attributes and overall scores. Pearson correlation coefficients were used to determine the correlations between variables. A multilayer perceptron neural network was used to predict nursing students' readiness to use AI in nursing practice. The network architecture consisted of 3 interconnected layers: an input layer, a single hidden layer, and an output layer. The input layer comprised categorical demographic factors (automatically dummy-coded into distinct operational units covering age categories, sex, study levels, internet duration, previous semester grade point average [GPA], and technological proficiency levels) and continuous variables (computer self-efficacy, perception, attitude, and computer anxiety). Of the 307 participants, the dataset was randomly partitioned into training (n=215, 70.0%) and testing (n=92, 29.9%) sets to prevent overfitting. The model was trained and its performance was evaluated using the sum of squares error (SSE) and relative error. The choice of a multilayer perceptron neural network was justified by its capacity to describe complicated nonlinear relationships and interactions between factors that typical linear regression models may not sufficiently capture. While regression and structural equation modeling are frequently used and have good interpretability, they often assume linearity and may be limited in capturing higher-order interactions between psychological and environmental variables [30].

Results

Participant Characteristics

The sample (N=307) distributions are presented in Table 1. It reveals that the mean age of undergraduate nursing students was 20.0 (SD 1.05) years, with 69.4% (n=213) being 20 years or younger. Most students were female (n=185, 60.3%). Most nursing students (n=143, 46.6%) had an excellent GPA during the semester, and 44.6% (n=137) had a moderate level of technology proficiency. While a clear majority of undergraduate nursing students (n=219, 71.3%) possessed general, baseline

XSL•FO
RenderX

knowledge about AI as a broad consumer concept, a substantially smaller proportion (n=126, 41.0%) demonstrated awareness regarding the specific, practical applications of AI within health care systems.

Table . Demographic characteristics of undergraduate nursing students (N=307).

Demographic characteristics	Values, n (%)
Age (y)	
≤20	213 (69.4)
>20	94 (30.6)
Sex	
Male	122 (39.7)
Female	185 (60.3)
Previous grade	
Poor	5 (1.6)
Fair	15 (4.9)
Good	30 (9.8)
Very good	114 (37.1)
Excellent	143 (46.6)
Technical proficiency	
Bad	5 (1.6)
Accepted	41 (13.4)
Middle	137 (44.6)
Above the average	86 (28.0)
Excellent	38 (12.4)
Awareness of AI in health care (yes/no)	
Yes	126 (41.0)
No	181 (59.0)
Awareness of AI in nursing practice (yes/no)	
Yes	125 (40.7)
No	182 (59.3)
Previous knowledge about AI (yes/no)	
Yes	219 (71.3)
No	88 (28.7)

Students aged over 20 years, studying at the fourth level, and having an excellent GPA in the last semester had a higher mean self-efficacy with computers, with statistically significant *P* values of .04, .04, and .001, respectively. Nursing students' readiness to adopt AI in nursing practice was higher among male students, those with excellent technological proficiency, those aware of AI use in health care and nursing care, and those with prior knowledge of AI, with statistically significant *P* values of .02, <.001, <.001, <.001, and <.001, respectively.

Nursing students' perception of AI had a higher mean score among those aware of AI's use in health care and nursing care, with a statistically significant *P* value of <.001. However, barriers to accessing AI technology were higher among students at the fourth level, with a statistically significant *P* value of .006. Nursing students' attitudes toward AI were more positive among those with excellent technological proficiency, those aware of AI's use in health care and nursing care, and those with prior knowledge about AI, with statistically significant *P* values of .04, .03, .008, and <.001, respectively. Finally, computer anxiety was higher among students aged 20 years or younger and those with fair GPAs in the last semester, with statistically significant *P* values of .01 and <.001, respectively. Further details are provided in Table 2.

XSL•FO
RenderX

Table . Association between demographic characteristics with self-efficacy, readiness, perception, barriers, attitude, and anxiety.

Variables	Self-efficacy	Readiness	Perception	Barriers	Attitude	Anxiety
Age, mean (SD)						
≤20 years	20 (4.0)	15 (4.6)	13.2 (3.7)	14.4 (2.9)	59.2 (8.6)	10.5 (3.4)
>20 years	21.2 (4.8)	14.1 (5.0)	13.8 (3.8)	14.8 (3.1)	59.3 (10.3)	9.4 (3.5)
P value	.04	.12	.22	.26	.93	.01
Sex, mean (SD)						
Male	20.5 (4.5)	15.5 (4.8)	13.8 (3.8)	14.7 (2.7)	60.2 (9.1)	10 (3.5)
Female	20.3 (4.1)	14.2 (4.7)	13.1 (3.7)	14.4 (3.1)	58.6 (9.2)	10.2 (3.4)
P value	.62	.02	.16	.37	.12	.61
Study level, mean (SD)						
First	19.4 (3.9)	14.8 (4.3)	12.9 (3.8)	14.2 (2.7)	59.5 (7.7)	10.7 (3)
Second	20.8 (4)	15.4 (4.6)	13.7 (3.2)	14.8 (2.9)	59.7 (9.5)	10.3 (3.9)
Third	20.4 (5.2)	13.6 (5)	13 (4.1)	13.8 (3.3)	57.4 (10.7)	9.4 (3.3)
Fourth	21.3 (4)	14.7 (5.2)	14.1 (3.9)	15.5 (3)	60.1 (8.7)	9.8 (3.3)
P value	.04	.12	.16	.006	.31	.08
Internet duration (h), mean (SD)						
<3	20 (4.6)	14.2 (5.0)	13 (3.7)	14.2 (3.5)	58.8 (10.5)	10.7 (3.6)
3 - 6	20.4 (4)	15.4 (4.4)	13.4 (3.9)	14.5 (2.6)	58.8 (7.6)	10.1 (3.4)
>6	20.7 (4.5)	14.1 (5.0)	13.6 (3.7)	15 (3)	60.2 (10.1)	9.7 (3.4)
P value	.57	.08	.60	.25	.52	.12
Last semester GPA[a], mean (SD)						
Poor	18.8 (6.7)	12 (5.1)	11 (4.9)	16.4 (6.1)	54 (11.3)	8.8 (1.9)
Fair	16.7 (3)	13.5 (3.8)	14 (3.2)	15.4 (2.4)	59.6 (11.3)	14.5 (4.3)
Good	18.9 (4)	15 (4.3)	13.2 (3.3)	13.9 (2.5)	59.9 (8.3)	10.2 (3.2)
Very good	20.4 (4.2)	14.2 (4.9)	13.7 (3.9)	14.4 (2.6)	58.8 (8.4)	10.1 (3.2)
Excellent	21.1 (4.3)	15.2 (4.7)	13.2 (3.8)	14.6 (3.3)	59.6 (9.7)	9.8 (3.4)
P value	.001	.23	.50	.31	.68	<.001
Technological proficiency, mean (SD)						
Bad	20 (4.8)	8.2 (3.4)	13 (6.0)	12.4 (2.9)	51.4 (11.8)	11 (2.5)
Accepted	20.5 (4.6)	13.8 (4.7)	14 (4.2)	14.4 (4.1)	58.3 (10.9)	9.8 (3.8)
Middle	19.9 (4.3)	13.7 (4.5)	13.6 (3.8)	14.7 (3)	58.6 (8.6)	10.6 (3.5)
Above the average	20.9 (3.7)	16.3 (4.6)	12.9 (3.2)	14.7 (2.5)	59.7 (8.9)	9.6 (3)
Excellent	20.9 (5.1)	16.5 (4.4)	13.1 (4.0)	14.1 (2.5)	62.6 (8.5)	10 (3.8)
P value	.43	<.001	.53	.36	.04	.29
Awareness AI used in health care, mean (SD)						
Yes	20.9 (4.3)	16.4 (4.8)	14 (3.90)	14.3 (2.9)	60.6 (8.2)	10 (3.3)
No	20 (4.3)	13.5 (4.3)	12.5 (3.4)	14.8 (3)	58.2 (9.7)	10.2 (3.6)
P value	.09	<.001	<.001	.16	.03	.65
Awareness AI used in nursing care, mean (SD)						
Yes	20.8 (4.2)	16.5 (4.8)	14 (3.9)	14.2 (2.9)	60.9 (8.2)	10 (3.3)
No	20.1 (4.3)	13.4 (4.3)	12.5 (3.4)	14.8 (3.0)	58.1 (9.6)	10.2 (3.5)
P value	.12	<.001	<.001	.09	.008	.65
Previous knowledge about AI, mean (SD)						

XSL•FO
RenderX

Variables	Self-efficacy	Readiness	Perception	Barriers	Attitude	Anxiety
Yes	20.5 (4.1)	15.4 (4.7)	13.2 (3.6)	14.6 (2.8)	60.6 (8.2)	10.1 (3.4)
No	20.2 (4.7)	13 (4.4)	13.7 (4.1)	14.5 (3.3)	55.8 (10.6)	10.2 (3.6)
P value	.59	<.001	.32	.95	<.001	.84

[a]GPA: grade point average.

Increased self-efficacy with computers was correlated with decreased barriers to accessing AI technology, lower computer anxiety scale scores (r=−0.27, P<.001 and r=−-0.568, P<.001, respectively), and an increase in the perception of using AI (r=0.27, P<.001). Meanwhile, nursing students' readiness to adopt AI in nursing practice was negatively associated with barriers to accessing AI technology (r=−0.20, P<.001) and positively associated with attitudes toward and perceptions of using AI (r=0.32, P<.001 and r=0.14, P=.01, respectively). An increase in barriers to accessing AI technology was associated with a negative attitude toward AI and the perception of using AI by nursing students (r=−0.34, P<.001 and r=−0.39, P<.001, respectively), as shown in Table 3.

Table . The correlation among nursing internship students' readiness, self-efficacy, barriers, attitudes, perceptions, and anxiety to use AI.

Variable and correlations	Readiness	Self-efficacy	Barriers	Attitude	Perception	Anxiety
Readiness						
r	1	0.103	−0.200	0.321	0.144	−0.031
P value	—[a]	.07	<.001	<.001	.01	.59
Self-efficacy						
r	0.103	1	−0.270	0.118	0.269	−0.568
P value	.07	—	<.001	.04	<.001	<.001
Barriers						
r	−0.200[b]	−0.270[b]	1	−0.336	−0.387	0.111
P value	<.001	<.001	—	<.001	<.001	.05
Attitude						
r	0.321	0.118	−0.336[b]	1	0.196	−0.041
P value	<.001	.04	<.001	—	<.001	.48
Perception						
r	0.144[c]	0.269[b]	−0.387[b]	0.196[b]	1	−0.087
P value	.01	<.001	<.001	<.001	—	.13
Anxiety						
r	−0.031	−0.568[b]	0.111	−0.041	−0.087	1
P value	.59	<.001	.05	.48	.13	—

[a]Not applicable.

[b]Correlation is significant at the .01 level (2-tailed).

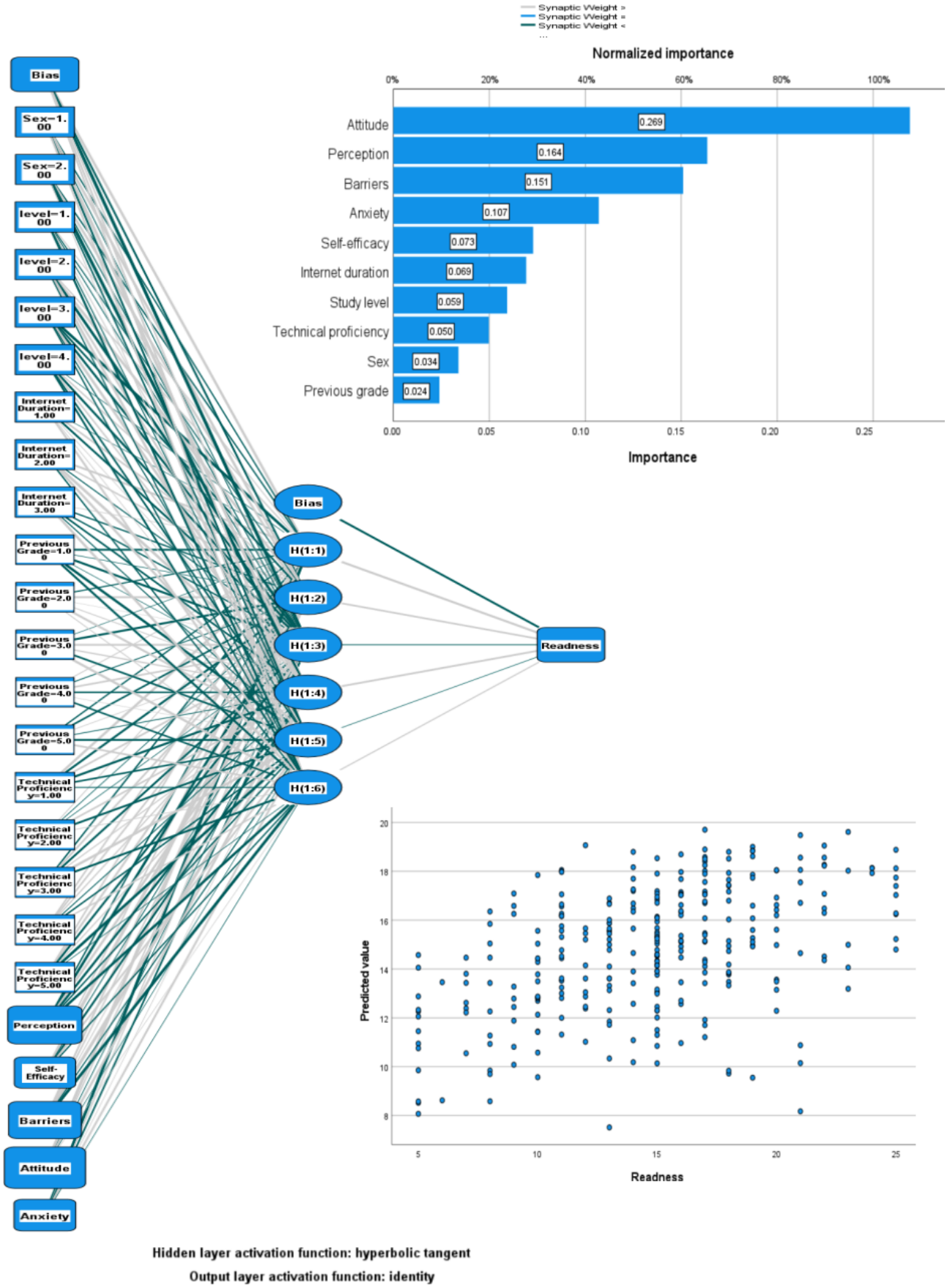
[c]Correlation is significant at the .05 level (2-tailed).

Predictors of Nursing Internship Students' Readiness to Use AI

A multilayer neural network was used to predict nursing internship students' readiness to adopt AI, based on specified independent factors. The model identified barriers (relative importance=0.27), attitudes (relative importance=0.16), and perceptions (relative importance=0.15) as the strongest predictors of readiness. Self-efficacy (relative importance=0.11) and anxiety (relative importance=0.07) demonstrated comparatively smaller contributions despite their nonsignificant bivariate correlations with readiness. The findings shown in Figure 1 suggest that these factors collectively influenced nurses' openness to AI integration. The model demonstrated strong predictive validity, minimized error profiles, and high stability. While the training SSE was 65.93, the model achieved a low training relative error of 0.62, indicating high explanatory adequacy. The generalization and predictive ability of the network were confirmed by a testing SSE of 35.49 and a low testing relative error, establishing that the model was structurally stable and free from overfitting constraints.

XSL•FO
RenderX

Figure 1. Factors affecting nursing internship students' readiness to use AI.



XSL•FO
RenderX

Discussion

Principal Findings

This study aimed to assess how nursing students' attitudes, perceptions, self-efficacy, barriers, and anxiety influenced their readiness to adopt AI in nursing practice. The readiness of nursing students to incorporate AI in their practice is influenced by several factors, such as gender, technology skills, knowledge of the applicability of AI, and most importantly, knowledge background. Comprehensively addressing these factors through oriented teaching may equip nursing students with better ways to incorporate AI into their future practice and consequently enhance patient outcomes. Students aged over 20 years, studying at the fourth level, and having excellent GPAs in the last semester had higher mean self-efficacy with computers.

The readiness of nursing students to adopt AI in nursing practice was higher among male students, those with excellent technological proficiency, those aware of the use of AI in health care and nursing care, and those with prior knowledge of AI. Additionally, students with significant technological abilities tend to be more self-assured in their capacity to study and apply technologies such as AI, which has been clearly linked to their preparedness to embrace AI in nursing practice [31]. According to previous studies, male students are more likely to be ready to operate with AI than female students [31,32]. This is likely due to sociocultural factors that encourage males to take an interest in technological aspects and be less risk averse [32].

Nursing students' perceptions of AI had a higher mean score among those who were aware of the use of AI in health care and nursing care. However, barriers to accessing AI technology were higher among students in the fourth level. This suggests that nursing students who are familiar with the application of AI in the health sector may tend to have favorable sentiments toward it. However, especially during the graduate year, fourth-year students may not be able to access and use AI technology for reasons such as graduation pressure, insufficient time, and resources. Students exposed to further uses of AI in a clinical environment develop a more favorable attitude toward AI [33]. Nevertheless, as many positive perceptions arise as a result of awareness, obstacles to the appreciation of AI technology are a real source of concern, mostly for fourth-year undergraduate nurses. Fourth-year students may experience increased academic and career-related pressures, which could contribute to perceived barriers to engaging in AI technologies. This is in line with those who argue that, due to the underdevelopment of certain components of nursing education programs, such as health informatics and AI incorporating education, feelings of inferiority may arise and inhibit students' ability to use these technologies [34]. This educational gap may contribute to the barriers experienced by these students, as they may not feel adequately prepared to engage with AI technologies in clinical settings.

Nursing students' attitudes toward AI were more positive among those with excellent technological proficiency, those aware of AI's use in health care and nursing care, and those with prior knowledge of AI. Additionally, computer anxiety was higher among students aged 20 years or younger and among those with a fair GPA in the previous semester. This finding indicates that younger nursing students may experience greater apprehension when using computer technologies, highlighting the importance of early exposure to digital technologies and targeted educational support to enhance confidence and reduce technology-related anxiety. Owing to the increasing integration of AI worldwide, there is a dire need for nursing education that integrates AI knowledge to enhance the willingness of technology-based students [34]. It is also pertinent to note that awareness of the application of these technologies among nursing students also has an impact on their attitudes toward nursing. Previous studies have noted that students and professionals perceive AI to be more positivistic and feasible if they are aware of implications such as improved patient outcomes and decision-making [11,17]. Considering the prospects of AI in health care, it is essential to focus on educational activities that will allow future nurses to fully apply this knowledge in practice.

Meanwhile, nursing students' readiness to adopt AI in nursing practice was associated with fewer barriers to accessing AI technology, a more positive attitude toward AI, and a more positive perception of AI use. Greater familiarity with AI implementation was associated with more favorable attitudes toward AI technology, suggesting that increased exposure to and experience with AI may enhance students' acceptance of and confidence in using these technologies. These students were confirmed to be more positive about AI technologies and wanted to use them in their work [17,35]. From this, one can deduce that attention directed toward the progression of students' understanding of AI leads to an increased possibility of supporting positive change.

In contrast, high-level barriers to entering AI technology correlate with negative attitudes toward and perceptions of AI. For instance, nursing students who do not have easy access to AI tools may develop negative feelings, such as anxiety and negligence toward these technologies, which could discourage them from using them. This was consistent with a study conducted by Chen et al [36] in which perceived barriers were shown to reduce students' intent to engage in AI-related activities. In addition, this study highlights the need to overcome these barriers because students' attitudes toward AI are important for the acceptance and subsequent use of AI-powered applications in nursing [37]. This study used neural network modeling to ascertain the participants' AI adoption styles. There was a higher degree of accuracy in predicting readiness, with barriers, attitudes, perceptions, self-efficacy, and anxiety emerging as the top 5 manifest variables. In the case of readiness, there were also variables such as lack of access to AI technology, attitudes toward AI technology, beliefs about the usefulness of the technology, feelings of self-efficacy, and feelings of anxiety about the new technology. AI technology is positively and directly associated with perceptions of usefulness [20]. For example, it has been reported that nursing students' perceptions of AI technologies predict more positive attitudes toward the use of AI in nursing practice [17]. Indeed, training and support remove barriers and improve self-efficacy, thereby increasing the propensity to use AI [38]. Obenza et al [39] supported the notion that self-efficacy is an important determinant of individuals' willingness and confidence in

XSL•FO
RenderX

adopting emerging technologies. There are various estimations among nursing students regarding the implementation of AI when providing nursing care, and these vary in the adoption of AI owing to their attitudes and concerning factors. For instance, high perceived barriers to using AI technology tend to correlate negatively with positive feelings toward AI and the perception of its usefulness. In cases where nursing students possess barriers such as poor orientation and deficit of resources or scanty institutional support, they are likely to develop a certain disbelief in the usage of these AI-based tools. In such cases, anxiety about these problems is more likely to arise, which, according to [17], is negatively related to self-efficacy and willingness to accept changes in practices, including new technologies [17] in a negative sense. However, because of this anxiety, there is a high chance that students will be reluctant to use AI, which will lead to a negative attitude toward the adoption of new technology, which is likely to result in a negative loop, as pointed out [40]. Therefore, it is essential to address students' anxiety regarding AI to eliminate this negative loop and establish better interactions between nursing students and their modern technologies.

Despite the contribution of self-efficacy and anxiety to the neural network model, they did not correlate substantially with readiness in bivariate studies. This disparity may be explained by the capacity of neural networks to identify interaction effects and nonlinear interactions that are missed by straightforward correlations. Therefore, rather than being direct indicators of preparedness, their involvement should be carefully understood as indirect or context dependent. Self-efficacy is an important determinant of nursing students' perceptions of AI. A recent study indicated that students with greater self-efficacy regarding AI tend to approach and use AI technologies more positively [39]. It has been noted that self-efficacy is an important variable that affects the use of any technology, suggesting that students with such self-efficacy will be able to break through barriers and embrace new technology optimization [41]. Therefore, if the root causes of low self-efficacy concerning AI use in nursing practice are addressed with appropriate educational measures, there is a high likelihood that nurses will accept and use this technology while performing their duties. When self-efficacy is improved through adequate training programs and AI barriers are uplifted, nursing teachers can positively alter students' attitudes toward using AI. This approach also resonates with an earlier research suggestion that health care professionals should work together to facilitate effective integration of AI technologies into nursing practice [42]. Furthermore, the literature highlights the need to consider students' perceptions and attitudes toward AI. Anxiety factors, such as fear of incomprehension and subordination, are negatively influenced by the intention to use AI in nursing students [17]. Filling students' minds with information on the availability and advantages of AI technologies can increase their acceptance of and intention to use such tools. This is especially pertinent for nursing education, where incorporating AI into curricula could help students better comprehend it and decrease the fear associated with its application. However, the use of self-reported measures increases the risk of social desirability effects and common method bias, which could exaggerate the observed correlations between variables. The correlations may have been overstated because multiple categories (such as attitudes, self-efficacy, and perception) were measured at a single time point using the same methodology. The prediction performance of a neural network model may also be affected by inflation. Therefore, the strength of this relationship should be cautiously interpreted.

Limitations

Despite revealing some important information, the limitations of this study include country restrictions, use of self-reporting methods, use of only survey instruments, a lack of emphasis on ethical and cultural dimensions, the absence of tools within the study that might be needed, absence of data regarding AI use by students, and consideration of only short-term issues. To mitigate or even eliminate such limitations, it is recommended that future studies cover more, use what is available, relate the variables, detail ethics and cultures, check the clients' tools, look for students' embedment in AI, and assess their anxiety at a distance. This helps to expand the knowledge of how nursing students perceive AI in their future practice and helps form clinically effective approaches to the integration of AI into nursing practice. Although the AI readiness scale was previously updated for nursing students, it was originally designed for different professional groups. In addition, the instruments used in this study were originally developed and validated in different cultural contexts. Although administered in English, the absence of formal cross-cultural validation in the study population may have affected the generalizability and cultural equivalence of the findings. Additionally, the use of a neural network model, which is advantageous for prediction, limits interpretability compared to traditional statistical methods, which may affect the explanatory depth of the findings. Moreover, this study relied solely on self-reported data, which are susceptible to frequent method bias and social desirability effects. These biases may result in inflated effect sizes and overestimation of correlations between variables, as well as potentially improve model prediction accuracy. To avoid these biases, future research should use objective metrics, multiple data sources, or longitudinal approaches.

Implications for Nursing

Nursing students' competence in using their technological knowledge and skills to adopt and implement AI in their careers is explained by their positive attitudes toward and understanding of the benefits of AI. However, factors such as anxiety and limited access can reduce competence. To solve these problems, it is necessary to embed AI in nursing education, provide practical experience, eliminate fear, change public perception, and highlight its benefits. Furthermore, AI education should address ethical and cultural issues explicitly. To guarantee ethical use, issues such as data privacy, algorithmic bias, patient safety, and cultural sensitivity in AI-supported decision-making are crucial. Educational practices should be context-sensitive and prioritize fairness and accessibility due to differences in access to technology and institutional resources. Such integration will prepare future nurses to accept AI in their practice to advance the level of care patients receive in the health care system.

XSL•FO
RenderX

Conclusions

Several contributing factors influence nursing students' readiness to embrace AI, with barriers, attitudes, and perceptions emerging as the most consistent factors, while self-efficacy and anxiety may play indirect roles. The receptive attitudes of students with more skills, awareness, computer proficiency, and positive beliefs have encouraged the use of AI. Conversely, barriers, particularly higher levels of anxiety among younger students (≤20 y) and those with a lower GPA, were associated with lower readiness to use AI. A neural network model affirms the effects of various factors in predicting readiness to embrace AI. Creating a supportive educational environment that fosters AI readiness, self-efficacy, and positive attitudes toward AI may facilitate appropriate integration of AI into nursing education and future clinical practice.

Acknowledgments

We would like to thank all nursing internship students who participated in this study. The authors confirm that no generative AI tools were used in the preparation of the article.

Funding

The authors declared no financial support was received for this work.

Data Availability

The data are available upon request from the corresponding author.

Authors' Contributions

Conceptualization: SA, AMA, A Alharbi, LA, BA, SME, A Abdulkareem
Data curation: SA, TAA, AMA, LA, BA, A Abdulkareem, SME
Formal analysis: SA, A Abdulkareem
Investigation: SA, TAA, LA, BA, A Abdulkareem, SME
Methodology: SA, A Alharbi, LA, BA, A Abdulkareem, SME
Resources: SA
Validation: SA
Visualization: SA
Writing – original draft: SA, WTA, BS, HA, EP-A, SME
Writing – review & editing: SA, WTA, TAA, BS, HA, AMA, A Alharbi, EP-A

Conflicts of Interest

None declared.

References

1. Collins C, Dennehy D, Conboy K, Mikalef P. Artificial intelligence in information systems research: a systematic literature review and research agenda. Int J Inf Manage 2021 Oct;60:102383. [doi: 10.1016/j.ijinfomgt.2021.102383]
2. Al-Qerem W, Eberhardt J, Jarab A, et al. Exploring knowledge, attitudes, and practices towards artificial intelligence among health professions' students in Jordan. BMC Med Inform Decis Mak 2023 Dec 14;23(1):288. [doi: 10.1186/s12911-023-02403-0] [Medline: 38098095]
3. Rony MKK, Parvin MR, Ferdousi S. Advancing nursing practice with artificial intelligence: enhancing preparedness for the future. Nurs Open 2024 Jan;11(1). [doi: 10.1002/nop2.2070] [Medline: 38268252]
4. El-Gazar HE, Zoromba MA, Fayed SM, et al. Nurturing success: e-learning readiness and academic self-efficacy in nursing students. BMC Nurs 2024 Jul 19;23(1):495. [doi: 10.1186/s12912-024-02125-2] [Medline: 39030566]
5. Van Bulck L, Couturier R, Moons P. Applications of artificial intelligence for nursing: has a new era arrived? Eur J Cardiovasc Nurs 2023 Apr 12;22(3):e19-e20. [doi: 10.1093/eurjcn/zvac097] [Medline: 36269804]
6. Ahmed SK. Artificial intelligence in nursing: current trends, possibilities and pitfalls. J Med Surg Public Health 2024 Aug;3:100072. [doi: 10.1016/j.glmedi.2024.100072]
7. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on nursing education: scoping review. JMIR Nurs 2021;4(1):e23933. [doi: 10.2196/23933] [Medline: 34345794]
8. Labrague LJ, Aguilar-Rosales R, Yboa BC, Sabio JB, de Los Santos JA. Student nurses' attitudes, perceived utilization, and intention to adopt artificial intelligence (AI) technology in nursing practice: a cross-sectional study. Nurse Educ Pract 2023 Nov;73:103815. [doi: 10.1016/j.nepr.2023.103815] [Medline: 37922736]
9. El-Ashry AM, Al Saleh NS, AlOtaibi NG, et al. The impact of artificial intelligence attitudes and acceptance on critical thinking motivation among nursing students in Saudi Arabia. SAGE Open Nurs 2025;11:23779608251369564. [doi: 10.1177/23779608251369564] [Medline: 40809301]

XSL•FO
RenderX

10. El-Ashry AM, Hermis AH, Al-Salih SSH, et al. Attitudes toward artificial intelligence and impostor phenomenon among nursing students: a five-country cross-sectional study. Nurse Educ Today 2026 Jun;161:107011. [doi: 10.1016/j.nedt.2026.107011] [Medline: 41653898]
11. Abuzaid MM, Elshami W, Fadden SM. Integration of artificial intelligence into nursing practice. Health Technol (Berl) 2022;12(6):1109-1115. [doi: 10.1007/s12553-022-00697-0] [Medline: 36117522]
12. Alnawafleh KA, Almagharbeh WT, Alfanash HA, et al. Exploring the ethical dimensions of AI integration in nursing practice: a systematic review. J Nurs Regul 2025 Oct;16(3):228-237. [doi: 10.1016/j.jnr.2025.08.001]
13. Alrasheeday A, ALhawsawi AA, Alshammari B, et al. Knowledge, attitude, benefits, risks, barriers, professional impact, and preparedness of nursing students toward the utilization of artificial intelligence in healthcare. Nurs Rep 2026 May 4;16(5):154. [doi: 10.3390/nursrep16050154] [Medline: 42188647]
14. Abou Hashish EA, Alnajjar H. Digital proficiency: assessing knowledge, attitudes, and skills in digital transformation, health literacy, and artificial intelligence among university nursing students. BMC Med Educ 2024 May 7;24(1):508. [doi: 10.1186/s12909-024-05482-3] [Medline: 38715005]
15. Glauberman G, Ito-Fujita A, Katz S, Callahan J. Artificial intelligence in nursing education: opportunities and challenges. Hawaii J Health Soc Welf 2023 Dec;82(12):302-305. [doi: 10.62547/YEIS2105] [Medline: 38093763]
16. Labrague LJ, Aguilar-Rosales R, Yboa BC, Sabio JB. Factors influencing student nurses' readiness to adopt artificial intelligence (AI) in their studies and their perceived barriers to accessing AI technology: a cross-sectional study. Nurse Educ Today 2023 Nov;130:105945. [doi: 10.1016/j.nedt.2023.105945] [Medline: 37625351]
17. Kwak Y, Seo YH, Ahn JW. Nursing students' intent to use AI-based healthcare technology: path analysis using the unified theory of acceptance and use of technology. Nurse Educ Today 2022 Dec;119:105541. [doi: 10.1016/j.nedt.2022.105541] [Medline: 36116387]
18. Alrasheeday AM, Alshammari B, Alkubati SA, Pasay-An E, Albloushi M, Alshammari AM. Nurses' attitudes and factors affecting use of electronic health record in Saudi Arabia. Healthcare (Basel) 2023 Aug 25;11(17):2393. [doi: 10.3390/healthcare11172393] [Medline: 37685427]
19. Ali AS, Alhirsan SM, Elshazly M, et al. Artificial intelligence in physiotherapy education: a multi-country Middle East survey of AI acceptance and barriers among university students. Euro J of Education 2026 Mar;61(1):e70476. [doi: 10.1111/ejed.70476]
20. Kelly S, Kaye SA, Oviedo-Trespalacios O. What factors contribute to the acceptance of artificial intelligence? A systematic review. Telemat Inform 2023 Feb;77:101925. [doi: 10.1016/j.tele.2022.101925]
21. OpenEpi. URL: https://www.openepi.com/Menu/OE_Menu.htm [accessed 2026-07-31]
22. Shen C, Wang MP, Ho HCY, et al. Test-retest reliability and validity of a single-item self-reported Family Happiness Scale in Hong Kong Chinese: findings from Hong Kong Jockey Club FAMILY Project. Qual Life Res 2019 Feb;28(2):535-543. [doi: 10.1007/s11136-018-2019-9] [Medline: 30284182]
23. Verster JC, Mulder KE, Verheul MC, et al. Test-retest reliability of single-item assessments of immune fitness, mood, and quality of life. Heliyon 2023 Apr;9(4):e15280. [doi: 10.1016/j.heliyon.2023.e15280] [Medline: 37089305]
24. Knezek G. Internal consistency reliabilities for 14 computer attitude scales. In: Willis J, Price J, McNeil S, Robin B, Willis D, editors. Society for Information Technology & Teacher Education International Conference: Association for the Advancement of Computing in Education (AACE); 1997:877-880 URL: https://www.learntechlib.org/primary/p/47210/ [accessed 2026-07-31]
25. Ayanwale MA, Sanusi IT, Adelana OP, Aruleba KD, Oyelere SS. Teachers' readiness and intention to teach artificial intelligence in schools. Comput Educ Artif Intell 2022;3(1):100099. [doi: 10.1016/j.caeai.2022.100099]
26. Swan BA. Assessing the knowledge and attitudes of registered nurses about artificial intelligence in nursing and health care. Nurs Econ 2021;39(3):139. [doi: 10.62116/NEC.2021.39.3.139]
27. Schepman A, Rodway P. Initial validation of the general attitudes towards Artificial Intelligence Scale. Comput Hum Behav Rep 2020;1:100014. [doi: 10.1016/j.chbr.2020.100014] [Medline: 34235291]
28. Cicero L, Russo A, Di Stefano G, Zammitti A. The General Attitudes towards Artificial Intelligence Scale (GAAIS): validation and psychometric properties analysis in the Italian context. BMC Psychol 2025 Jul 1;13(1):641. [doi: 10.1186/s40359-025-02935-2] [Medline: 40597372]
29. Lester D, Yang B, James S. A short computer anxiety scale. Percept Mot Skills 2005 Jun;100(3 Pt 2):964-968. [doi: 10.2466/pms.100.3c.964-968] [Medline: 16158683]
30. Ostovar A, Davari DD, Dzikuć M. Determinants of design with multilayer perceptron neural networks: a comparison with logistic regression. Sustainability 2025;17(6):2611. [doi: 10.3390/su17062611]
31. Dai Y, Chai CS, Lin PY, Jong MSY, Guo Y, Qin J. Promoting students' well-being by developing their readiness for the artificial intelligence age. Sustainability 2020;12(16):6597. [doi: 10.3390/su12166597]
32. Tubadji A, Denney T, Webber DJ. Cultural relativity in consumers' rates of adoption of artificial intelligence. Econ Inq 2021 Jul;59(3):1234-1251. [doi: 10.1111/ecin.12978]
33. O'Connor S, Yan Y, Thilo FJS, Felzmann H, Dowding D, Lee JJ. Artificial intelligence in nursing and midwifery: a systematic review. J Clin Nurs 2023 Jul;32(13-14):2951-2968. [doi: 10.1111/jocn.16478] [Medline: 35908207]

XSL•FO
RenderX

34. Ronquillo CE, Peltonen LM, Pruinelli L, et al. Artificial intelligence in nursing: priorities and opportunities from an international invitational think-tank of the Nursing and Artificial Intelligence Leadership Collaborative. J Adv Nurs 2021 Sep;77(9):3707-3717. [doi: 10.1111/jan.14855] [Medline: 34003504]
35. Alrashedi H, Alderaan SM, Alnomasy N, Lamine H, Saleh KA, Alkubati SA. Insights into factors affecting nurses' knowledge of and attitudes toward AI and implications for successful AI integration in critical care: cross-sectional study. JMIR Nurs 2026 Jan 16;9:e85649. [doi: 10.2196/85649] [Medline: 41544103]
36. Chen SY, Su YS, Ku YY, Lai CF, Hsiao KL. Exploring the factors of students' intention to participate in AI software development. Library Hi Tec 2024 May 24;42(2):392-408. [doi: 10.1108/LHT-12-2021-0480]
37. Hassan Mekawy S, Ali Mohamed Ismail S, Zayed Mohamed M. Digital health literacy (DHL) levels among nursing baccalaureate students and their perception and attitudes toward the application of artificial intelligence (AI) in nursing. Egypt J Health Care 2020 Mar 1;11(1):1266-1277. [doi: 10.21608/ejhc.2020.274757]
38. Jia XH, Tu JC. Towards a new conceptual model of AI-enhanced learning for college students: the roles of artificial intelligence capabilities, general self-efficacy, learning motivation, and critical thinking awareness. Systems 2024;12(3):74. [doi: 10.3390/systems12030074]
39. Obenza BN, Baguio J, Bardago KMW, et al. The mediating effect of AI trust on AI self-efficacy and attitude toward AI of college students. Int J Metaverse 2024;2(1):1-10. [doi: 10.54536/ijm.v2i1.2286]
40. Hu D, Kong Y, Li W, et al. Frontline nurses' burnout, anxiety, depression, and fear statuses and their associated factors during the COVID-19 outbreak in Wuhan, China: a large-scale cross-sectional study. EClinicalMedicine 2020 Jul;24:100424. [doi: 10.1016/j.eclinm.2020.100424] [Medline: 32766539]
41. Pan X. Technology acceptance, technological self-efficacy, and attitude toward technology-based self-directed learning: learning motivation as a mediator. Front Psychol 2020;11:564294. [doi: 10.3389/fpsyg.2020.564294] [Medline: 33192838]
42. Laï MC, Brian M, Mamzer MF. Perceptions of artificial intelligence in healthcare: findings from a qualitative survey study among actors in France. J Transl Med 2020 Jan 9;18(1):14. [doi: 10.1186/s12967-019-02204-y] [Medline: 31918710]

Abbreviations

GPA: grade point average
SSE: sum of squares error

Edited by E Borycki; submitted 30.Jan.2026; peer-reviewed by JU Almazan, P Saaristo; revised version received 20.Jun.2026; accepted 22.Jul.2026; published 18.Aug.2026.

Please cite as:
Alkubati SA, T Almagharbeh W, A Alqalah T, Salameh B, Albaqawi H, M Alrasheeday A, Alharbi A, Alshammari L, Alshammari B, Pasay-an E, Abdulkareem A, Mohammed Elsayed S
Factors Influencing Nursing Internship Students' Readiness to Use AI: Cross-Sectional Study Using Neural Network Analysis
JMIR Nursing 2026;9:e92533
URL: https://nursing.jmir.org/2026/1/e92533
doi:10.2196/92533

© Sameer A Alkubati, Wesam T Almagharbeh, Talal A Alqalah, Basma Salameh, Hamdan Albaqawi, Awatif M Alrasheeday, Abdulhafith Alharbi, Layla Alshammari, Bushra Alshammari, Eddieson Pasay-an, Anwar Abdulkareem, Shimmaa Mohammed Elsayed. Originally published in JMIR Nursing (https://nursing.jmir.org), 18.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Process Evaluation of a Digital Smartphone Intervention (URApp) to Support Adherence to Bladder Training in Community-Based Specialist Pediatric Continence Care in England: Mixed Methods Study

Lucy Beasant[1], PhD; Jenny Ingram[1], PhD; Davina Richardson[2], BSc; Joanna Clothier[3], BChir, MA, MB, MRCPCH; Carol Joinson[1], PhD

[1]Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, England, United Kingdom

[2]Bladder & Bowel UK, 10 Priestley Road, Burrows House, Manchester, England, United Kingdom

[3]Department of Paediatric Nephrology, Evelina London Children's Hospital, SE1 7EH, Westminster Bridge Road, London, England, United Kingdom

Corresponding Author:
Lucy Beasant, PhD
Bristol Medical School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol, England, United Kingdom

Abstract

Background: Daytime urinary incontinence is a common childhood condition that can persist into adolescence, and it is associated with psychological distress, social and educational difficulties, and an increased risk of later mental disorders. The co-designed smartphone app "URApp" was developed in partnership with young people aged 10 to 18 years to support their self-management of bladder symptoms, but it is not yet used by community pediatric continence clinics to facilitate ongoing patient bladder training.

Objective: The study aims to understand the barriers and facilitators to embedding a digital health intervention (URApp) designed to aid adherence to bladder training in children and young people with daytime urinary incontinence in community pediatric continence clinics.

Methods: We conducted a mixed methods process evaluation using the Normalization Measure Development (NoMAD) questionnaire at 4 time points to measure behavior change in nursing teams at 6, 12, and 18 weeks. Teams' responses to questionnaire items were summarized as median scores and plotted at each time point using petal plots. Qualitative data were collected through focus groups and interviews and analyzed, drawing upon thematic analysis.

Results: Twenty pediatric continence nurses from 5 National Health Service community pediatric continence clinics in England were recruited through purposive maximum variation sampling. URApp was well-received, and nurses were positive about introducing it to patients and felt that URApp's functionality aligned with existing bladder training provided by National Health Service community clinics. Nurses reported that URApp was easy to introduce, age-appropriate for children aged 10 to 18 years, and accessible to patients with neurodivergent conditions. However, community teams found integration into existing care pathways challenging and did not fully use URApp for urotherapy and bladder diary data collection. Teams lacked confidence in changing from paper to electronic bladder diaries, and although they were positive about young people taking ownership of their bladder condition, they struggled to encourage URApp uptake and were concerned about the lack of parental oversight of bladder diary data in URApp. Successful uptake and patient engagement with URApp were also negatively impacted by barriers in the school environment, such as lack of access to toilets and smartphones.

Conclusions: The successful integration of URApp into community pediatric continence clinics requires cultural change in schools regarding toilet access and healthy drinking. Future implementation of URApp into existing clinical care pathways will need to consider how electronic bladder diary data from URApp can be used more effectively by clinical teams, in partnership with patients and their parents/carers to facilitate the shift from analog to a "digital-first health service."

(JMIR Nursing 2026;9:e95418) doi:10.2196/95418

KEYWORDS

adolescents; digital health interventions; behavior change; daytime urinary incontinence; community pediatric continence care; mixed methods; process evaluation; schools; smartphone; standard urotherapy; URApp; urinary bladder

XSL•FO
RenderX

Introduction

Daytime urinary incontinence (UI), the involuntary leakage of urine during the day, is a common child health problem that can persist into adolescence. International prevalence estimates of UI vary across studies due to differing definitions, but approximately 5% to 10% of school-aged children experience daytime UI, with prevalence declining to around 1% to 3% during adolescence [1]. Similar patterns have been reported in UK cohort studies, demonstrating that continence problems persist into adolescence for a clinically important minority of young people [2]. There is robust evidence that the prevalence of daytime UI is higher among girls (4.2% of females and 1.3% of males) at the age of 14 years [3-5]. Daytime UI in children and young people (hereafter referred to as young people) leads to psychological distress, social isolation, adverse impacts on education, and negative self-beliefs such as "feeling like an outsider" [6]; young people with UI are at increased risk of mental disorders in early adulthood [7,8]. Most cases of UI in children and adolescents are not due to underlying neurological or anatomical abnormalities, and the causes are unknown [7]. Despite the prevalence and adverse impacts of UI, most young people feel that they are not adequately supported by current services [9-11].

There are huge variations across the United Kingdom in pediatric continence service provision, and demand on these already overstretched services has increased since the COVID-19 pandemic [12]. Bladder training (standard urotherapy) is a behavior modification intervention that aims to correct bladder filling and voiding function [13]. It is an established conservative first-line intervention for UI that is recommended by the National Health Service (NHS) [14], charities supporting people affected by incontinence (ERIC [The Children's Bladder and Bowel Charity] and Bladder and Bowel, United Kingdom), and the International Children's Continence Society [15,16]. Bladder training involves establishing a regular schedule of fluid intake and timed toileting to improve the bladder function and reduce UI [13]. Bladder training can be effective, but success strongly depends on self-motivation, commitment, and ongoing clinical support [9,17]. Young people with UI often feel pessimistic about their prognosis and find it challenging to adhere to bladder training [9]. Treatment failure has cost implications for the NHS and threatens the mental well-being and academic attainment of young people [6,7,9,18,19].

Our research with young people with UI found that they have a strong desire to be involved in decisions about their treatment and to feel supported in self-managing their bladder symptoms [9]. In response to this, we co-designed the smartphone app "URApp" in partnership with young people aged 10 to 18 years [20]. URApp's core functions are based on existing bladder training approaches [13,21] and on empirical evidence about behavior change techniques to improve adherence [22-25]. These include techniques such as goal setting and reviewing, action planning, personalized feedback, and rewards to aid motivation to adhere to bladder training and to help young people work toward a personalized drinking goal. URApp provides discreet reminders for toilet visits and drinking; allows users to record fluid intake, frequency of toilet visits, and urine volume output; and provides charts to track progress, and it includes a diary to record other useful information such as wetting accidents, mood, types of drinks, or stressful events.

In August 2024, URApp was updated to allow users to export a bladder diary to share with their clinic team by text or email; click on *View the App Guide* [20]. Young people involved in the development and formative evaluation of URApp reported that it was acceptable, usable, engaging, and effective in supporting their drinking and toilet schedules, with many requesting to continue using the app beyond the study [26]. Health care professionals working in secondary care settings also reported that URApp was clinically appropriate and aligned with the guidelines for current bladder training and would use it within their own clinics [26].

Despite URApp already being recommended and used by some health care professionals in pediatric continence care [27], it had not yet been evaluated in community pediatric continence care settings. Digital health interventions, including smartphone apps, could facilitate the shifting of care from hospital to community and from analog/paper-based processes by providing continuous digital support, empowering patients to be active participants in their care, and digitizing patient information for health care professionals, in line with the NHS 10-year health plan.

The aim of the current study was to conduct a process evaluation of URApp in pediatric continence clinics to understand how it is embedded into usual care. We gathered evidence to understand the feasibility and practicalities of embedding URApp in this real-world setting, identifying barriers and facilitators to implementation.

Methods

Study Design

This process evaluation used a mixed methods approach based on Normalization Process Theory (NPT). NPT is a framework used to guide the implementation and evaluation of complex interventions [28-30] and to understand how staff routinely incorporate a complex intervention into practice in a specific setting [31]. The Normalization Measure Development (NoMAD) questionnaire [32,33] operationalizes NPT across its 4 constructs: coherence, cognitive participation, collective action, and reflexive monitoring. Coherence refers to key stakeholders' understanding of and opinions about an intervention. Cognitive participation involves stakeholders' engagement with the intervention. Collective action involves putting the intervention into practice, and reflexive monitoring focuses on stakeholders' appraisal of the intervention and its effects [28].

Participant Characteristics

Pediatric continence nurses who introduced bladder training to young people aged 10 years and older in Children's Community Continence Services were eligible to participate.

XSL•FO
RenderX

Study Setting, Sample, and Recruitment

Children's Community Continence Services are Tier 2 community specialist services comprising nurse-led, multidisciplinary teams that provide more specialized assessment and management of bladder and bowel issues in young people. They aim to reduce unnecessary referrals to secondary care. Clinical coapplicants DR and JC assisted with identifying sites in different regions of England serving young people from diverse socioeconomic and ethnic backgrounds. URApp was not routinely embedded within participating Children's Community Continence Services before the study. As part of this process evaluation, teams were asked to familiarize themselves with URApp and were encouraged to offer it to eligible young people receiving bladder training during the 18-week study period. Decisions about whether and how URApp was used with individual patients remained at the discretion of pediatric continence nurses. Therefore, the study evaluated the implementation and integration of URApp into routine practice rather than testing a mandated implementation protocol.

We recruited Children's Community Continence Service nurses from 5 centers to provide a purposive maximum variation sample of 20 participants. DR distributed the study advert to the lead nurse at each site via email (Multimedia Appendix 1), who provided consent to be contacted by LB if interested in finding out more about the study. LB emailed the participant information sheet to each interested site and provided further information at an online meeting. NHS staff completed consent forms via REDCap (Vanderbilt University), a secure encrypted web application for building and managing online surveys and databases [34,35]. All participating team members were asked to familiarize themselves with URApp using *View the App Guide* and by downloading URApp onto a mobile device [20].

Data Collection and Analysis

Staff completed an online NoMAD questionnaire (Multimedia Appendix 2) via REDCap at baseline and 6, 12, and 18 weeks. The NoMAD questionnaire is a set of 20 questions for assessing implementation processes among participants directly involved in implementing complex interventions in health care [32,33]. Staff at each site were also invited to take part in focus groups or interviews. Focus groups and interviews (lasting around 30 - 45 min) were conducted online via Teams, audio-recorded with consent from participants, and transcribed, and data were deidentified. A topic guide (Multimedia Appendix 3) was developed in consultation with the clinical coapplicants. The topic guide drew on the NPT theory and had the flexibility to explore teams' NoMAD questionnaire responses (eg, if there was "low agreement" with NPT subconstructs at a particular site/time point).

Responses to NoMAD items were analyzed descriptively as follows: "strongly agree" and "agree" were interpreted as "high agreement" (5 and 4), "Neither agree nor disagree" as a "neutral" response (3), and "disagree" and "strongly disagree" as "low agreement" (2 and 1). NoMAD item-10 was reverse-coded so that the higher value indicated more agreement for ease of interpretation. Measures of the "normalization" of practice from the NoMAD questionnaires were summarized as median scores and plotted at each time point using petal plots. Higher agreement with NPT subconstructs and better integration of an intervention are characterized by more uniform "petals," closer to the outer limits of each petal plot (see Multimedia Appendix 4, Figures SA-SE, petal plots). Qualitative data analysis used the NVivo software (version 14; Lumivero) to organize data and drew on reflexive thematic analysis [36] with a deductive approach to explore how data fitted with NPT constructs and subconstructs. NoMAD outputs were cross-referenced with qualitative data to gain an in-depth understanding of the optimization of URApp in community clinic settings.

Patient and Public Involvement

The original idea for developing URApp came from our previous research with young people with UI, who told us that they needed more support to manage their symptoms [26] and felt they were not adequately supported by current services. URApp was co-designed and iteratively optimized through extensive work with young people with incontinence to ensure it is age-appropriate and meets their needs [26].

Ethical Considerations

Ethical approval was granted by the University of Bristol Faculty of Health Science Research Ethics Committee (reference 17560) and by the Health Research Authority and Health and Care Research Wales (IRAS project ID 338625, protocol number 2023 - 7539).

Results

Participant Characteristics

From April to October 2024, 5 Children's Community Continence Service teams in England were recruited to take part (Table S1 in Multimedia Appendix 5), with participants at each site completing NoMAD questionnaires from June 2024 to January 2025. Most participants were band 6 or above (specialist) pediatric continence nurses (n=14) and had worked in a pediatric continence clinic setting for up to 5 years (n=12; Table S2 in Multimedia Appendix 5). At baseline, half of the participants (n=10) reported that they had "never used" URApp; 8 participants stated that since the study began, they "sometimes" use it with patients, and 1 had started to "routinely use" the app with patients (Table S3 in Multimedia Appendix 5). Only 3 participants responded "never use" when they completed their final NoMAD questionnaire.

NoMAD Questionnaire Data

Twelve participants completed NoMAD questionnaires at 4 time points: baseline and 6, 12, and 18 weeks (sites A-D). The remaining 7 participants (site E) completed NoMAD questionnaires at baseline and 6 and 12 weeks (Multimedia Appendix 4). Broadly speaking, sites showed collective agreement with "Coherence" NoMAD items at all time points and high levels of neutrality (score of 3, neither agree nor disagree) in relation to cognitive participation, collective action, and reflexive monitoring items. A detailed descriptive analysis of NoMAD questionnaire data, participants' free-text response data, NoMAD median scores by sites at each time point, and raw NoMAD data is provided in Multimedia Appendix 6.

XSL•FO
RenderX

Focus Group and Interview Data

Findings from NoMAD questionnaire data were expanded upon during focus group and interview discussions to gain a more in-depth understanding of why teams struggled to integrate URApp. Five focus groups were conducted with 16 participants, and 3 participants took part in one-to-one interviews. One additional participant who did not complete a NoMAD questionnaire participated in a focus group (pediatric continence nurse or nonnurse, band 5 or below), and one participant who did complete NoMAD questionnaires did not take part in a focus group or interview (specialist pediatric continence nurse, band 6 or above). Illustrative themes, corresponding NPT constructs, and quotes can be found in Table 1 and Multimedia Appendix 7.

Table . Illustrative themes, corresponding Normalization Process Theory (NPT) constructs, and quotes (N=19).

Theme (NPT construct)/subtheme	Supporting quotation
1. URApp has potential (coherence)	
Positive view of app as age appropriate	*it is nice to have something for them I would say, because with all the toileting information, well a lot of it, is based on younger children in a way, and I felt that this was something a bit more relatable to them.* (participant 4, site B-Interview[a])
Paper bladder diaries not returned	*[bladder diaries] very few and far between they get returned as well, so compliance is an issue with completing charts. That's why we thought the app would be a bit easier, because they're constantly on the phone, even parents.* (participant 7, site C)
2. It's easy to introduce… But (cognitive participation)	
Easy to introduce	*I think it's easy to talk to them about it. in terms of actually putting it out there for them to try it, that's quite a straightforward thing to do.* (participant 12, site D)
Signposting to URApp not effective	*Quite often what I find, if I signpost a young person to it, then they, you see them or talk to them again and they haven't downloaded it or haven't looked at it.* (participant 19, site E)
Asking young people to download URApp in clinic more effective	*Most of them we get them to download it in clinic. So we know they've got it.* (participant 7, site C)
3. Behavior change is hard (collective action)	
Lack of engagement from young people demotivating for staff	*I'm not seeing the massive change I thought I would see, unfortunately, I'll still be suggesting it.* (participant 15, site E)
Lack of confidence in changing from paper to electronic bladder diaries	*…other concern from parents and carers would be that actually if you've got the young person directly collecting the data how is the parent or carer going to monitor that. Whereas when they've got a paper diary invariably it goes up on the fridge, or it gets stuck up in their bedroom.* (participant 2, site A)
4. Using URApp in a school environment (reflexive monitoring)	
Impact of school culture on engagement with bladder training	*there are quite a lot of barriers to doing the good toileting and the good drinking in school*. (participant 9, site D)
Lack of toilet access	*Massive issues over toilet access in the school, and drinking. If the education sector got that on-board we probably wouldn't have a job. they could probably use their phone more than they're allowed to go to the toilet.* (participant 7, site C)
Not able to access phone at school	*Then there's been at least four that I've spoken to about it, and the person's just gone, "Oh they're not allowed their mobile phones, they're not allowed the phones in school." when they can't use it in school they're very much like no, can't use this phone, and a little bit dismissive about it.* (participant 11, site D)
More usable bladder diary data with URApp	*that [neurodivergent male] I said, you know, normally you sit here and you shrug and you don't know, you don't know. But we've got that phone now, and I've seen more data now than I ever have done. I was surprised he was doing it. So he wasn't one I thought, actually, you're going to just download it and off you're going to go because that's normally what happens. But now I bought him back sort of the, I think two months later. Yeah. and we've got some data.* (participant 6, Site B-Interview[a])

[a]Indicates interview, all other participants took part in focus groups.

XSL•FO
RenderX

Thematic Analysis With Links to Core NPT Constructs

Theme 1: URApp Has Potential (Coherence—Understanding and Opinion of the Intervention)

Although some team members were aware of URApp and other apps that could be used to support hydration and set reminders, none were actively using URApp as part of their clinical practice prior to their involvement in the study. All team members were able to download URApp onto a phone and were enthusiastic about the prospect of introducing it to patients. Teams had a positive view of app functionality (internalization) and felt that URApp had the potential to improve young people's engagement with bladder training techniques (internalization). All discussed ways in which URApp could be integrated into clinical practice to support diagnosis, by recording fluid input and output, and to aid adherence to ongoing bladder training through drink and toilet reminders and working toward a drinking goal (communal specification).

Recording data digitally in URApp differed from teams' usual practice of asking families to record fluid input/output using paper bladder diaries (differentiation). All teams reflected on the fact that tracking young people's progress using paper-based bladder diaries was not effective because families were either not completing them accurately or frequently did not return them at all (individual specification). Teams could see how URApp reminder functions were similar to their past use of basic alarms to remind young people to use the toilet and/or drink; however, all acknowledged that using a digital app with all these functions was potentially more appealing and accessible for young people and inclusive in terms of neurodiversity (individual specification). Because URApp was a new process, 2 teams acknowledged that they sometimes forgot to discuss it with eligible patients, which had hindered its integration. Team members began to discuss other ways in which URApp could be integrated into their current practice, including supporting young people with an overactive bladder and bowel issues, enabling them to record and share when relevant symptoms were occurring via the "daily diary" function (eg, damp pants and tummy ache) and by allowing young people to develop their awareness of the need to drink and use the toilet regularly, despite not being able to use URApp in school (individual specification).

Theme 2: It's Easy to Introduce… But (Cognitive Participation—Engagement With the Intervention)

All sites reported that URApp was easy to introduce because it closely aligned with the bladder training information they would routinely provide (legitimation). Nurses enjoyed introducing URApp because it was more engaging than paper-based bladder diaries, and most stated that they would continue to introduce the app in the future (activation). Nurses began by signposting families to the URApp website and asked young people to download the app in their own time at home due to limited time during clinic appointments. However, at their next clinic appointment, there were reports of young people having forgotten to download URApp. Those who were not already doing so began to ask patients to download URApp onto their phones during the clinic appointment (legitimation). This was viewed by many as a key initial step in engaging young people and ensuring they did not simply forget to download the app when they got home. Across sites, participants identified practical strategies that appeared to support the introduction and use of URApp in routine care. These strategies were subsequently refined with input from the clinical coauthors (DR and JC) to produce a set of suggested best practice recommendations (Textbox 1). In addition to introducing URApp and explaining how it could be used to support bladder training techniques, 2 sites also made changes to standard team templates, such as including the URApp website link in letters sent to parents/carers after their child's clinical appointment (site E) or developing a leaflet with the URApp link and logo (site C), to display in waiting areas for families to take away (activation). Two further sites stated that they would consider adding URApp to team templates in the future (sites A and D).

Textbox 1. Best practice when using URApp in Children's Community Continence Services.

1. Ask young person to download URApp onto their phone in their clinic appointment
2. Send URApp weblink in the clinic letter after the initial assessment so that parent/carers also have access
3. Integrate URApp into a new or existing patient care plan
4. Show young person how to record fluid intake and urine output using the "volume" function in URApp
5. Use of URApp to collect volumes of drinking and voiding and data prior to clinic visits/at intervals for 48-hour periods as an electronic bladder diary
6. Ongoing use of URApp to support bladder training—use over a weekend, during school holidays
7. Provide a National Health Service email address so that young person or parent/carer can record/email fluid intake and urine output data to the clinical team

Although one person at each site acted as a "site contact" with the URApp research team, this person had limited time to act as a "champion" for URApp in terms of integrating it more widely into team processes (initiation). At the 18-week point, site D was still "neutral" in relation to NoMAD item 5: "'there are key people who drive URApp forward and get others involved." Although, according to NoMAD questionnaires, most stated they were open to working with colleagues in new ways to use URApp, engagement with the intervention did not progress to discussions at team meetings about more structured integration of URApp into existing care pathways (enrollment). Teams felt that training and support in the form of a QR code and animation would be useful to facilitate the introduction and ongoing use of URApp. The development of an animation was

XSL•FO
RenderX

discussed as something that young people, parents, and staff might use collectively, highlighting key app functions (such as downloading data) and providing useful examples about why certain aspects of bladder training are important. Despite team members liking URApp functionality, believing that it has potential benefits for use in the community clinic setting (coherence), and enjoying introducing it to families (cognitive participation), teams reported poor uptake and lack of long-term engagement with the app from young people. This in turn impacted how teams engaged with URApp in the long term (collective action).

Theme 3: Behavior Change Is Hard (Collective Action—Putting the Intervention Into Practice)

Teams were positive about the theory and behavior change techniques that URApp was based on but found these hard to implement in practice. When nurses reengaged with young people who had been introduced to URApp but not downloaded it or had downloaded it but had not continued to use it in the long term, it was demotivating for staff who had initially thought that introducing URApp to patients would motivate young people to engage with bladder training (lack of relational integration, NoMAD item 11). Although nurses discussed more examples of feedback about URApp from female patients, they did not report any specific gender differences in relation to URApp uptake or use; therefore, this may have been due to the fact that the prevalence of daytime UI is higher among females. Some commented that URApp was appropriate for and liked by patients with neurodivergent conditions who prioritized routines or needed reminders in relation to toileting (interactional workability), while others reflected that older teenagers were not engaging with URApp. However, some highlighted that these were often patients who had been managing their condition for a long period of time and had not previously engaged with bladder training. Others felt that younger teenagers liked the rewards (eg, streaks/trophies) or were motivated by the novelty of having access to a smartphone to engage with URApp.

Although nurses discussed actively trying to encourage young people to take ownership of their bladder training, they were doubtful that young people would fully engage with data capture in URApp without monitoring and input from a parent/carer (skill set workability, NoMAD item 12). The strategy of asking young people to use URApp at their initial appointment and to remind or show their nurse how they had used it at their next clinic appointment (which could be 3 or more months later) was also not successful. One focus group highlighted that young people who volunteer to participate in co-creation research (ie, those who codeveloped URApp) may be more motivated than the patients they typically see in their day-to-day clinical practice. Parents/carers were described by all teams as leading or being actively involved in the management of their child's bladder training; particularly, in relation to recording fluid input/output using a paper-based bladder diary for 2 to 3 days prior to their child attending an initial assessment appointment. Some suggested that URApp should allow a shared parent/child login or a secondary parent/carer URApp platform to minimize these issues.

Although parents/carers played a key role in their child's engagement with bladder training, one team discussed instances where they encouraged parents/carers to step back and allow their child to take the lead, highlighting the difficulties faced by community teams balancing clinical recommendations with parent/carer preferences and young people's autonomy. Patients who were proactive and engaged with bladder training were described as making quicker progress in terms of symptom improvement, but teams also reflected on some families' broader views that medication would be more effective than bladder training. There were also reports of some follow-up appointments being conducted remotely via a speaker or video phone call without the child/young person present. Because the child/young person was not present in these instances, nurses would be unable to discuss URApp since the parent would not have access to their child's URApp data on their smartphone (lack of interactional workability, NoMAD item 9).

Despite all sites reporting low engagement from patients and their parents/carers when using paper bladder diaries to gather data, several nurses stated that they would not feel confident collecting these data using URApp, and teams did not engage with the data export function available in URApp (lack of contextual integration, NoMAD item 14). Most did not recall that URApp allows young people to measure the volume (in mL) of their fluid intake and urine output, and all reported that the URApp data export function was not used by young people (lack of skill set workability, NoMAD item 13). Currently, URApp does not automatically share bladder diary data with community clinical teams, but young people can download data and send it by email. Teams generally viewed URApp data on the young person's phone during clinic appointments. None of the teams asked young people to download or email their URApp data for review during the consultation or for inclusion in the patient's notes (lack of contextual integration, NoMAD item 14).

Although all teams reported that some young people still lacked motivation, they also acknowledged that "it takes a while for parents and children to engage" (P4B) in bladder training. Staff reflected on the timing of NoMAD questionnaires at 6 and 12 weeks, as they noted that patients' next appointment could be 3 or more months after the appointment in which they had been introduced to URApp. Therefore, nurses may not yet have reviewed a patient who was actively using URApp, potentially resulting in neutral responses to "Collective action" NoMAD items. Nurses also commented that the independent use of URApp for recording input/output over a 3-month period without clinical support would not be manageable for many young people who were typically only asked to record their fluid intake and output for a fixed period (eg, a 2 - 3 days paper-based bladder diary). Teams found the integration of URApp into existing care pathways challenging and lacked confidence in changing from paper to electronic bladder diaries. Although they were positive about young people taking ownership of their bladder condition, they were concerned about the lack of parental oversight of bladder diary data in URApp (item 11: Relational integration, item 12: Skill set workability).

XSL•FO
RenderX

Theme 4: Using URApp in a School Environment (Reflexive Monitoring—Appraisal of the Intervention)

Nurses' appraisal of URApp as an effective way of collecting data going forward was also based on reports from young people that they were unable to use URApp effectively in their school environment (a lack of effective external systematization). Patients who were engaging with URApp were largely doing so outside of school, either at weekends or because they were not currently attending school. Nurses were frustrated when patients reported restrictive school toilet policies and inadequate or unpleasant toilet facilities that did not promote healthy bladder and bowel habits or did not enable bladder training adherence. Inadequate toilet facilities resulted in young people with bladder conditions not wanting to drink fluids at regular intervals so that they did not need to use the toilet during the school day. Young people also reported significant issues with some school toilets being locked during lesson time and only accessible during breaks, when significant queues formed. There were also reports of teachers restricting access during lesson time, only allowing those with a medical or toilet pass to leave the classroom, and the use of chaperones. Young people reported that they did not want to use a medical/toilet pass to gain access to toilets during lesson time because this might highlight their condition and lead to unwanted questions from peers.

Nurses highlighted that many schools now have strict phone use policies; they either do not allow young people to use their smartphones at all during school hours or they do not allow phone use during lessons (lack of positive communal appraisal of URApp). Since toilets in some schools were only open during breaks and at lunchtime, pupils drank water less to avoid the need to use the toilet, and there were also reports of young people having limited access to drinking water during lessons (eg, not being allowed to drink water at their desks). Despite these challenges, all teams were very positive about continuing to introduce URApp to young people, and NoMAD questionnaires showed agreement from nurses on how they could modify and work with URApp in the future (item 20: reconfiguration). One team was moving toward patient-initiated follow-up and felt that URApp could potentially fit well with this new approach. Alternatives to using URApp in the school environment were suggested, such as advising young people to focus on using URApp at weekends when they have less structure in their day and struggle to remember to drink regularly and use the toilet. Individual appraisal, in terms of "valuing the effects that URApp has had on my work," was apparent when a team member had a direct experience of seeing data output from URApp when it was used effectively by patients, particularly by those who had not previously engaged with bladder training.

Discussion

Principal Findings

This is the first study to explore facilitators and barriers to embedding a digital health intervention (URApp) in Children's Community Continence Services in England. We used a mixed methods approach combining results from the NoMAD questionnaire with qualitative data from interviews/focus groups. The main findings of this study were that URApp's functionality, alignment with bladder training techniques, age appropriateness, and accessibility (eg, supporting young people with neurodivergent conditions) facilitated staff engagement with URApp. Although nurses found URApp engaging and easy to introduce, they reported broader challenges relating to young people's engagement with bladder training. School toilet access policies were reported as a key barrier to successful engagement with bladder training. Nurses did not engage with the bladder diary data export function in URApp and had concerns about relying on young people to self-manage their bladder training and digital data without access being granted to parents/carers. We found that nurses acknowledged the importance of young people taking ownership of managing their bladder symptoms; however, they also stressed their belief that successful management should be part of a shared decision-making model that includes positive support from parents/carers. Although our sample was not large enough to draw firm conclusions, our findings tentatively suggest that engagement with URApp may vary by age, with younger children appearing more responsive to app-based motivational features than older teenagers. This may reflect age-related differences in motivation and preferences but could also be influenced by changing attitudes toward smartphone use among young people, including increasing awareness of screen time and digital well-being. Policies limiting mobile phone use in schools were major barriers to engagement with URApp. These factors warrant further investigation when optimizing and implementing digital bladder training interventions.

Strengths and Limitations

This is the first study in this area to use a process evaluation approach guided by NPT; we provide valuable insights into the challenges and facilitators of embedding a digital health intervention in community pediatric continence services. Using a mixed methods approach enabled us to capture behavior change over an 18-week period and to gain an in-depth understanding of current clinical practice. Challenges included clinical teams' resistance to changing usual practice (eg, paper-based bladder diaries) and school-based barriers that prevented young people's engagement with bladder training (policies relating to toilet access and mobile phone use). Our study included Children's Community Continence Service teams in different geographical regions of England, serving rural and urban populations, but transferability of our findings is limited by the small sample size and the small number of teams involved in the study. We aimed to recruit nurses with varying levels of experience, but our sample mainly comprised nurses with up to 5 years of experience in Children's Community Continence Services. We were only able to capture behavior change in the clinical teams over a relatively short period of time (at baseline and 6, 12, and 18 wk). Constraints included the short period of funding for the study (1 y) and the length of time taken to gain capacity and capability in the 5 participating NHS sites. Data collection time points did not align well with the time period (approximately 3 months) between the initial clinic appointment, in which URApp was introduced, and routine follow-up appointments. Some team members reported that they had not yet seen patients who had previously been introduced to URApp

XSL•FO
RenderX

at a follow-up appointment. Although young people were involved throughout the development and optimization of URApp [26], patient and public involvement was not incorporated into the design, conduct, or interpretation of this process evaluation. The absence of young people and parents/carers as study participants may have limited our understanding of implementation barriers and facilitators from the perspective of end users.

Links With the Existing Literature

Empirical evidence indicates that bladder training can be effective [13,37,38]; however, the success of bladder training depends on motivation and continued engagement [39,40]. There is evidence that digital health interventions can support pediatric bladder training [26,41,42], but barriers include individuals' beliefs about the effectiveness of conservative interventions, clinicians' resistance to adopting new approaches, and environmental barriers.

Engagement With Conservative Interventions

One of our key findings was that children and young people experienced difficulties engaging with bladder training because of concerns about standing out from their peers at school and dissatisfaction with school toilet facilities. This aligns with findings from a European focus group study of children aged 9 to 13 years, which identified fear of attracting unwanted attention and dissatisfaction with school toilets as important influences on engagement with bladder training [39]. We also found that participants described frustration with symptom relapse and a lack of long-term engagement with URApp. These findings are consistent with previous UK research in which young people reported frustration with bladder training due to treatment failure and relapse, alongside disengagement from treatment following unsuccessful symptom management [9]. Participants described how school-related barriers could limit adherence to bladder training, which may contribute to perceptions among parents and young people that medication is a more practical treatment option. Similar barriers have been reported in studies of children's bladder health in school settings [43]. However, evidence from adult populations suggests that bladder training may be as effective as, or more effective than, anticholinergic medication for improving bladder symptoms [44]; therefore, further research is needed in pediatric populations.

Resistance to Adopting Digital Systems

Although health care professionals recognized the potential advantages of electronic bladder diaries, many remained reluctant to replace established paper-based approaches. This finding is consistent with implementation research showing that the adoption of digital innovations depends on staff confidence, organizational support, and the presence of local champions who facilitate new ways of working [45-47]. Evidence from mobile health research similarly highlights the important role of health care professionals in supporting engagement with digital interventions [48]. Our findings therefore suggest that implementation efforts should focus not only on patient engagement but also on enhancing health care professionals' confidence and readiness to adopt digital continence technologies.

Environmental Barriers

Although a recent systematic review reported that mobile apps for urological conditions can improve the quality of life and symptoms in adults through real-time data capture and self-monitoring [49], our findings suggest that the benefits of digital bladder training interventions may be limited by the school environment. Participants described school policies that limit access to toilets (only during morning/afternoon break or at lunchtime) and reported that toilet facilities were inadequate and unhygienic. These findings are consistent with previous research demonstrating that school environments can hinder healthy toileting and drinking behaviors [6,50-53] and are further supported by a recent systematic review, showing that restricted toilet access, poor toilet conditions, and privacy concerns represent significant barriers to healthy toileting behaviors in a school environment [54]. These findings highlight the need for greater awareness among school staff of the ongoing toileting requirements associated with bladder training, particularly in school environments where toilet access is closely regulated [55].

Participants also highlighted restricted access to mobile phones during the school day as a potential barrier to engagement with URApp. This challenge may become increasingly important as schools adopt stricter policies regarding phone use [56,57]. Although smartphone ownership among young people is high [58], more than half of children aged 8 to 17 years are not permitted to use their phones during school hours [59]. While UK guidance emphasizes schools' responsibilities to support pupils with medical conditions, including access to drinking water and toilet breaks when needed, it does not address the use of digital health tools to support self-management [55,57]. Given growing evidence that mobile health interventions can improve adherence and self-management across a range of health conditions [60], further research is needed to determine how school phone policies can support equitable access to digital health interventions for young people with medical conditions.

Implications for Policy, Practice, and Future Research

The successful implementation of digital bladder training interventions requires action at multiple levels. Within Children's Community Continence Services, efforts should focus on increasing NHS staff confidence in digital tools and integrating electronic bladder diary data into routine practice. Although health care professionals recognized the potential value of URApp, reluctance to replace established paper-based approaches suggests that staff training, organizational support, and local implementation champions may be important for successful adoption [45,61]. Our findings also highlight the importance of involving young people, parents/carers, and health care professionals in shared decision-making. Previous research found that young people wished to take ownership of their continence management and viewed smartphone apps as a supportive tool for self-management [26]. Future research should explore how shared decision-making influences engagement with digital interventions and adherence to bladder training.

XSL•FO
RenderX

Beyond health care settings, participants in this study consistently identified the school environment as a major barrier to bladder training adherence. Schools should be supported to develop policies that balance educational priorities with pupils' needs for regular toilet access and hydration [6,52,54]. In addition, schools should consider how digital health technologies can be accommodated for pupils who rely on them to manage long-term health conditions [55,57]. As schools increasingly adopt restrictions on mobile phone use, alternative implementation strategies may be required, including negotiating health-related exemptions, permitting their use during school holidays, or delaying data entry outside school hours. More broadly, our findings highlight the importance of considering organizational, educational, and policy contexts when implementing digital continence interventions.

Conclusions

Using URApp in community clinics has the potential to align pediatric continence care with the NHS 10-year health plan [62] by focusing on early intervention in a community setting and using digital technology to capture and store data. URApp is free to download, available for both iOS and Android phones, and has the potential to empower young people to self-manage their symptoms between clinic appointments. Digitizing bladder diaries may lead to improved treatment outcomes, improved patient experience, and reduced barriers to accessing health care. Future implementation of URApp into existing clinical care pathways will need to consider how electronic bladder diary data from URApp can be used more effectively by clinical teams, in partnership with patients and their parents/carers.

Acknowledgments

The authors thank all the nurses who participated in this study. Generative AI was used to assist with language editing, including improving grammar, readability, and clarity of the manuscript. The authors reviewed and revised all AI-generated suggestions and take full responsibility for the content, interpretation, and conclusions presented in the manuscript.

Funding

This work was supported by the Medical Research Council Impact Acceleration Account and the University of Bristol internal funding scheme and was peer-reviewed by an internal panel during the selection process.

Data Availability

Descriptive quantitative data are available in Multimedia Appendix 6. The qualitative data are not publicly available to protect the anonymity of staff but are available from the corresponding author upon reasonable request. Illustrative themes, corresponding Normalization Process Theory constructs, and quotations can be found in Multimedia Appendix 7. Anonymized research data will be stored for 10 years and have been deposited in the University of Bristol Research Data Repository, with controlled access. CJ is the custodian for all data generated by this study and will decide whether/which research data are appropriate to be shared upon request.

Authors' Contributions

Conceptualization: CJ (lead), LB (supporting), JI (supporting)
Data curation: LB (lead), CJ (supporting)
Formal analysis: LB (lead), CJ (supporting), JI (supporting)
Funding acquisition: CJ
Investigation: LB (lead), CJ (supporting), JI (supporting)
Methodology: LB (lead), CJ (equal), JI (equal)
Project administration: LB (lead), CJ (supporting), JI (supporting)
Resources: CJ (lead), DR (supporting), JC (supporting), LB (supporting)
Supervision: CJ
Validation: CJ (lead), JI (supporting)
Visualization: LB (lead), CJ (supporting), JI (supporting)
Writing – original draft: LB (lead), CJ (supporting)
Writing – review and editing: LB (lead), CJ (equal), JI (supporting), DR (supporting), JC (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study advert and email subject.
[DOCX File, 36 KB - nursing_v9i1e95418_app1.docx]

Multimedia Appendix 2

XSL•FO
RenderX

Normalization Measure Development Questionnaire.
[DOCX File, 38 KB - nursing_v9i1e95418_app2.docx]

Multimedia Appendix 3
Topic guide.
[DOCX File, 37 KB - nursing_v9i1e95418_app3.docx]

Multimedia Appendix 4
Petal plots.
[DOCX File, 731 KB - nursing_v9i1e95418_app4.docx]

Multimedia Appendix 5
Participant characteristics.
[DOCX File, 39 KB - nursing_v9i1e95418_app5.docx]

Multimedia Appendix 6
Normalization Measure Development Questionnaire data descriptive analysis.
[DOCX File, 86 KB - nursing_v9i1e95418_app6.docx]

Multimedia Appendix 7
Illustrative themes and quotes.
[DOCX File, 52 KB - nursing_v9i1e95418_app7.docx]

References

1. Nieuwhof-Leppink AJ, Schroeder RPJ, van de Putte EM, de Jong T, Schappin R. Daytime urinary incontinence in children and adolescents. Lancet Child Adolesc Health 2019 Jul;3(7):492-501. [doi: 10.1016/S2352-4642(19)30113-0] [Medline: 31060913]
2. Heron J, Grzeda MT, von Gontard A, Wright A, Joinson C. Trajectories of urinary incontinence in childhood and bladder and bowel symptoms in adolescence: prospective cohort study. BMJ Open 2017 Mar 14;7(3):e014238. [doi: 10.1136/bmjopen-2016-014238] [Medline: 28292756]
3. Sureshkumar P, Jones M, Cumming R, Craig J. A population based study of 2,856 school-age children with urinary incontinence. J Urol 2009 Feb;181(2):808-815. [doi: 10.1016/j.juro.2008.10.044] [Medline: 19110268]
4. Swithinbank LV, Brookes ST, Shepherd AM, Abrams P. The natural history of urinary symptoms during adolescence. Br J Urol 1998 May;81 Suppl 3:90-93. [doi: 10.1046/j.1464-410x.1998.00016.x] [Medline: 9634028]
5. Yüksel S, Yurdakul A, Zencir M, Çördük N. Evaluation of lower urinary tract dysfunction in Turkish primary schoolchildren: an epidemiological study. J Pediatr Urol 2014 Dec;10(6):1181-1186. [doi: 10.1016/j.jpurol.2014.05.008] [Medline: 25001292]
6. Whale K, Cramer H, Joinson C. Left behind and left out: the impact of the school environment on young people with continence problems. Br J Health Psychol 2018 May;23(2):253-277. [doi: 10.1111/bjhp.12284] [Medline: 29228510]
7. Gordon K, Warne N, Heron J, von Gontard A, Joinson C. Continence problems and mental health in adolescents from a UK cohort. Eur Urol 2023 Nov;84(5):463-470. [doi: 10.1016/j.eururo.2023.05.013] [Medline: 37248139]
8. Joinson C, Richardson D, Rayner J, Fowler S, Lynden A. Paediatric incontinence: early identification and treatment needed to prevent mental health problems. : University of Bristol, PolicyBristol; 2023 URL: https://www.bristol.ac.uk/media-library/sites/policybristol/briefings-and-reports-pdfs/2023-briefings/PolicyReport89_Joinson_incontinence_mental_health.pdf [accessed 2026-09-11]
9. Whale K, Cramer H, Wright A, Sanders C, Joinson C. "What does that mean?": a qualitative exploration of the primary and secondary clinical care experiences of young people with continence problems in the UK. BMJ Open 2017 Oct 16;7(10):e015544. [doi: 10.1136/bmjopen-2016-015544] [Medline: 29042374]
10. Advice for children with daytime bladder problems. : ERIC (The Children's Bladder and Bowel Charity); 2024 URL: https://eric.org.uk/wp-content/uploads/2026/01/Advice-for-children-with-daytime-bladder-problems.pdf.pdf [accessed 2026-09-11]
11. Parker A. Urinary incontinence guideline. : Ashford and St Peter's Hospitals, NHS Foundation Trust; 2022 URL: https://ashfordstpeters.net/Guidelines_Paediatrics/Incontinence-Guideline-Oct-2022.pdf [accessed 2026-09-11]
12. Palmer C, Richardson D, Rayner J, Drake MJ, Cotterill N. Professional perspectives on impacts, benefits and disadvantages of changes made to community continence services during the COVID-19 pandemic: findings from the EPICCC-19 national survey. BMC Health Serv Res 2022 Jun 15;22(1):783. [doi: 10.1186/s12913-022-08163-3] [Medline: 35705988]
13. Assis GM, da Silva CPC, Martins G. Urotherapy in the treatment of children and adolescents with bladder and bowel dysfunction: a systematic review. J Pediatr (Rio J) 2019;95(6):628-641. [doi: 10.1016/j.jped.2019.02.007] [Medline: 31009619]

XSL•FO
RenderX

14. Bedwetting (enuresis): diagnosis—assessment. National Institute for Health and Care Excellence (NICE). URL: https://cks.nice.org.uk/topics/bedwetting-enuresis/diagnosis/assessment/#primary-bedwetting-with-daytime-symptoms [accessed 2026-09-11]
15. Day-time incontinence. : Swedish Enuresis Academy; 2022 URL: https://i-c-c-s.org/wp-content/uploads/2022/06/Incontinence-Leaflet.pdf [accessed 2026-09-11]
16. Children and young people—bladder and bowel health. Bladder and Bowel UK. 2025. URL: https://www.bbuk.org.uk/children-young-people [accessed 2025-12-15]
17. Buckley BS, Sanders CD, Spineli L, Deng Q, Kwong JS. Conservative interventions for treating functional daytime urinary incontinence in children. Cochrane Database Syst Rev 2019 Sep 18;9(9):CD012367. [doi: 10.1002/14651858.CD012367.pub2] [Medline: 31532563]
18. Joinson C, Sullivan S, von Gontard A, Heron J. Early childhood psychological factors and risk for bedwetting at school age in a UK cohort. Eur Child Adolesc Psychiatry 2016 May;25(5):519-528. [doi: 10.1007/s00787-015-0756-7] [Medline: 26294078]
19. Warne N, Heron J, von Gontard A, Joinson C. Mental health problems, stressful life events and new-onset urinary incontinence in primary school-age children: a prospective cohort study. Eur Child Adolesc Psychiatry 2024 Mar;33(3):871-879. [doi: 10.1007/s00787-023-02211-x] [Medline: 37095371]
20. URApp. URL: https://urapp.org.uk/ [accessed 2026-09-11]
21. Schäfer SK, Niemczyk J, von Gontard A, Pospeschill M, Becker N, Equit M. Standard urotherapy as first-line intervention for daytime incontinence: a meta-analysis. Eur Child Adolesc Psychiatry 2018 Aug;27(8):949-964. [doi: 10.1007/s00787-017-1051-6] [Medline: 28948380]
22. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232. [doi: 10.2196/jmir.7126] [Medline: 28663162]
23. O'Cathain A, Croot L, Duncan E, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8):e029954. [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
24. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30. [doi: 10.2196/jmir.4055] [Medline: 25639757]
25. Holt L, Denford S, Bowers H, et al. The person-based approach to intervention development: a scoping review of methods and applications. Digit Health 2025;11:20552076241305934. [doi: 10.1177/20552076241305934] [Medline: 39801581]
26. Whale K, Beasant L, Wright AJ, et al. A smartphone app for supporting the self-management of daytime urinary incontinence in adolescents: development and formative evaluation study of URApp. JMIR Pediatr Parent 2021 Nov 15;4(4):e26212. [doi: 10.2196/26212] [Medline: 34779780]
27. Cotterill N, Schroeder K. CONfidence: developing an app to promote bladder and bowel health. Proc Inst Mech Eng H 2024 Jun;238(6):628-632. [doi: 10.1177/09544119241246848] [Medline: 38627995]
28. May C, Finch T, Mair F, et al. Understanding the implementation of complex interventions in health care: the normalization process model. BMC Health Serv Res 2007 Sep 19;7:148. [doi: 10.1186/1472-6963-7-148] [Medline: 17880693]
29. May CR, Albers B, Bracher M, et al. Translational framework for implementation evaluation and research: a normalisation process theory coding manual for qualitative research and instrument development. Implement Sci 2022 Feb 22;17(1):19. [doi: 10.1186/s13012-022-01191-x] [Medline: 35193611]
30. May CR, Mair F, Finch T, et al. Development of a theory of implementation and integration: Normalization Process Theory. Implement Sci 2009 May 21;4:29. [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
31. Forster DA, Newton M, McLachlan HL, Willis K. Exploring implementation and sustainability of models of care: can theory help? BMC Public Health 2011 Nov 25;11(Suppl 5):S8. [doi: 10.1186/1471-2458-11-S5-S8] [Medline: 22168585]
32. Rapley T, Girling M, Mair FS, et al. Improving the normalization of complex interventions: part 1—development of the NoMAD instrument for assessing implementation work based on Normalization Process Theory (NPT). BMC Med Res Methodol 2018 Nov 15;18(1):133. [doi: 10.1186/s12874-018-0590-y] [Medline: 30442093]
32. Finch TL, Girling M, May CR, et al. Improving the normalization of complex interventions: part 2—validation of the NoMAD instrument for assessing implementation work based on Normalization Process Theory (NPT). BMC Med Res Methodol 2018 Nov 15;18(1):135. [doi: 10.1186/s12874-018-0591-x] [Medline: 30442094]
34. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
35. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
36. Braun V, Clarke V. Thematic Analysis: A Practical Guide: Sage Publications Ltd; 2021. URL: https://www.sagepub.com/shop/buy-a-book/thematic-analysis-1-248481 [accessed 2026-09-11]
37. Jessen AS, Hagstroem S, Borch L. Comparison and characteristics of children successfully treated for daytime urinary incontinence. J Pediatr Urol 2022 Feb;18(1):24. [doi: 10.1016/j.jpurol.2021.11.013] [Medline: 34930690]

XSL•FO
RenderX

38. Morizawa Y, Mieda K, Tachibana A, et al. Characteristics of children successfully treated for daytime urinary incontinence. Int J Urol 2025 Jun;32(6):633-638. [doi: 10.1111/iju.70020] [Medline: 40079335]
39. Bladt L, Nieuwhof-Leppink A, Blomme RF, et al. Understanding pediatric patient experiences with urotherapy tools: qualitative focus group study. JMIR Pediatr Parent 2025 Oct 16;8:e79142. [doi: 10.2196/79142] [Medline: 41100523]
40. Caldwell PHY, Kerr M, Hamilton S, Teixeira-Pinto A, Craig JC. An alarm watch for daytime urinary incontinence: a randomized controlled trial. Pediatrics 2022 Jan 1;149(1):e2021053863. [doi: 10.1542/peds.2021-053863] [Medline: 34907443]
41. Caldwell PHY, Richards D, Hamilton S, et al. A randomized controlled trial of a web-based management support system for children with urinary incontinence: the eADVICE trial. J Urol 2024 Mar;211(3):364-375. [doi: 10.1097/JU.0000000000003832] [Medline: 38150394]
42. Bladt L, Vermeulen J, Vermandel A, De Win G, Van Campenhout L. Innovative, technology-driven, digital tools for managing pediatric urinary incontinence: scoping review. Interact J Med Res 2025 May 5;14:e66336. [doi: 10.2196/66336] [Medline: 40324170]
43. Venkatapuram P, Angulo-Lozano JC, Spinzi S, Brown-Johnson C, Phord-Toy A, Kan KM. Parental knowledge, beliefs, practices, and barriers related to children's bladder health in the school environment. J Sch Health 2024 Oct;94(10):965-974. [doi: 10.1111/josh.13456] [Medline: 38711264]
44. Funada S, Yoshioka T, Luo Y, Sato A, Akamatsu S, Watanabe N. Bladder training for treating overactive bladder in adults. Cochrane Database Syst Rev 2023 Oct 9;10(10):CD013571. [doi: 10.1002/14651858.CD013571.pub2] [Medline: 37811598]
45. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation Sci 2009 Dec;4(1):50. [doi: 10.1186/1748-5908-4-50]
46. Wynn M, Garwood-Cross L, Vasilica C, Griffiths M, Heaslip V, Phillips N. Digitizing nursing: a theoretical and holistic exploration to understand the adoption and use of digital technologies by nurses. J Adv Nurs 2023 Oct;79(10):3737-3747. [doi: 10.1111/jan.15810] [Medline: 37530425]
47. Hammerton M, Benson T, Sibley A. Readiness for five digital technologies in general practice: perceptions of staff in one part of southern England. BMJ Open Qual 2022 Jun;11(2):e001865. [doi: 10.1136/bmjoq-2022-001865] [Medline: 35768171]
48. Alkhuzaimi F, Rainey D, Brown Wilson C, Bloomfield J. The impact of mobile health interventions on service users' health outcomes and the role of health professions: a systematic review of systematic reviews. BMC Digit Health 2025;3(1):3. [doi: 10.1186/s44247-024-00143-3]
49. Rosa D, Villa G, Bonetti L, et al. The use of mobile applications in urology. A systematic review. Int J Urol Nurs 2021 Nov;15(3):105-116. [doi: 10.1111/ijun.12282]
50. Palmer MH, Athanasopoulos A, Lee KS, Takeda M, Wyndaele JJ. Sociocultural and environmental influences on bladder health. Int J Clin Pract 2012 Dec;66(12):1132-1138. [doi: 10.1111/ijcp.12029] [Medline: 23163494]
51. Haines Lyon C, Little A, Dobson E, et al. Toilet talk: using a students as researchers approach to problematize and co-construct school toilet policy and practice. Gend Educ 2024 Oct 2;36(7):801-816. [doi: 10.1080/09540253.2024.2389108]
52. Jørgensen CS, Breinbjerg AS, Rittig S, Kamperis K. Dissatisfaction with school toilets is associated with bladder and bowel dysfunction. Eur J Pediatr 2021 Nov;180(11):3317-3324. [doi: 10.1007/s00431-021-04111-1] [Medline: 33999258]
53. Voices for change report: a co-production insights and engagement project on bowel and bladder health and the challenges young people face. : ERIC (The Children's Bowel and Bladder Charity); 2025 URL: https://eric.org.uk/voices-for-change-report/#voices-for-change-report [accessed 2026-09-11]
54. Li A, Hillion JA, Godeau E. School toilet use and avoidance: a systematic review of barriers, facilitators, and impact on child and adolescent health. J Public Health (Berl) 2026. [doi: 10.1007/s10389-026-02694-3]
55. Supporting pupils at school with medical conditions. : Department for Education; 2015 URL: https://assets.publishing.service.gov.uk/media/5ce6a72e40f0b620a103bd53/supporting-pupils-at-school-with-medical-conditions.pdf [accessed 2026-09-11]
56. Mobile phones in schools. : Department for Education; 2026 URL: https://dera.ioe.ac.uk/id/eprint/41821/1/Mobile%20phones%20in%20schools%20-%20GOV.UK.pdf [accessed 2026-09-11]
57. Creating a mobile phone-free school environment. : Department for Education; 2024 URL: https://dera.ioe.ac.uk/id/eprint/40536/2/Creating_a_mobile_phone-free_school_environment.pdf [accessed 2026-09-11]
58. Mobile phones in schools: are they being banned? The Education Hub. 2024. URL: https://educationhub.blog.gov.uk/2024/10/mobile-phones-in-schools-are-they-being-banned/ [accessed 2026-09-11]
59. Children and parents: media use and attitudes report. : Ofcom; 2025 URL: https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/children/childrens-media-use-and-attitudes-report-2025/childrens-media-literacy-report-2025.pdf?v=396621 [accessed 2026-09-11]
60. Badawy SM, Barrera L, Sinno MG, Kaviany S, O'Dwyer LC, Kuhns LM. Text messaging and mobile phone apps as interventions to improve adherence in adolescents with chronic health conditions: a systematic review. JMIR Mhealth Uhealth 2017 May 15;5(5):e66. [doi: 10.2196/mhealth.7798] [Medline: 28506955]

XSL•FO
RenderX

61. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. Milbank Q 2004;82(4):581-629. [doi: 10.1111/j.0887-378X.2004.00325.x] [Medline: 15595944]
62. Fit for the future: 10 year health plan for England. : Department of Health and Social Care (DHSC); 2025 URL: https://assets.publishing.service.gov.uk/media/6888a0996478525675738f3a/fit-for-the-future-10-year-health-plan-for-england-executive-summary.pdf [accessed 2026-09-11]

Abbreviations

ERIC: The Children's Bladder and Bowel Charity
NHS: National Health Service
NoMAD: Normalization Measure Development
NPT: Normalization Process Theory
UI: urinary incontinence

Edited by E Borycki; submitted 16.Mar.2026; peer-reviewed by CH Lyon, Z Davey; revised version received 02.Sep.2026; accepted 07.Sep.2026; published 27.Oct.2026.

Please cite as:
Beasant L, Ingram J, Richardson D, Clothier J, Joinson C
Process Evaluation of a Digital Smartphone Intervention (URApp) to Support Adherence to Bladder Training in Community-Based Specialist Pediatric Continence Care in England: Mixed Methods Study
JMIR Nursing 2026;9:e95418
URL: https://nursing.jmir.org/2026/1/e95418
doi:10.2196/95418

© Lucy Beasant, Jenny Ingram, Davina Richardson, Joanna Clothier, Carol Joinson. Originally published in JMIR Nursing (https://nursing.jmir.org), 27.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI Use, Perceptions, and Perceived Impact Among Nursing Students: Cross-Sectional Study

Ilda Taka[1], PhD; Elona Hasalla[2*], Doc; Albana Sula[3*], PhD; Blerina Bahiti[4*], PhD; Rajmonda Oboni[3*], Doctoral Researcher; Blerina Bani[2*], PhD

[1]Department of Medical Technical Specialties, Faculty of Medical Technical Sciences, University of Elbasan, 3 Dëshmorët Street, Tirana, Albania

[2]Department of Paraclinical Subjects, Faculty of Medical Technical Sciences, University of Elbasan, Elbasan, Albania

[3]Department of Nursing, Faculty of Medical Technical Sciences, University of Elbasan, Elbasan, Albania

[4]Department of Clinical Subjects, Faculty of Medical Technical Sciences, University of Elbasan, Elbasan, Albania

[*]these authors contributed equally

Corresponding Author:
Ilda Taka, PhD
Department of Medical Technical Specialties, Faculty of Medical Technical Sciences, University of Elbasan, 3 Dëshmorët Street, Tirana, Albania

Abstract

Background: AI is increasingly being integrated into education and health care, offering opportunities to improve learning, understanding of clinical cases, and students' self-confidence. However, it remains necessary to assess nursing students' perceptions of AI and its impact on their academic and professional development.

Objective: The aim of the study was to assess AI use among nursing students, their perceptions of AI, and its impact on learning, academic performance, and professional preparation.

Methods: A descriptive cross-sectional study with analytical components was conducted at the Faculty of Medical Technical Sciences in Elbasan, Albania. Data were collected through a structured questionnaire administered via Google Forms, which assessed AI use and its perceived impact on learning and professional preparation. Data were analyzed using SPSS (version 23.0). Descriptive statistics, chi-square tests for associations between variables (P<.05), and logistic regression to estimate crude odds ratios (CORs) and adjusted odds ratios (AORs) with 95% CIs were used.

Results: A total of 279 nursing students participated in the study (mean age 22.4, SD 5.7 years), the majority of whom were female (273/279, 97.8%), lived in urban areas (156/279, 55.9%), and were enrolled in the bachelor's program (198/279, 71%). Overall, 83.9% (234/279) reported using AI, mainly virtual assistants such as ChatGPT or similar tools (183/234, 78.2%). The most common reason was information searching (216/234, 92.3%), followed by studying and understanding lecture content (96/234, 41%). Bivariate analysis showed no significant associations between AI use and residence or study cycle, whereas grade point average (GPA) was significantly associated with AI use. In the multivariable analysis, GPA remained the only independent predictor of AI use. Students with a GPA of 6.0 to 6.9 (AOR 5.55, 95% CI 1.53-20.14; P=.009) and those with a GPA of 8.0 to 8.9 (AOR 5.73, 95% CI 1.26-26.00; P=.02) were significantly more likely to use AI than the reference group. Students perceived AI as having a moderate impact on learning, particularly understanding lectures (mean 2.58, SD 1.10) and exam preparation (mean 2.58, SD 1.02), whereas its impact on self-confidence (mean 1.97, SD 1.17) and empathy (mean 1.90, SD 1.12) was perceived as low. Although AI was considered useful for supporting learning (mean 2.83, SD 1.12), students expressed concerns regarding the reliability of AI-generated information (mean 3.18, SD 1.22), dependence on AI (mean 2.75, SD 1.28), and its impact on critical thinking (mean 2.80, SD 1.18).

Conclusions: AI is widely used among nursing students, primarily supporting learning and academic performance. However, its impact on professional and interpersonal competencies remains limited. These findings suggest the need for integrating AI into nursing education curricula, with a focus on critical use and the development of students' professional competencies.

(JMIR Nursing 2026;9:e97832) doi:10.2196/97832

KEYWORDS

artificial intelligence; nursing students; impact; perception; nursing education

XSL•FO
RenderX

Introduction

In the last decade, AI has been one of the most important developments in the field of technology, and it has been integrated into various fields, including health care [1-3]. Its use is changing not only the way health care services are delivered but also the way future health care professionals are prepared, especially in the context of decision-making and patient management [4].

AI is also having a visible impact on nursing education. Today, students no longer rely solely on traditional learning methods but increasingly use digital methods to understand and assimilate their theoretical and practical knowledge [5]. AI-based technologies such as virtual simulations, virtual patients, and intelligent systems supporting decision-making have enabled an interactive and safe environment for the development of critical thinking and clinical reasoning [6,7]. These tools also contribute to the development of nursing care plans and an improved understanding of complex clinical processes [8].

Recently, a special role has also been played by various generative AI platforms such as ChatGPT and similar tools, which have become part of students' daily lives [9]. These tools offer quick access to information, help explain difficult concepts, support the processing of academic materials, and generate structured responses similar to those produced by humans [10-12]. For nursing students, these technologies provide potential support for academic performance and professional development by linking theoretical, practical, and clinical knowledge [13].

Various studies have shown that the use of AI improves clinical reasoning abilities, critical thinking, and self-confidence in practice [1,14]. AI also provides personalized and interactive information, adapting to students' needs and offering rapid feedback [15,16]. At the same time, the use of technology can help improve the quality of patient care through continuous support in monitoring and decision-making [17].

However, beyond its benefits, the literature shows that the use of AI is associated with a series of challenges and concerns in the field of nursing education. Uncontrolled use of AI may lead to reduced active student involvement in the learning process, increased dependence on technology, and reduced professional competence [18-20]. Other concerns relate to the accuracy of information, academic privacy, and the risk of plagiarism [21,22], issues that are particularly sensitive in the field of nursing.

Another aspect that should not be neglected is the need to develop new competencies among nursing students, including the critical use of technology and data analysis [23]. Importantly, AI cannot replace the human elements of the nursing profession such as empathy, communication, and patient care [24].

Although the use and impact of AI in nursing education continue to grow, evidence on how students use and perceive it in practice remains limited. Most studies focus on evaluating the theoretical benefits of AI, while less attention has been paid to students' real-life experiences and the factors influencing their use of AI in the educational context [11]. Moreover, there are several barriers such as the lack of standardization in AI education and inequalities in access [25].

It is important to study nursing students' perceptions and experiences regarding the use of AI tools, with a focus on identifying benefits, challenges, and implications to develop educational policies for the most effective and ethical use of AI [26,27].

Therefore, the aim of this study was to assess the use of AI among nursing students, their perceptions of AI, and the perceived impact of AI tools on the learning process, academic performance, and professional preparation.

Methods

Ethical Considerations

The study was not submitted for formal ethics review because it was conducted through an anonymous, voluntary, minimal-risk web-based survey and did not involve the collection of personally identifiable or sensitive personal information. Electronic informed consent was obtained from all participants before accessing the questionnaire. Participants were informed about the purpose of the study, voluntary participation, confidentiality, and anonymity. No personally identifiable information was collected, and participants could withdraw at any time by choosing not to complete the survey. All data were collected, stored, and analyzed anonymously in accordance with the general principles of research integrity outlined in the institutional framework [28], and no compensation was provided.

Study Design and Setting

A descriptive cross-sectional study with analytical components was conducted among students enrolled in the General Nursing program at the Faculty of Medical Technical Sciences, University of Elbasan, Albania. The study was carried out between January 2026 and March 2026.

Study Population and Sampling

The study included all students enrolled in the bachelor's and master's General Nursing programs. Students from other study programs and other universities were excluded. Participants were selected using convenience sampling, and participation was voluntary. The minimum sample size was calculated using the Cochran formula with a 95% CI, a 5% margin of error, and a total population of 1020 students. The calculation resulted in a minimum required sample size of 279 participants.

Data Collection Procedure

Data were collected using a structured, self-administered questionnaire developed by the authors and distributed online through Google Forms (Alphabet Inc) in student communication groups. Participants were informed about the purpose of the study and were assured of the confidentiality and anonymity of their responses. Before completing the questionnaire, all participants provided electronic informed consent. Only fully completed questionnaires were included in the final analysis. Due to the method of questionnaire distribution, the exact

XSL•FO
RenderX

number of students who received the survey link could not be determined.

Study Instrument

The study instrument was developed by the authors based on the scientific literature regarding the use of AI in nursing education. The questionnaire was developed in Albanian and consisted of 3 sections.

The first section included demographic and academic information, including age, sex, place of residence (urban or rural), study cycle (bachelor's or master's), and academic grade point average (GPA). In the Albanian grading system, grades range from 4 to 10, with 10 representing the highest academic achievement. For analytical purposes, GPA was categorized into 5 groups: <6.0, 6.0 to 6.9, 7.0 to 7.9, 8.0 to 8.9, and 9.0 to 10.0. This section also included questions regarding AI use, frequency of use, types of AI tools used, and purposes of use. Questions related to AI tool types and purposes of use allowed multiple responses. AI tools were categorized as virtual assistants (ChatGPT [OpenAI], DeepSeek [DeepSeek AI], and Copilot [Microsoft Corp]), educational platforms (Learning Nurse; Steppingstones Partnership Inc), text-generation and text-enhancement tools (paraphrase tool), and other tools used for academic support and information retrieval. AI use was initially assessed using a screening question (yes or no). Only participants who reported using AI proceeded to the subsequent questionnaire sections, which assessed frequency of use, types of AI tools used, and the perceived impact of AI on learning and professional preparation. Participants who reported not using AI completed only the demographic and academic sections.

The second section assessed the perceived impact of AI on learning and professional preparation. This section included questions on AI's impact on understanding lectures, exam preparation, academic performance, time saving during studying, understanding clinical cases, nursing care, self-confidence during clinical practice, critical thinking, communication skills, and empathy toward patients. All questions were assessed using a 5-point Likert scale ranging from 1 ("no influence") to 5 ("very high influence"). For statistical analysis, these variables were grouped into 2 main dimensions: academic impact and professional impact.

The final section assessed students' perceptions regarding the use of AI in nursing education. It included statements about the usefulness of AI, its impact on learning quality, the risk of dependence on technology, the reliability of AI-generated information, and its potential impact on critical thinking. Responses were recorded using a 5-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree").

Validity and Reliability

To evaluate the clarity, comprehensibility, and relevance of the questionnaire items in relation to the study objectives, the questionnaire was reviewed by 4 academic staff members with expertise in nursing education and research methodology. Their suggestions were considered, and minor modifications were made to the wording of several questions. Subsequently, a pilot test was conducted with 20 nursing students who had characteristics similar to those of the study population to assess comprehensibility, feasibility, and the time required to complete the questionnaire. On the basis of participants' feedback, minor revisions were made to improve the instrument. Data collected during the pilot study were not included in the final analysis.

The internal consistency of the questionnaire was evaluated using the Cronbach α coefficient. The scale assessing the perceived impact of AI on learning outcomes (10 items) demonstrated excellent reliability (α=0.937). The scale assessing perceptions of AI use (6 items) demonstrated good reliability (α=0.810). These findings indicate satisfactory internal consistency for both scales, supporting their suitability for further statistical analyses.

Data Analysis

Data were analyzed using SPSS (version 23.0; IBM Corp). Statistical analysis was conducted in several steps. First, descriptive statistics were used to summarize students' sociodemographic and academic characteristics, as well as variables related to AI use. Categorical variables were presented as frequencies and percentages, whereas continuous variables were described using means and SDs. Variables measured using Likert scales were also analyzed descriptively and presented as frequencies and percentages for each response category, as well as means and SDs for each item.

The chi-square test was used to assess associations between AI use and categorical variables. Factors associated with AI use were further analyzed using logistic regression. Both crude odds ratios (CORs) and adjusted odds ratios (AORs), together with 95% CIs, were reported. A *P* value <.05 was considered statistically significant.

Results

General Characteristics

A total of 279 students enrolled in the General Nursing program participated in the study. Of these, 273 (97.8%) were female and 6 (2.2%) were male. The mean age of the participants was 22.4 (SD 5.7) years. More than half of the students (n=156, 55.9%) lived in urban areas, while 123 (44.1%) resided in rural areas. The majority of participants (n=198, 71%) were enrolled in the bachelor's program. Regarding academic performance, 123 (44.1%) students had a GPA of 7.0 to 7.9, followed by 78 (28%) with a GPA of 6.0 to 6.9. Additionally, 45 (16.1%) students reported a GPA of 8.0 to 8.9 and 12 (4.3%) had a high GPA of 9.0 to 10.0. Students with a GPA of <6.0 accounted for 7.5% (n=21) of the participants. Overall, 234 (83.9%) students reported using AI tools (Table 1).

XSL•FO
RenderX

Table . Sample characteristics of participants.

Variables and categories	Participants
Age (y), mean (SD)	22.4 (5.7)
Residence, n (%)[a]	
Urban	156 (55.9)
Rural	123 (44.1)
Sex, n (%)[a]	
Male	6 (2.2)
Female	273 (97.8)
Study cycle, n (%)[a]	
Bachelor's	198 (71)
Master's	81 (29)
Academic average (grade point average), n (%)[a]	
<6.0	21 (7.5)
6.0-6.9	78 (28)
7.0-7.9	123 (44.1)
8.0-8.9	45 (16.1)
9.0-10.0	12 (4.3)
AI use, n (%)[a]	
Yes	234 (83.9)
No	45 (16.1)

[a]N=279.

Use of AI

A high proportion of students (234/279, 83.9%) reported using AI. Among the 234 AI users, 105 (44.9%) reported daily use, while 69 (29.5%) used AI on a weekly basis. Less frequent use was reported by a smaller proportion of participants, with 20.5% (n=48) using AI rarely and 5.1% (n=12) using it once a month. Regarding the types of AI tools used, 183 (78.2%) participants reported using virtual assistants such as ChatGPT, DeepSeek, or Copilot. Educational platforms such as Learning Nurse were used by 63 (26.9%) students, while text-generation tools were reported by 36 (15.4%) participants. In addition, 21.8% (n=51) of participants reported using other AI tools. Regarding the purpose of AI use, the majority of students (n=216, 92.3%) used AI for information searching. Additionally, 41% (n=96) of participants used AI for studying and understanding lecture content, 16.7% (n=39) for academic assignments, and 12% (n=28) for understanding clinical cases (Table 2).

XSL•FO
RenderX

Table . Frequency, tools, and purpose of AI use.

Variables and categories	Participants, n (%)
Frequency of AI use[a]	
Every day	105 (44.9)
Every week	69 (29.5)
Every month	12 (5.1)
Rarely	48 (20.5)
Tools used[a,b]	
Virtual assistants (ChatGPT, DeepSeek, and Microsoft Copilot)	183 (78.2)
Educational platforms (Learning Nurse)	63 (26.9)
Text-generation tools (paraphrase tool)	36 (15.4)
Other	51 (21.8)
Purpose of AI use[a,b]	
Studying and learning lecture content	96 (41)
Academic assignments or essays	39 (16.7)
Understanding clinical cases	28 (12)
Information search	216 (92.3)

[a]N=234.

[b]Multiple-response questions; n summed in each section may be greater than N.

Factors Associated With the Use of AI

Bivariate logistic regression analysis showed that place of residence and study cycle were not significantly associated with the use of AI. Although students living in urban areas had higher odds of using AI than those living in rural areas, this association was not statistically significant (COR 1.56, 95% CI 0.82-2.96; *P*=.23). Similarly, no statistically significant difference was observed between bachelor's and master's students (COR 0.84, 95% CI 0.42-1.65; *P*=.74).

In contrast, academic performance (GPA) was significantly associated with AI use. Compared with students with a GPA <6.0, those with GPAs of 6.0 to 6.9, 7.0 to 7.9, and 8.0 to 8.9 had higher odds of using AI, whereas no statistically significant association was observed for students with a GPA of 9.0 to 10.0.

In the multivariable analysis, place of residence and study cycle remained statistically nonsignificant. After adjusting for the variables included in the model, GPA remained the only independent factor associated with AI use. Students with a GPA of 6.0 to 6.9 (AOR 5.55, 95% CI 1.53-20.14; *P*=.009) and those with a GPA of 8.0 to 8.9 (AOR 5.73, 95% CI 1.26-26.00; *P*=.02) were approximately 5 times more likely to use AI than the reference group (GPA<6.0). The association observed for the GPA category of 7.0 to 7.9 in the bivariate analysis was no longer statistically significant after adjustment (AOR 1.69, 95% CI 0.59-4.85; *P*=.33), suggesting a possible influence of confounding factors. Similarly, no statistically significant association was found between AI use and the GPA category of 9.0 to 10.0 (AOR 0.51, 95% CI 0.11-2.33; *P*=.38; Table 3).

XSL•FO
RenderX

Table . Factors associated with the use of AI (N=279).

Variables and categories	AI use, n (%)	No AI use, n (%)	Crude odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)	*P* value
Residence						
Rural (n=123)	99 (80.5)	24 (19.5)	1.00	—[a]	1.00	—
Urban (n=156)	135 (86.5)	21 (13.5)	1.56 (0.82-2.96)	.23	1.44 (0.71-2.92)	.31
Study cycle						
Bachelor's (n=195)	165 (84.6)	30 (15.4)	1.00	—	1.00	—
Master's (n=84)	69 (82.1)	15 (17.9)	0.84 (0.42-1.65)	.74	1.31 (0.61-2.84)	.49
Academic average (grade point average)						
<6.0 (n=21)	10 (47.6)	11 (52.4)	1.00	—	1.00	—
6.0-6.9 (n=78)	66 (84.6)	12 (15.4)	6.05 (2.11-17.36)	<.001	5.55 (1.53-20.14)	.009
7.0-7.9 (n=123)	111 (90.2)	12 (9.8)	10.18 (3.59-28.88)	<.001	1.69 (0.59-4.85)	.33
8.0-8.9 (n=45)	39 (86.7)	6 (13.3)	7.15 (2.13-24.06)	.002	5.73 (1.26-26.00)	.02
9.0-10.0 (n=12)	8 (66.7)	4 (33.3)	2.20 (0.50-9.61)	.29	0.51 (0.11-2.33)	.38

[a]Not applicable.

Students' Perceptions of the Impact of AI

Participants generally reported a low to moderate impact of AI on learning outcomes, with mean scores ranging from 1.90 to 2.95 on a 5-point Likert scale. The highest perceived impact was reported for understanding clinical cases (mean 2.95, SD 1.20), followed by exam preparation (mean 2.58, SD 1.02) and understanding lectures (mean 2.58, SD 1.10), suggesting that AI plays a primarily supportive role in cognitive and academic processes. AI was also perceived as helpful in saving time during studying (mean 2.49, SD 1.22).

The perceived impact was lower in psychosocial and professional dimensions, including academic performance, critical thinking, communication skills, and nursing care. The lowest mean scores were reported for empathy toward patients (mean 1.90, SD 1.12) and self-confidence (mean 1.97, SD 1.17; Table 4).

Table . Impact of AI on learning outcomes (n=234).

Items	1: no influence, n (%)	2: low influence, n (%)	3: moderate influence, n (%)	4: high influence, n (%)	5: very high influence, n (%)	Scores, mean (SD)[a]
Understanding lectures	52 (22.2)	49 (20.9)	84 (35.9)	43 (18.4)	6 (2.6)	2.58 (1.10)
Exam preparation	45 (19.2)	54 (23.1)	93 (39.7)	39 (16.7)	3 (1.3)	2.58 (1.02)
Academic performance	81 (34.6)	54 (23.1)	81 (34.6)	15 (6.4)	3 (1.3)	2.17 (1.02)
Time saving while studying	60 (25.6)	66 (28.2)	60 (25.6)	30 (12.8)	18 (7.7)	2.49 (1.22)
Understanding clinical cases	36 (15.4)	39 (16.7)	87 (37.2)	45 (19.2)	27 (11.5)	2.95 (1.20)
Nursing care	87 (37.2)	39 (16.7)	78 (33.3)	21 (9)	9 (3.8)	2.26 (1.16)
Self-confidence	117 (50)	42 (18)	48 (20.5)	18 (7.7)	9 (3.8)	1.97 (1.17)
Critical thinking	84 (35.9)	57 (24.4)	69 (29.5)	15 (6.4)	9 (3.8)	2.18 (1.10)
Communication skills	99 (42.3)	39 (16.7)	60 (25.6)	24 (10.3)	12 (5.1)	2.19 (1.23)
Empathy toward patients	126 (53.8)	33 (14.1)	54 (23.1)	15 (6.4)	6 (2.6)	1.90 (1.12)

[a]Higher scores indicate greater perceived impact.

XSL•FO
RenderX

Students' Perceptions of AI Use

Students' perceptions of the use of AI were generally characterized by a moderate level, with mean scores ranging from 2.45 to 3.18 on a 5-point Likert scale. The highest level of agreement was reported for concerns regarding the reliability of AI-generated information (mean 3.18, SD 1.22), reflecting a considerable degree of skepticism about its accuracy.

Overall, AI was perceived as a supportive tool for the learning process (mean 2.83, SD 1.12) and for improving the quality of studying (mean 2.77, SD 1.14), suggesting a generally neutral to slightly positive perception of its use in education. However, moderate concerns were also expressed regarding its impact on critical thinking (mean 2.80, SD 1.18) and the risk of developing a dependence on AI (mean 2.75, SD 1.28; Table 5).

Table . Students' perceptions of AI use (n=234).

Statements	1: strongly disagree, n (%)	2: disagree, n (%)	3: neutral, n (%)	4: agree, n (%)	5: strongly agree, n (%)	Scores, mean (SD)[a]
AI is a useful tool.	39 (16.7)	35 (15)	102 (43.6)	40 (17.1)	18 (7.7)	2.83 (1.12)
AI improves the quality of learning.	39 (16.7)	45 (19.2)	84 (35.9)	51 (21.8)	15 (6.4)	2.77 (1.14)
AI helps in preparation for professional practice.	54 (23.1)	75 (32.1)	78 (33.3)	15 (6.4)	12 (5.1)	2.45 (1.08)
AI negatively affects critical thinking.	39 (16.7)	61 (26.1)	72 (30.8)	39 (16.7)	23 (9.8)	2.80 (1.18)
AI may create dependency.	66 (28.2)	27 (11.5)	75 (32.1)	42 (17.9)	24 (10.3)	2.75 (1.28)
Information from AI is not always reliable.	33 (14.1)	6 (2.6)	84 (35.9)	60 (25.6)	51 (21.8)	3.18 (1.22)

[a]Higher scores indicate stronger agreement with the statement.

Discussion

Principal Findings

This study identified a high level of AI use among nursing students, with 83.9% (234/279) reporting that they used AI tools. This finding is consistent with previous studies reporting the widespread use of digital technologies in higher education [29]. However, daily use was reported by only a proportion of students, suggesting that AI is used primarily as a functional tool for specific academic tasks and information retrieval rather than as a continuously integrated component of the learning process [30].

The results showed that the most commonly used AI tools were virtual assistants, particularly ChatGPT and similar applications (78.2%), reflecting a global trend toward the use of AI as a learning support tool. This finding is consistent with the existing literature, which suggests that AI facilitates access to information and supports the understanding of complex academic concepts [31-34]. However, the fact that AI was used mainly for information searching rather than for developing clinical skills or critical thinking suggests that students perceive AI primarily as an informational resource rather than an educational tool [9].

Overall, students perceived the impact of AI on both academic and professional outcomes as low to moderate. The greatest perceived impact was reported for understanding clinical cases and preparing for examinations, whereas the lowest perceived impact was reported for empathy toward patients, self-confidence, communication skills, and critical thinking. These findings suggest that AI is primarily perceived as a supportive tool for accessing and organizing information, while its role in developing interpersonal and professional competencies remains limited.

The greater perceived impact on understanding clinical cases may be attributed to the ability of AI tools to provide structured explanations, practical examples, and simplified clinical information. Similarly, the moderate perceived impact on understanding lectures and preparing for examinations suggests that students use AI as a complementary learning resource rather than as a transformative factor in academic performance [11,35,36].

Conversely, the limited perceived impact on critical thinking may reflect students' uncertainty regarding the role of AI in developing analytical reasoning and reflective skills. Likewise, the low perceived impact on empathy and self-confidence suggests that students do not consider AI to be an important contributor to their professional development. These findings support the view that although AI can facilitate the learning process, it cannot replace the human dimensions of nursing practice, such as empathy, communication, and the nurse-patient relationship [37,38].

A considerable proportion of students also expressed concerns regarding the reliability of AI-generated information, the risk of dependence on AI, and its potential impact on critical thinking. These perceptions suggest that uncontrolled AI use may limit the development of clinical judgment and encourage reliance on ready-made answers without sufficient critical reflection [18,27].

XSL•FO
RenderX

At the same time, students reported positive perceptions regarding the benefits of AI in improving learning efficiency and facilitating access to information. This finding reflects a balance between the perceived benefits and risks of AI, with acceptance of the technology accompanied by awareness of its limitations [39-41]. These results suggest that nursing education is currently in a transitional phase of AI adoption, in which students recognize its potential while remaining cautious about its academic and professional implications.

Logistic regression analysis showed that academic performance (GPA) was the only independent factor associated with AI use. Students with GPAs ranging from 6.0 to 8.9 were more likely to use AI than those with lower GPAs. These findings suggest that students with better academic performance are more likely to integrate digital tools into their learning strategies, consistent with previous studies reporting a positive association between academic achievement and technology use in education [33,42].

However, the absence of a statistically significant association among students with the highest GPA (9.0-10.0) suggests that this relationship may not be linear. One possible explanation is that high-achieving students may be more selective and critical in their use of AI, preferring traditional learning methods. It should also be noted that the small number of participants in this subgroup may have reduced the statistical power of the analysis.

In contrast, place of residence was not significantly associated with AI use, suggesting relatively equal access to digital technologies among students from urban and rural areas. Similarly, no association was found between AI use and study cycle, indicating that AI use was comparable among both bachelor's and master's students. These findings suggest that AI use in this context is influenced more by individual factors, such as exposure to technology and personal attitudes toward innovation, than by the level of academic study [29,43].

Overall, the findings of this study indicate that AI is already widely integrated into the learning process of nursing students; however, its use remains primarily supportive and informational. Although students recognize its benefits for accessing and organizing information, important limitations remain regarding its use for developing clinical and professional competencies. These findings highlight the need for a more structured integration of AI into nursing curricula through educational approaches such as simulation-based learning, virtual clinical cases, and supervised AI-assisted learning activities, with the aim of strengthening students' critical thinking, clinical reasoning, and professional competencies.

Study Limitations

This study has several limitations. First, as the data were based on self-reported responses, the findings may be subject to reporting bias and subjectivity in students' perceptions and evaluations of the impact of AI. Furthermore, the use of an online questionnaire may have favored participation by students who were more familiar with or interested in technology. In addition, the cross-sectional design and the conduct of the study at a single institution did not allow for the assessment of causal relationships, thereby limiting the generalizability of the findings. Therefore, future studies should use more comprehensive research approaches, including qualitative methods such as interviews, to gain a deeper understanding of nursing students' perceptions and experiences regarding the use of AI. Moreover, multicenter studies with larger and more diverse samples are recommended to improve the representativeness and generalizability of the findings.

Conclusions

The findings of this study showed that nursing students had a generally positive attitude toward the use of AI and perceived it as a supportive tool in the learning process. However, its impact was greater on theoretical aspects, whereas its impact on their professional and interpersonal skills was more limited. These results emphasize the need for a more advanced integration of AI into nursing education. Training, experience, the development of digital competencies, and a supportive academic environment are essential for these tools to be used effectively.

Acknowledgments

The authors thank all nursing students from the Faculty of Medical Technical Sciences in Elbasan who participated voluntarily in this study and contributed to the data collection.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The data generated and/or analyzed in this study are available from the corresponding author upon reasonable request. The data are not publicly available to preserve the privacy and confidentiality of the participants.

Conflicts of Interest

None declared.

Checklist 1
CHERRIES checklist.

XSL•FO
RenderX

[DOCX File, 16 KB - nursing_v9i1e97832_app1.docx]

References

1. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on nursing education: scoping review. JMIR Nurs 2021;4(1):e23933. [doi: 10.2196/23933] [Medline: 34345794]
2. Labrague LJ, AL Sabei S, AL Yahyaei A. Artificial intelligence in nursing education: a review of AI-based teaching pedagogies. Teach Learn Nurs 2025 Jul;20(3):210-221. [doi: 10.1016/j.teln.2025.01.019]
3. Rony MK, Ahmad S, Das DC, et al. Nursing students' perspectives on integrating artificial intelligence into clinical practice and training: a qualitative descriptive study. Health Sci Rep 2025 Apr;8(4):e70728. [doi: 10.1002/hsr2.70728] [Medline: 40260042]
4. Fernandes F, Shinners L, Mota M, Santos P, Sá L. Contributions of artificial intelligence to decision making in nursing: a scoping review. Nurs Health Sci 2026 Mar;28(1):e70308. [doi: 10.1111/nhs.70308] [Medline: 41709566]
5. Cucci F, Marasciulo D, Romani M, et al. The contribution of artificial intelligence in nursing education: a scoping review of the literature. Nurs Rep 2025 Aug 1;15(8):283. [doi: 10.3390/nursrep15080283] [Medline: 40863670]
6. Fung TC, Chan SL, Lam CF, et al. Effects of generative artificial intelligence (GenAI) patient simulation on perceived clinical competency among global nursing undergraduates: a cross-over randomised controlled trial. BMC Nurs 2025 Jul 17;24(1):934. [doi: 10.1186/s12912-025-03492-0] [Medline: 40676632]
7. Sriram A, Ramachandran K, Krishnamoorthy S. Artificial intelligence in medical education: transforming learning and practice. Cureus 2025 Mar;17(3):e80852. [doi: 10.7759/cureus.80852] [Medline: 40255837]
8. Wei Q, Pan S, Liu X, Hong M, Nong C, Zhang W. The integration of AI in nursing: addressing current applications, challenges, and future directions. Front Med (Lausanne) 2025;12:1545420. [doi: 10.3389/fmed.2025.1545420] [Medline: 40007584]
9. Zgambo M, Costello M, Buhlmann M, Maldon J, Anyango E, Adama E. Artificial intelligence and academic integrity in nursing education: a mixed methods study on usage, perceptions, and institutional implications. Nurse Educ Today 2025 Oct;153:106796. [doi: 10.1016/j.nedt.2025.106796] [Medline: 40517666]
10. Guirguis M, Fotsing S, Fevry J, et al. Artificial intelligence in health professions education: qualitative study of student experiences. J Med Internet Res 2026 Apr 2;28:e82432. [doi: 10.2196/82432] [Medline: 41926317]
11. Khlaif ZN, Salameh N, Ajouz M, et al. Using generative AI in nursing education: students' perceptions. BMC Med Educ 2025 Jul 1;25(1):926. [doi: 10.1186/s12909-025-07416-z] [Medline: 40598139]
12. Lifshits I, Rosenberg D. Artificial intelligence in nursing education: a scoping review. Nurse Educ Pract 2024 Oct;80:104148. [doi: 10.1016/j.nepr.2024.104148] [Medline: 39405792]
13. George Pallivathukal R, Kyaw Soe HH, Donald PM, Samson RS, Hj Ismail AR. ChatGPT for academic purposes: survey among undergraduate healthcare students in Malaysia. Cureus 2024 Jan;16(1):e53032. [doi: 10.7759/cureus.53032] [Medline: 38410331]
14. Labrague LJ, Aguilar-Rosales R, Yboa BC, Sabio JB, de Los Santos JA. Student nurses' attitudes, perceived utilization, and intention to adopt artificial intelligence (AI) technology in nursing practice: a cross-sectional study. Nurse Educ Pract 2023 Nov;73:103815. [doi: 10.1016/j.nepr.2023.103815] [Medline: 37922736]
15. Glauberman G, Ito-Fujita A, Katz S, Callahan J. Artificial intelligence in nursing education: opportunities and challenges. Hawaii J Health Soc Welf 2023 Dec;82(12):302-305. [Medline: 38093763]
16. Lane SH, Haley T, Brackney DE. Tool or tyrant: guiding and guarding generative artificial intelligence use in nursing education. Creat Nurs 2024 May;30(2):125-132. [doi: 10.1177/10784535241247094] [Medline: 38651267]
17. Martinez-Ortigosa A, Martinez-Granados A, Gil-Hernández E, Rodriguez-Arrastia M, Ropero-Padilla C, Roman P. Applications of artificial intelligence in nursing care: a systematic review. J Nurs Manag 2023;2023:3219127. [doi: 10.1155/2023/3219127] [Medline: 40225652]
18. Dagcan Sahin N, Yildirim M. Nursing students' experiences with artificial intelligence: a qualitative study on education, clinical practice, and future expectations. J Eval Clin Pract 2026 Mar;32(2):e70413. [doi: 10.1111/jep.70413] [Medline: 41833057]
19. Pailaha AD. The impact and issues of artificial intelligence in nursing science and healthcare settings. SAGE Open Nurs 2023;9:23779608231196847. [doi: 10.1177/23779608231196847] [Medline: 37691725]
20. Sun GH, Hoelscher SH. The ChatGPT storm and what faculty can do. Nurse Educ 2023;48(3):119-124. [doi: 10.1097/NNE.0000000000001390] [Medline: 37043716]
21. Le Lagadec D, Jackson D, Cleary M. Artificial intelligence in nursing education: prospects and pitfalls. J Adv Nurs 2024 Oct;80(10):3883-3885. [doi: 10.1111/jan.16276] [Medline: 38864369]
22. Lund B, Mannuru NR, Teel ZA, et al. Student perceptions of AI-assisted writing and academic integrity: ethical concerns, academic misconduct, and use of generative AI in higher education. AI Educ 2025 Sep 2;1(1):2. [doi: 10.3390/aieduc1010002]
23. Wynn MO. Beyond competency: developing critical digital capabilities in nursing students through Freirean pedagogy. Nurs Inq 2025 Apr;32(2):e70011. [doi: 10.1111/nin.70011] [Medline: 40077858]
24. Montemayor C, Halpern J, Fairweather A. In principle obstacles for empathic AI: why we can't replace human empathy in healthcare. AI Soc 2022;37(4):1353-1359. [doi: 10.1007/s00146-021-01230-z] [Medline: 34054228]

XSL•FO
RenderX

25. El Arab RA, Alshakihs AH, Alabdulwahab SH, et al. Artificial intelligence in nursing: a systematic review of attitudes, literacy, readiness, and adoption intentions among nursing students and practicing nurses. Front Digit Health 2025;7:1666005. [doi: 10.3389/fdgth.2025.1666005] [Medline: 41079691]
26. Abuadas M, Albikawi Z, Rayani A. The impact of an AI-focused ethics education program on nursing students' ethical awareness, moral sensitivity, attitudes, and generative AI adoption intention: a quasi-experimental study. BMC Nurs 2025 Jul 1;24(1):720. [doi: 10.1186/s12912-025-03458-2] [Medline: 40597065]
27. Dodson TM, Thompson-Hairston K, Reed JM. Nursing students' AI literacy and ethical understanding of AI in nursing education. Teach Learn Nurs 2025 Oct;20(4):390-394. [doi: 10.1016/j.teln.2025.07.004]
28. Kodi i etikës i Universitetit të Elbasanit "Aleksandër Xhuvani" [Article in Albanian]. University of Elbasan "Aleksandër Xhuvani.". 2020. URL: https://uniel.edu.al/images/35_Kodi_i_Etikes_UE.pdf [accessed 2026-08-27]
29. Han S, Kang HS, Gimber P, Lim S. Nursing students' perceptions and use of generative artificial intelligence in nursing education. Nurs Rep 2025 Feb 14;15(2):68. [doi: 10.3390/nursrep15020068] [Medline: 39997804]
30. Johnston H, Wells RF, Shanks EM, Boey T, Parsons BN. Student perspectives on the use of generative artificial intelligence technologies in higher education. Int J Educ Integr 2024;20:2. [doi: 10.1007/s40979-024-00149-4]
31. Arkan B, Dallı Ö, Varol B. The impact of ChatGPT training in the nursing process on nursing students' problem-solving skills, attitudes towards artificial intelligence, competency, and satisfaction levels: single-blind randomized controlled study. Nurse Educ Today 2025 Sep;152:106765. [doi: 10.1016/j.nedt.2025.106765] [Medline: 40334550]
32. Bai˙doo-anu D, Owusu Ansah L. Education in the era of generative artificial intelligence (AI): understanding the potential benefits of ChatGPT in promoting teaching and learning. J AI 2023;7(1):52-62. [doi: 10.61969/jai.1337500]
33. Gonzalez-Garcia A, Bermejo-Martinez D, Lopez-Alonso AI, Trevisson-Redondo B, Martín-Vázquez C, Perez-Gonzalez S. Impact of ChatGPT usage on nursing students education: a cross-sectional study. Heliyon 2024;11(1):e41559. [doi: 10.1016/j.heliyon.2024.e41559] [Medline: 39850430]
34. Saatçi G, Korkut S, Ünsal A. The effect of the use of artificial intelligence in the preparation of patient education materials by nursing students on the understandability, actionability and quality of the material: a randomized controlled trial. Nurse Educ Pract 2024 Nov;81:104186. [doi: 10.1016/j.nepr.2024.104186] [Medline: 39520840]
35. Shen M, Shen Y, Liu F, Jin J. Prompts, privacy, and personalized learning: integrating AI into nursing education-a qualitative study. BMC Nurs 2025 Apr 29;24(1):470. [doi: 10.1186/s12912-025-03115-8] [Medline: 40301862]
36. Vieriu AM, Petrea G. The impact of artificial intelligence (AI) on students' academic development. Educ Sci 2025 Mar 11;15(3):343. [doi: 10.3390/educsci15030343]
37. Ronquillo CE, Peltonen LM, Pruinelli L, et al. Artificial intelligence in nursing: priorities and opportunities from an international invitational think-tank of the Nursing and Artificial Intelligence Leadership Collaborative. J Adv Nurs 2021 Sep;77(9):3707-3717. [doi: 10.1111/jan.14855] [Medline: 34003504]
38. Gülırmak Güler K, Şen Atasayar B. The relationship between nursing students' attitudes toward artificial intelligence and their creative personality traits. Int Nurs Rev 2025 Mar;72(1):e70008. [doi: 10.1111/inr.70008] [Medline: 40070203]
39. Farghaly Abdelaliem SM, Dator WL, Sankarapandian C. The relationship between nursing students' smart devices addiction and their perception of artificial intelligence. Healthcare (Basel) 2022 Dec 30;11(1):110. [doi: 10.3390/healthcare11010110] [Medline: 36611570]
40. Mersal FA, Mersal NA, Ibrahim NM, et al. Nursing students' trust in artificial intelligence (AI) clinical recommendations: a multicenter cross-sectional study of risk-benefit perceptions across Saudi Arabian universities. Digit Health 2026;12:20552076261429671. [doi: 10.1177/20552076261429671] [Medline: 41783285]
41. Thomson SR, Pickard-Jones BA, Baines S, Otermans PC. The impact of AI on education and careers: what do students think? Front Artif Intell 2024;7:1457299. [doi: 10.3389/frai.2024.1457299] [Medline: 39610850]
42. Bakarman SS, Al-Shammari A, Aboshaiqah A. Nursing students' perception of and readiness for artificial intelligence in Saudi Arabia. Nurs Open 2025 Dec;12(12):e70386. [doi: 10.1002/nop2.70386] [Medline: 41329788]
43. Zeng Q, Zhu J, Wang Y, Su S, Huang Y. Artificial intelligence self-efficacy and attitudes among nursing students: a multicenter network analysis of educational stratification. BMC Med Educ 2026 Jan 26;26(1):299. [doi: 10.1186/s12909-026-08685-y] [Medline: 41588383]

Abbreviations

AOR: adjusted odds ratio
COR: crude odds ratio
GPA: grade point average

XSL•FO
RenderX

Edited by E Borycki; submitted 10.Apr.2026; peer-reviewed by KL Ajee, S Llewellyn; revised version received 01.Jul.2026; accepted 28.Jul.2026; published 31.Aug.2026.

Please cite as:
Taka I, Hasalla E, Sula A, Bahiti B, Oboni R, Bani B
AI Use, Perceptions, and Perceived Impact Among Nursing Students: Cross-Sectional Study
JMIR Nursing 2026;9:e97832
URL: https://nursing.jmir.org/2026/1/e97832
doi:10.2196/97832

© Ilda Taka, Elona Hasalla, Albana Sula, Blerina Bahiti, Rajmonda Oboni, Blerina Bani. Originally published in JMIR Nursing (https://nursing.jmir.org), 31.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Social Media Listening to Understand the Pressure Injury Experience: A Qualitative Descriptive Study

Jessica Zihui Song[1,2], BMSc, MSc; Isabel Tsang[1], BMSc, MBBS; Parthena Mouratidis[3], BA; Manahil Saeed[3], BSc; Alma A Ornelas L[3], MD; Alaul Islam[1], PhD; Sharon Gabison[1,2,4], BSc, BScPT, MSc, PhD, IIWCC

[1]KITE Research Institute, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada

[2]Rehabilitation Sciences Institute, Temerty Faculty of Medicine, University of Toronto, 500 University Avenue, Toronto, ON, Canada

[3]Faculty of Liberal Arts and Science, Humber Polytechnic, Toronto, ON, Canada

[4]Department of Physical Therapy, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Jessica Zihui Song, BMSc, MSc
KITE Research Institute, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada

Abstract

Background: Pressure injuries (PIs) are a common complication in people with reduced mobility or sensation and can be burdensome for individuals with PIs and their caregivers. Valuable insights and real-world challenges faced by individuals living with PIs can be captured through candid accounts posted on social media. Social media listening (SML) is a tool that can enhance the understanding of those with lived experience by offering firsthand accounts that are irreproducible from controlled studies.

Objective: This study aims to capture the candid experiences of individuals with PIs and caregivers through social media.

Methods: A noninterventional qualitative descriptive analysis was conducted using SML. Social media posts made on X (formerly Twitter), Reddit, and YouTube between January and December 2022 were compiled using SML tools X Pro (formerly TweetDeck) and Awario, and using Boolean search terms. Posts were manually screened for relevance, and duplicates were removed. Relevant posts were hand-coded by two independent reviewers. Inductive content analysis was used to analyze the posts.

Results: The search yielded 666 relevant posts from 498 unique social media users. Most posts were made in the United States (170/666, 25.5%), the United Kingdom (150/666, 22.5%), and Canada (62/666, 9.3%). Social media users provided detailed descriptions of the PIs, including the setting in which the PI occurred, the cause of the PI, and how the PI was managed. The majority of PIs (197/666, 29.6%) were reported to have occurred in the hospital setting due to a perceived lack of care from care providers, and local wound care was often cited (99/666, 14.9%) as a PI management strategy. Three key themes were developed regarding living with or caring for someone with a PI: (1) challenges experienced when living with or caring for a PI, (2) needs related to PI prevention and management, and (3) emotions experienced when living with or caring for a PI. Social media users frequently discussed challenges associated with living with a PI, including negative personal impacts and poor perceived treatment quality. Users also described a critical need for health care, education, and social support. Finally, users often expressed anger and/or sadness related to living with or caring for a PI.

Conclusions: SML captured candid insights into the experiences, challenges, and needs of individuals living with PIs and their caregivers globally that may not be gleaned from controlled studies. Individuals with lived experience and their caregivers often struggled with negative personal impacts regarding their physical health and daily functioning related to PIs, further highlighting the urgent need to address barriers to appropriate PI care. Clinicians and policymakers should consider practices and policies that optimize the delivery of person-centered PI care in order to overcome challenges and needs identified in this study.

(JMIR Nursing 2026;9:e76682) doi:10.2196/76682

KEYWORDS

social media listening; social media; pressure injury; pressure ulcer; bedsore; caregiver

Introduction

Pressure injuries (PIs) occur due to damage to the integrity of the skin and underlying soft tissue resulting from severe and persistent localized pressure and shear, usually over a bony prominence [1-3]. Those at risk of developing PIs include older individuals and individuals with reduced mobility or sensation [1,3]. The global prevalence rate of PIs has remained stagnant at around 12% over the past few decades [4,5], which is notable because most PIs are avoidable if appropriate prevention techniques are applied [6]. This prevalence is especially troubling since PIs impose severe physical, mental, social, and

XSL•FO
RenderX

emotional impacts on those living with or caring for them [7-11]. Individuals living with PIs often experience physical discomfort, social isolation, and challenges when performing daily activities [8]. Similarly, caregivers of individuals with PIs may experience a high caregiving burden and a low quality of life [12,13]. PIs also cause tremendous financial strain on affected individuals and the US health care system, with an estimated yearly expenditure of US $26.8 billion on hospital-acquired PIs [14].

The burdens and challenges of living with a PI may go unrecognized or unaddressed by the health care teams caring for individuals with PIs [8,11]. Patient-centered care, which incorporates patients' and caregivers' perspectives, preferences, needs, and goals, is essential for addressing patients' challenges and improving treatment adherence and patient outcomes [15,16]. An interventional study of hospitalized patients admitted for acute medical or surgical treatment comparing usual care with patient-centered interventions revealed that a patient-centered educational intervention resulted in greater confidence in medication adherence [17]. Similarly, an interventional study conducted in an inpatient ward specializing in care for patients with neurological, orthopedic, or musculoskeletal conditions revealed that the active involvement of patients and their colleagues during nursing clinical handovers resulted in a decrease in hospital-acquired complications [18]. Thus, understanding the perspectives of individuals living with PIs and their caregivers may provide insights into improving the delivery of patient-centered PI care [15,16].

Over the past few decades, social media has become a vital resource where communities of individuals living with health conditions and their caregivers from around the world share unfiltered opinions, seek health information, and connect with one another [19]. Social media platforms such as YouTube, X (formerly Twitter), and Reddit are popular forums for individuals to seek social support and health information or share their personal experiences with their health conditions and subsequent treatment [20]. The increasing use of social media by patient and caregiver communities presents a valuable opportunity for researchers to elucidate candid patient and caregiver perspectives and experiences as reported on social media sites [19]. This is made possible through social media listening (SML) [21,22], a research method that involves performing a systematic search of public social media platforms. SML uses tools such as Talkwalker, Salesforce Social Studio [23,24], and X Pro (formerly TweetDeck) to identify real-world, firsthand accounts of individuals' experiences with disease and treatment, quality of life, and unmet needs [23,25].

SML can be used as a complementary research method alongside traditional qualitative or quantitative studies and has been shown to yield several benefits [19]. Given that patient recruitment is not required, SML can reduce recall and reporting biases that may be present in interviews or retrospective data collection methods by removing the researcher from the discussion [19]. SML also allows researchers to collect data from patients without geographic restrictions, providing large-scale insight into a topic [19]. SML has been shown to be an effective method of understanding patient and caregiver perspectives, unmet medical needs, and adherence to treatment that may not have been previously identified in the literature [23-28]. For instance, Perić et al [23] explored the needs and lived experiences of patients with graft-versus-host disease by using Talkwalker to search for relevant posts on Twitter, Facebook, Instagram, and YouTube. Similarly, Kline et al [24] used Salesforce Social Studio and Talkwalker to understand unmet needs, barriers to treatment, patient journey, and treatment options for patients living with amblyopia. Other similar studies have investigated individuals' experiences with various medical conditions, providing critical insights into individuals' quality of life, treatment, and perception of their quality of care [27-30].

While previous studies have explored the challenges faced by individuals living with PIs and caregivers who provide support [8,11,13,31], their perceptions of their role in PI care [9,16,32], as well as their knowledge, information needs, and preferences for PI education [15,33-35], none have used an SML approach to explore their experiences. Given the benefits of SML and the pervasive use of social media in the present day, SML may provide important, unique, and unfiltered insights into the experiences, challenges, and needs of individuals with PIs and caregivers globally that may not be elucidated from traditional studies. Thus, our study aimed to capture the experiences, challenges, and needs of individuals with PIs and caregivers using SML, an emerging research method that could reveal aspects of the PI experience not easily observed in controlled research environments.

Methods

Study Design

This study implemented a noninterventional qualitative descriptive approach to analyze publicly available social media posts made on X (formerly Twitter), Reddit, and YouTube using SML software Awario and X Pro. The study design was conceived by postgraduate students and scientists with extensive experience in both qualitative and quantitative research. We used the SRQR (Standards for Reporting Qualitative Research) guideline [36] to draft this manuscript and the SRQR reporting checklist [37] when editing (Checklist 1).

Search Strategy

Overview

Between May and June 2023, comprehensive searches were conducted on Reddit and YouTube using Awario. Searches were simultaneously conducted on X using X Pro. Awario and X Pro enable users to search web-based platforms for phrases and keywords in public posts in a similar manner to a literature search [38,39]. Both Awario and X Pro identified English-language posts made between January 1 and December 31, 2022. X, Reddit, and YouTube were selected as popular open-access social media websites where users commonly share personal opinions and experiences, enabling us to feature the perspectives of a wide range of individuals. Predefined Boolean search strings were used to identify relevant posts on each social media platform, and Boolean operators (AND/OR) were used to combine each keyword within the strings (Multimedia Appendix 1). The search string included key terms related to pressure injury and alternate words for pressure injury, such as

XSL•FO
RenderX

"bedsore" and "pressure ulcer," as well as care recipients' or caregivers' identities (ie, grandparents, parents, sister).

Definitions

A "user" or "social media user" referred to individuals with lived experience (ie, individuals who had experience living with a PI), caregivers (ie, individuals who provided direct, unpaid care to someone with a PI), or observers (ie, individuals who were not explicitly a caregiver but had witnessed someone else's PI experience as their friend or relative) who posted content on social media. Observers were distinguished from caregivers by their level of participation in providing PI care: caregivers explicitly provided care for someone with a PI, whereas an observer did not provide direct care for an individual with a PI but appeared to be actively involved in the life of an individual with a PI (eg, friend, colleague, or family member). A "post" was defined as any social media content identified through the search. A "mention" was defined as a reference to a specific topic within posts, which was used for coding in the analysis.

Inclusion and Exclusion Criteria

Social media posts from X, Reddit, and YouTube were included for analysis if they were written or recorded in English by individuals with lived PI experience, caregivers, or observers. In order to be included in the study, the social media post had to contain content relating to personal experiences with PIs (ie, experiences of living with a PI, caring for someone with a PI, or observations of another person's experience with PIs). There were no restrictions in terms of the user's geographic location.

Posts were excluded if they mentioned the development of a PI from medical masks or if the user did not fit the definition of an individual with lived PI experience, caregiver, or observer. Posts were also excluded if the user was a paid caregiver or health care professional (HCP) and if the post contained product endorsements or advertisements.

Data Screening

The results of each Boolean search were entered into a Microsoft Excel spreadsheet. Each X, Reddit, and verbatim transcript of the YouTube post was manually reviewed for relevance based on the inclusion and exclusion criteria before being input into a separate spreadsheet for content analysis. Any posts that did not meet the inclusion criteria or the exclusion criteria (eg, product endorsements) were removed. Duplicate posts that contained identical text were identified using the sort function of the spreadsheet and removed manually by researchers. To ensure interplatform data consistency, we applied common inclusion and exclusion criteria across platforms. Screening and verbatim transcription were performed in duplicate by two independent researchers to enhance trustworthiness.

Data Collection: Post Characteristics

Users' relationships with individuals living with a PI, geographic location, and intended audience were documented whenever such information was available from the users' social media profiles or could be inferred from the posts. Otherwise, users were classified as "unknown." Post details including the posting date, retrieval date, source platform (ie, X, Reddit, or YouTube), post format (ie, video, forum post, or tweet), and post hyperlinks were also documented on a Microsoft Excel spreadsheet.

Summative Content Analysis

Overview

Segments of text within each social media post were assigned a label (ie, code). Given that multiple segments of text within each social media post may have been present, social media posts were often assigned more than one code. Summative content analysis was used [40,41] through inductive coding; theme generation; and code, subcode, and theme frequency measurement.

Theme Generation

Inductive coding was used to formulate a comprehensive codebook. Primary codes, corresponding subcodes, example posts for reference, and code definitions were determined through an iterative process. The initial creation of codes and subcodes was performed independently by four research team members (PM, Pilar Tabuenca [PT], Uchechukwu Akunna [UA], and SG) based on post content, and the final codebook was developed following an exhaustive discussion by the team and the refinement of initial codes and subcodes. Disagreements were resolved through a discussion between coders or by an additional coder, if needed. Codes and subcodes were assigned to the relevant posts on a Microsoft Excel spreadsheet. Emotion codes were assigned to each post based on Plutchik's emotion wheel [42], which identifies eight primary emotion codes—joy, trust, fear, surprise, sadness, disgust, anger, and anticipation—along with their combinations. To ensure the reliability of the coding, these code assignments were later double-checked by four independent researchers (AI, AAO, IT, and JZS) and input into NVivo 12 (QSR International), a qualitative data management software.

The codes and subcodes were then consolidated into major themes representing the experiences of individuals with lived PI experience as reported on social media following further discussion by the research team. Posts were double coded under the same theme when they reflected multiple aspects of the theme; for instance, posts often described both personal impacts and perceived quality of treatment that were encountered while caring for a PI, which both fell under the theme of "Challenges experienced when living with or caring for a PI." As a result, the number of mentions (ie, codes) per theme surpassed the number of total posts in some cases.

Frequency Measurement

Code, subcode, and theme frequencies were calculated using NVivo 12. Frequency was similarly calculated and summarized for users' demographic information (ie, geographic location, intended audience, and relationship to the individual living with a PI).

Ethical Considerations

This study was exempt from human subject research ethics review because all data were obtained from public, internet-based sources, as per University Health Network [43] and University of Toronto [44] policies. No human participants were directly involved in this study, and as such, compensation

XSL•FO
RenderX

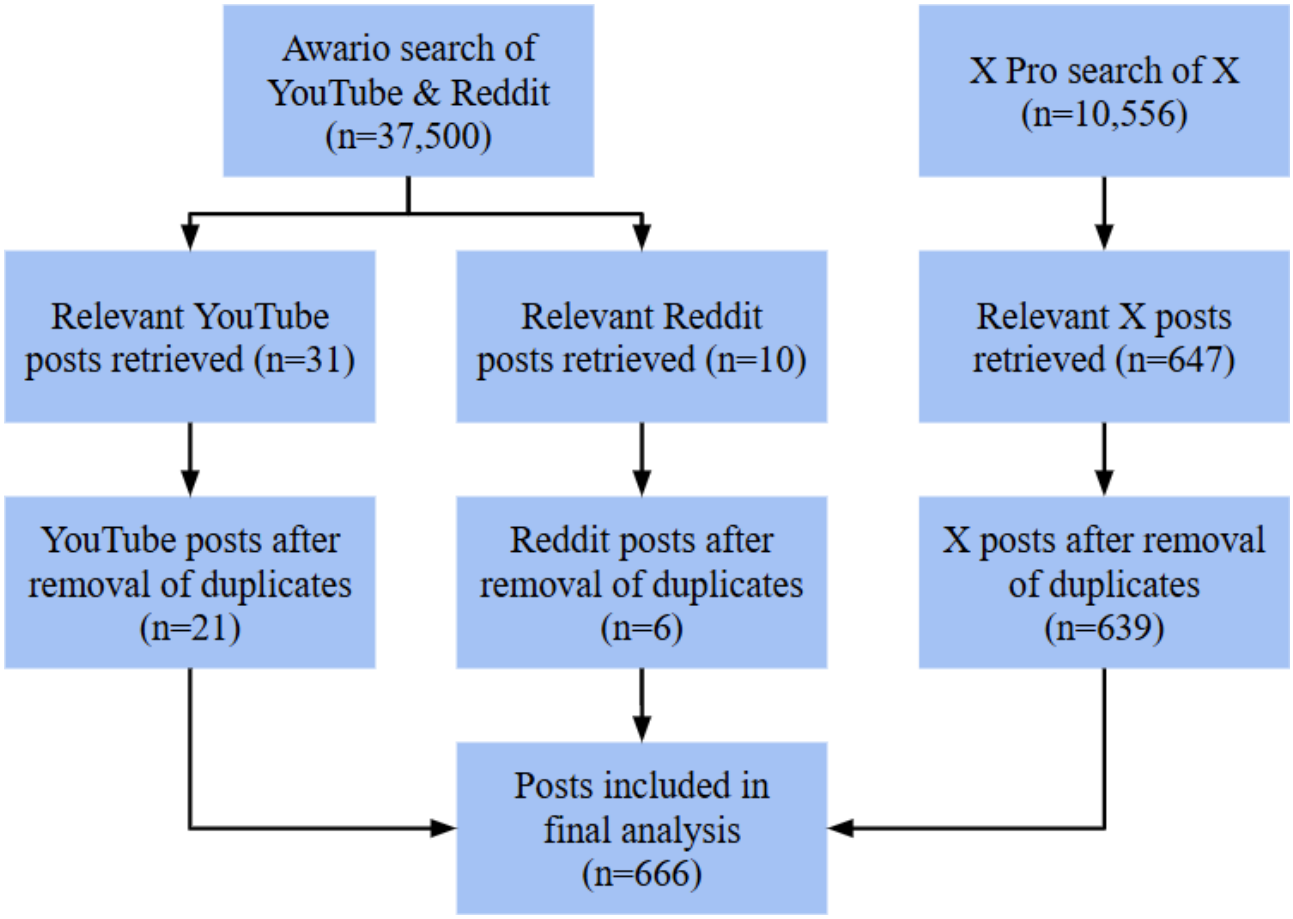
was not provided. Study data were deidentified, and any identifiable information was removed from the study analysis, manuscript, figures, and tables.

Results

Search Strategy and Data Screening

The initial search retrieved 37,500 and 10,556 posts from Awario and X Pro, respectively (Figure 1). After three independent researchers (PM, PT, UA) manually screened the posts, 41 relevant posts were identified from Awario, and 647 relevant posts were identified from X Pro. Duplicate posts were then removed. Ultimately, 639 X posts, 6 Reddit posts, and 21 YouTube videos were identified as relevant to the study objectives.

Figure 1. Flow diagram depicting the process of selecting relevant English-language social media posts made between January 1 and December 31, 2022, using social media listening (SML) software programs Awario and X Pro.

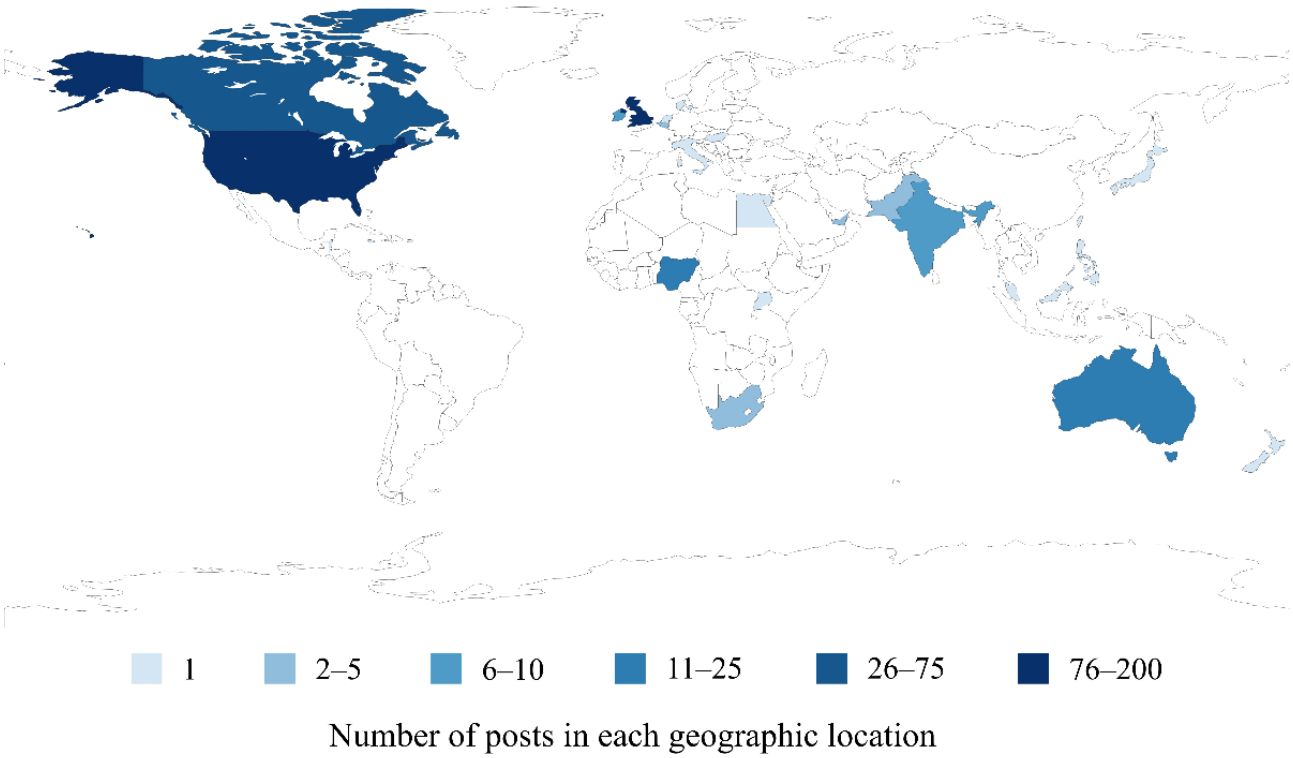


Post Characteristics

Our study identified 498 unique social media users across X, YouTube, and Reddit who generated a total of 666 posts. Of the 666 posts, 361 (54.2%) were posted by observers, 219 (32.9%) were posted by caregivers, and 86 (12.9%) were posted by individuals with lived PI experience. Caregivers and observers were often the child or other relative (eg, grandchild, sibling, parent) of the person with a PI. Most posts (n=631, 94.7%) were intended for the general public, although a few were targeted at specific health care institutions or other individuals (ie, governments, support groups). Most posts (n=481, 72.2%) were intended to share information about PIs with others (eg, warnings, knowledge about PI progression, causes, and complications), while some posts were intended to blame or advocate for others. Posts most frequently originated from the United States (n=170, 25.5%), the United Kingdom (n=150, 22.5%), or Canada (n=62, 9.3%), as seen in Figure 2.

XSL•FO
RenderX

Figure 2. Heat map displaying the distribution of social media posts according to geographic location.



Overview: The Pressure Injury Experience

Table 1 displays characteristics related to the PI experience, including the description of the PI, risk factors and causes of PIs, PI symptoms, the setting where the PI occurred, PI management, and commentary about PI healing. In posts (n=666), users frequently discussed the severity of the PI (n=120, 18%) and the location of the PI on the body (n=116, 17.4%). Several users described comorbidities (eg, dementia, obesity) as a risk factor for developing a PI (n=105, 15.8%), and many users cited a lack of care from HCPs or caregivers as the direct cause of their PI (n=156, 23.4%). Additionally, pain and discomfort were commonly reported symptoms (n=56, 8.4%). Many PIs were reported to have been hospital-acquired (n=197, 29.6%), and the most commonly reported PI management strategy was local wound care (n=99, 14.9%; eg, debridement, cleaning). A few users described the healing process of their PI (n=41, 6.2%).

XSL•FO
RenderX

Table . Frequency of posts that reported elements of the pressure injury (PI) experience, including PI description, risk factors, causes, symptoms, setting, management, and healing[a].

Topic	Posts, n	Relevant quotes
Pressure injury description		
Severity	120	"… my back, bottom & legs were a mass of pressure sores, two so severe that you could fit a large man's fist into the holes."
Location on body	116	"My Mum had a bedsore at the bottom of her spine right through to the spine. It killed her …"
Size	50	"… she had heel pressure ulcers the size of a silver dollar and deep ..."
Appearance	36	"… her bottom became an un-stageable bedsore. After surgery to remove [the] blackened and dead flesh, you could actually see her lower spine."
Number	34	"I'm alive, but I am dealing with three separate pressure ulcers …"
Duration	33	"Almost seven years of dealing with that pressure sore and it's finally healed …"
Risk factors of pressure injury		
Comorbidities	105	"She keeps getting UTIs and now she has a pressure sore that's slow to heal. She needs a hoist to get up. Her dementia is getting worse ..."
Immobility	50	"… She was essentially immobile, and fat. It took three hospital staff every time she needed to be repositioned to prevent bedsores."
COVID	15	"My mom got COVID… A few months later, she got an infection from a bedsore, got sepsis, and passed away. I believe COVID weakened her immune system …"
Body mass	15	"[He] was about 1200 pounds, bedridden, and ended up having a huge wound as a result of a bedsore."
Thin skin	7	"My mom went in hospital and got pressure ulcers like that within days ... Her skin was very thin."
Cause of pressure injury		
Lack of care	156	"She has two pressure sores on one foot as the carers in the home didn't move her when she was out of it for a couple of days."
Immobility	103	"... [My grandpa was] completely paralyzed. Ended up dying from a bedsore because of it."
Comorbidities	63	"…due to my spin a (sic) bifida and lack of sensation in some areas of my body I developed pressure ulcers on my sacrum and was hospitalized for wound care and sepsis."
Equipment-related	39	"I got a pressure sore on the back of my heel from my cast …"
Symptoms related to pressure injury		
Pain or discomfort	56	"My gran was constantly crying due to pressure ulcers …"
Bleeding or exudate	12	"… We are bringing my dad to see someone about some bedsores we are worried about because they're starting to bleed."
Odor	3	"… I could smell the infection as soon as I entered the room."

XSL•FO
RenderX

Topic	Posts, n	Relevant quotes
Setting where pressure injury occurred		
Hospital (ie, ICU[b], ER[c])	197	"... I've seen how they treated my mother in the hospitals ... She came home with a bedsore."
Care home	64	"Just like my loved one in LTC a couple of years back … She was yelling and crying out and the staff dismissed it as a regular dementia behavior … She ended up with a terrible pressure sore …"
Home	63	"Got my first pretty bad pressure sore since coming home, I do not want to go back to the spinal unit man."
Pressure injury prevention and management strategies		
Local wound care	99	"... [My dad] had surgery to debride bedsores."
Use of equipment (eg, cushion, mattress)	38	"... [My mom] had a scary bedsore & was given a gel pad for chair & bed. I have a gel mattress which is great for not getting bedsores …"
Offloading	21	"... You are the person who turns her over every 2 hours or she will get bedsores."
Bed rest	14	"[He] has been trying to heal a pressure sore by being on bed rest for nearly a year."
Inspection or assessment	14	"… I do skin checks every morning when I get up, every evening when I go to bed, and every time I go to the bathroom, I check the vulnerable areas."
Medication	14	"Just beginning my 3rd month in hospital as a result of a pressure sore. 9 weeks of IV antibiotics due to osteomyelitis."
Nutrition	11	"When I was recovering from severe damage related to my pneumonia/sepsis episode (stage 3 pressure ulcers, crushed immune system, etc.) they made me aim for 120 [grams of protein] every day. Three medical Ensures, special feeding tube formula ..."
Healing of pressure injury		
Healing process	41	"At the start the pressure sore needed 3 packing strings. Now, only half of 1 fits in."
Healing time	25	"My grandma (86) got a really bad bedsore … After nearly 20 months, it's finally almost healed."
Lack of healing	22	"I have a pressure sore right now that's been open for several months and not close to healing …"
Worsening	12	"My downfall was a pressure sore which rapidly went from bad to worse …"

[a]Some posts contained multiple constructs of the PI experience.

[b]ICU: intensive care unit.

[c]ER: emergency room.

Theme 1: Challenges Experienced When Living With or Caring for a Pressure Injury

Within the 666 posts, there were 807 mentions related to the challenges associated with living with a PI, caring for someone with a PI, or receiving care for a PI. A total of 423 out of 807 (52.4%) mentions identified the personal impact or impacts of living with a PI or caring for someone with a PI, with 221 out of 423 (52.2%) mentions discussing the challenges associated with PI complications, such as infection and death. A total of 158 out of 423 (37.4%) mentions related to the personal impact of living with a PI reported a negative impact of PI or caregiving on overall functioning (eg, work or employment, mobility, daily activities) and perceived quality of life. A total of 23 out of 423 (5.4%) mentions related to the personal impact of living with a PI reported negative financial impacts (ie, struggling to afford

XSL•FO
RenderX

treatment), and 21 out of 423 (5%) discussed negative mental health impacts such as depression and social isolation.

> *My dad died from overwhelming sepsis due to osteomyelitis caused by his bedsores [and] of course everyone in the ICU had [pneumonia]. I checked his chart each day [and] it was sad to see signs of life slowly slipping away. Pressors couldn't keep him up.*
>
> *Mum has to stay in bed most of the day cause of pressure sore/open wound, we only transfer her to chair for a few hours.*

Additionally, there were 241 out of 807 (29.9%) mentions of the perceived quality of PI treatment received, with 205 out of 241 (85.1%) mentions describing poor care practices such as lack of care and misdiagnosis from HCPs that commonly led to PI complications for the individual. A total of 36 out of 241 (14.9%) mentions related to the perceived quality of treatment reported witnessing poor attitudes of caregivers and HCPs when treating the individual with a PI.

> *I had no idea [my mom] was suffering from a level 4 bedsore [because] the facility didn't tell me. The hospital told me. Shameful. This is the norm in most LTC facilities. This has to change. Now. #nursinghomeneglect.*
>
> *I haven't even mentioned some of the worst parts of this. My mom developed multiple new pressure ulcers at [healthcare facility]. I've had nurses and therapists privately tell me how bad things are at [healthcare facility]. I was told a nurse practitioner sighed in disappointment that my mom was improving.*

Furthermore, there were 102 out of 807 (12.6%) mentions related to perceived health care system issues when accessing services to treat PIs. Of these 102 mentions, 57 (55.9%) reported issues with accessing equipment or services (eg, difficulties obtaining pressure mattresses, long hospital wait times) and 26 (25.5%) mentions described prolonged hospitalization or inability to be discharged due to these challenges with health care access. Additionally, 19 out of 102 (18.6%) mentions related to perceived health care system issues described witnessing critical staffing shortages at health care facilities that contributed to the deterioration of the individual with lived PI experience.

> *Please help. Nobody will give us an air mattress for my elderly sick bed bound father! It's a disgrace when he is starting to get bedsores. He's under palliative care but without complex needs!!!!*
>
> *My poor dad ended up with a giant bedsore while on the barely staffed dementia ward at a [location] care home. He died of sepsis because of it. I was always stunned that there never seemed to be anyone up there working! So effed up!*

Finally, there were 41 out of 807 (5.1%) mentions related to a perceived lack of knowledge from HCPs, caregivers, or individuals with lived PI experience. Of these 41 mentions, 22 (53.7%) described a perceived lack of HCP's PI knowledge that contributed to the development or worsening of a PI. Moreover, 19 out of 41 (46.3%) mentions related to a lack of knowledge indicated the individual with lived PI experience or caregivers' lack of knowledge related to PIs or their prevention and management.

> *My mother had a horrendous [suppurating] bedsore on the back of her head after 8 weeks solid lying prone. We didn't know what it was but asked a nurse who shrugged and said, "Oh, that's weird. I don't know what it is" and walked off. It took 18 months to heal once home.*
>
> *I can't wait to look at my birthday messages, but [right now] I'm going to see my grandma [because] she's going to the hospital [for] bedsores that we don't know how to take care of.*

Theme 2: Needs Related to Pressure Injury Prevention and Management

There were 110 mentions about the users' needs related to PI prevention and management. Of these 110 mentions, 39 (35.5%) described a need for health care support. These mentions shared the users' need to receive PI care from HCPs, as well as the difficulty in accessing care.

> *[I have] infected pressure ulcers. Can't get ahold (sic) of family doctor. Virtual care clinics are booked up for multiple days—no way to get myself to a clinic.*
>
> *No, I had been trying to get treatment for pressure sores, but both my family members had covid at the time and couldn't take me to any appointments due to being covid positive.*

Moreover, of the mentions related to the user's needs related to PI prevention and management, 38 out of 110 (34.5%) mentions were related to a need for educational support. These mentions described users' need for advice from HCPs, health care organizations, or the general public related to PI prevention or management.

> *My dad is in a care home. He has dementia and now bad grade bedsores. Doctor said they are bad and could turn sepsis and [there is] nothing else they can do. Do I need a second opinion? Please help. Thanks.*
>
> *[My dad's] bedsores are bad which the Doctor has said could lead to sepsis. Any help or advice, please.*

In addition, there were 26 out of 110 (23.6%) mentions related to the need for social support. These mentions often described a need for spiritual guidance, prayers, or words of encouragement.

> *Please pray for me, I'm suffering with a pressure sore that won't go away.*
>
> *I've been dealing with a pressure sore for a [little] bit now. Today I found out it's infected. It's very triggering as pressure sores is how I lost my legs. Please send good vibes [or] distractions and stuff. I feel awful.*

Finally, there were 7 out of 110 (6.4%) mentions related to users' need for financial support. In these mentions, users requested assistance with medical bills related to PI prevention or management.

XSL•FO
RenderX

[Please] help my family to pay my mom's hospital bill. There's another 16 [days] left for surgery but yesterday my mom got admitted due to sudden bedsores and she's in pain right now. We [do not have any] income to buy the medication.

Please help me. I'm getting pressure ulcers from having to lay in bed a lot due to feeling unwell from my cancer. I started a [GoFundMe] a while ago and it hasn't got too much traction. I love you all and hope you'll take a second to boost this.

Theme 3: Emotions Experienced Related to Living With or Caring for a Pressure Injury

There were 478 mentions related to emotions when experiencing a PI, including primary emotions (ie, joy, trust, fear, surprise, sadness, disgust, anger, and anticipation) and a combination of these primary emotions (eg, anger and sadness, fear, and anger). Of the 478 mentions related to emotions experienced, negative emotions were frequently mentioned: 91 (19%) of the emotions included anger, 36 (7.5%) included sadness, and 68 (14.2%) included a combination of the two. Combinations of sadness with other primary emotions (eg, disgust, surprise, anticipation) were also frequently observed in 82 (17.2%) of the mentions. In total, 62 (13%) of the mentions included a combination of anger with other emotions. Combinations of fear with other emotions were also common in 54 (11.3%) of the mentions.

My mother is in aged care. Now transferring to palliative care after neglect. Massive bedsores despite instructions from doctor on how to position her. $600k plus daily fees for neglect and malnutrition that will ultimately see her out. We are beyond furious.

I am a mess, honestly. My sister and I take turns everyday feeding and hydrating her. She has a nasty bedsore, bone exposed … We can't even get her up to take her for a walk. I am just sad.

I've been TERRIFIED of getting pressure sores for a decade [and] half now, given my super low ability to move even in the bed I'm stuck in. I'm sitting or lying almost 24/7 … I'm super high risk [and] when it finally happens ... The [doctors] don't care.

Positive emotions were less frequently mentioned and included joy in 11 out of 478 (2.3%) mentions, anticipation in 5 out of 478 (1%) mentions, or both in 21 out of 478 (4.4%) mentions. Often, these emotions were expressed when describing appreciation for the kindness of others or anticipation related to the PI recovery process or prognosis.

My heart bleeds seeing my dad this way, bedsore has taken over but your donations kept him till this time. I can't thank you all enough but I know my God will reward you for your good deeds amen.

Today's goal: I'm hoping to take him on a wheelchair expedition outside his room--maybe even the hospital--for the first time in 2 months! Hopefully his pressure ulcer and spinal pain will allow it.

A good day! Nurse says pressure sore looking good, don't need to bandage it up with padding anymore. I should be able to pick up drugs at Tesco later [and] I think I can try get driving assessed and car adapted so I can drive with foot drop while it's healing! STOKED!!! #traumarehab

Discussion

Overview

To our knowledge, this is the first study investigating the experiences and perspectives of individuals with lived PI experience as shared through public social media posts. Using SML, we were able to capture the candid experiences and perspectives of individuals with PIs and caregivers that may not be easily gathered from controlled studies. Based on our findings, three major themes were developed: challenges experienced when living with or caring for a PI, needs related to PI prevention and management, and emotions experienced when living with or caring for a PI.

Principal Findings

Perceived Challenges of Living With or Caring for a PI

The most frequently mentioned perceived challenge related to PIs was the personal impact of living with a PI or caring for someone with a PI; most mentions involved challenges associated with complications of PIs, such as infection and death. Previous work notes that individuals living with PIs have an estimated two times higher risk of mortality compared with individuals living without PIs, often due to the development of severe infections [45]. Individuals with PIs are at high risk for bacteremia, which can cause a mortality of over 50% [46]. The findings from this work may help guide the development of strategies to help address this concern among those with lived PI experience.

Social media users frequently mentioned the negative impacts of PI on the daily, physical, occupational, or general functioning of individuals living with PIs. Previous studies on caregivers and those living with PIs identified similar experiences, with individuals reporting a reduction in mobility, independent movement, and quality of life [47-49]. A recent systematic review found that individuals living with PIs often demonstrated declines in physical function and activity, along with low quality of life scores [50]. These deficits are exacerbated by PI complications, such as pain and odor, which may impede the performance of physical and social activities [50]. In alignment with best practices for PI prevention and management, a holistic interprofessional approach should be used to address not only the PI but also the complications of living with a PI to help optimize outcomes.

Perceived poor quality of treatment related to PI management by HCPs or caregivers was reported by social media users. Posts often mentioned poor attitudes of HCPs and caregivers and neglect of the individual with lived PI experience, leading to increased pain experienced by the individual with lived experience, the development or worsening of a PI, infection, and even death. Other studies have highlighted negative attitudes of HCPs toward PI prevention and/or management as well as HCPs who engage in inadequate PI care practices [51-55]. These deficits in practice are potentially related to reported low levels of knowledge of HCPs related to PI prevention and management

XSL•FO
RenderX

[56-60]. This sentiment was shared by social media users in this study who reported a lack of HCP knowledge related to PI management.

According to the knowledge, attitudes, and practices model, knowledge may correlate with attitudes and health practices [61,62]. Thus, the reported lack of knowledge among HCPs may have also contributed to the reported negative attitudes and suboptimal care practices reported by social media users in this study. This is supported by previous studies investigating nurses' attitudes and perceived barriers toward PI prevention and management, which found that inadequate nurse training and knowledge were major obstacles contributing to their negative attitudes and inadequate practices [7,63,64]. Since low knowledge about PI may lead HCPs to adopt poor attitudes and health care practices, the need to ensure that HCPs receive sufficient training in PI prevention and management should be considered.

In addition to a lack of training and knowledge, HCPs have cited a lack of time and equipment (ie, pressure-relieving devices) as prominent perceived barriers to care delivery [51,65-67]. Other studies have found that staffing shortages and inadequate HCP training were persistent barriers to PI prevention practices [68-71]. These findings align with results from the present study, as social media users commonly identified health care system issues, including staffing shortages and lack of availability of equipment (eg, specialized mattresses). Health care system–related barriers appear to be shared globally, and future initiatives should ensure that equipment for PI prevention and management is accessible in order to optimize health outcomes for individuals living with PIs.

PI-Related Needs Among Those With Lived Experience

HCPs play a pivotal role in ensuring that the health care needs and goals of the person living with the PI are met [72,73]. Therefore, it is not surprising that the most frequently mentioned need identified by social media users was health care support related to managing PIs. Although best practice guidelines include self-management where people living with PIs assume responsibility for their PI [74], the frequent sentiment of social media users was the need for health care support, highlighting the limitations of self-management alone. Additionally, users often reported a continued need for HCP support, suggesting the need for ongoing provision of services. One potential contributor to the prevalent need for support may be the inaccessibility of PI resources. Previous studies have demonstrated that individuals seeking specialized medical care experience barriers related to long wait times and high treatment costs, making it exceptionally challenging to obtain relevant services [75,76]. Thus, these barriers may impact individuals' access to satisfactory PI care, driving their self-reported need for health care support.

Social media platforms provide a resource for individuals seeking guidance on managing their illnesses, allowing them to benefit from the advice and experiences shared by others [77,78]. As such, our study found that one of the most commonly expressed needs of individuals living with PIs was their request for advice and information regarding PIs. Social media can be beneficial as an educational resource because authentic accounts of an individual's illness can make medical information more relevant to the user's situation and help them to better understand it [78]. However, given the large volume of information available on the internet, the quality of information can vary widely [79] depending on its source and other post characteristics (eg, length of the post, author credentials). A study evaluating the quality of health information shared on online discussion forum websites, such as Reddit, found that posts shared by medical doctors were "of reasonably good quality" [80]. A recent study by Bang et al [81] that examined the quality of PI health information on YouTube found that longer videos produced by a physician and other health personnel contained more accurate and reliable information about PI prevention and management. In contrast, health information on X may not be reliable: a recent retrospective cohort study revealed that X users were approximately 974 times more likely to encounter inaccurate than accurate information about monkeypox [82]. Similarly, another study found that while authenticated authors on X with large followings were more likely to produce accurate posts about COVID-19, over 25% of the posts were inaccurate, indicating that this information may not be reliable for educational purposes [83]. Future studies should evaluate the quality of PI health information on various social media platforms and elucidate the factors impacting information quality to provide guidance for users regarding where and how to obtain reliable and accurate information.

Social media serves not only for sharing individual experiences but also for offering and receiving support [78]. In our study, social support emerged as a frequently mentioned need, which aligns with previous study findings that many caregivers of adults with cancer expressed a desire for social support in the form of prayers and patient visitation [84]. Individuals living with PIs are particularly vulnerable to social isolation due to the challenges associated with immobility and body image issues due to characteristics of PIs such as exudate and malodor [7,85]. Hence, social media may be a valuable resource for those living with PIs, allowing them to connect with others who can offer support. This notion is further supported by a study on social media use among adults during the COVID-19 pandemic, which highlighted how social media platforms can be major sources of social support amid isolation, ultimately aiding in the reduction of loneliness [86]. Future work should consider developing interventions to address the social isolation that individuals living with PIs may face.

Emotions Experienced as a Result of a Pressure Injury

In this study, anger, sadness, and a combination of the two emotions were the most frequently cited among social media users. This aligns with previous research findings that individuals with lived PI experience and caregivers often feel negative emotions such as anger, frustration, and sadness due to the lengthy healing process of PIs, lack of care received, and PI symptoms such as pain [7,32,50,87-89]. In the qualitative study of Latimer et al. [32] that explored the perceptions of individuals with PIs regarding their role in PI prevention, participants reported feeling angry and frustrated about failing to receive appropriate PI preventative care and experiencing

XSL•FO
RenderX

negative interactions with HCPs. These findings suggest that the PI experience is often distressing and can incite intense negative emotions in those living with or caring for a PI. However, it is important to note that social media users tend to engage with and share negative emotions and experiences (eg, anger and distress) more frequently than positive ones on these platforms [90-95]. In addition, studies have shown that those in the online health community may be more likely to post negative content to seek emotional and informational support [96,97]. The platform X in particular has been associated with messages displaying highly negative emotions [92]. Given that X was our primary source of data, the potential likelihood of predominantly negative experiences reported on X was high, and as a result, our findings may be skewed. Nonetheless, our findings provide valuable insights into individuals' experiences living with or caring for a PI, highlighting the negative emotional impact of the pervasive challenges associated with preventing and managing PIs.

Implications for Digital Health

Our study identified the candid experiences, challenges, and needs of individuals with PIs and their caregivers, which have several actionable implications for digital health. In this study, individuals living with PIs and their caregivers frequently expressed a need for information and advice regarding PIs and PI care. Additionally, social media users often reported a continued need for HCP support throughout the PI journey, which they lacked access to. These findings suggest that researchers and clinicians should prioritize the development of accessible online PI educational materials or platforms for individuals with PI and caregivers. Online PI resources or platforms (eg, websites, mobile health apps, social media pages) could provide contact information for HCPs or health organizations, enabling individuals with PI and/or their caregivers to access PI information or support when needed. Disseminating critical health information in a centralized and accessible forum may help promote health literacy among individuals with PIs and caregivers who rely on social media sources to supplement knowledge gaps. In addition, given the widespread use of social media today, it may also be useful to develop social media campaigns to educate the general public about PIs to combat general misinformation.

SML allowed us to identify patterns in the experiences and struggles of individuals with PI and their caregivers, such as a perceived lack of adequate PI care received and lack of knowledge among HCPs. These issues may subsequently be targeted and addressed by relevant regulatory or health entities to improve patient experiences and outcomes. As such, we note that SML may be used in the future to complement digital health surveillance technology, as researchers may utilize SML to observe and track ongoing trends in the reported challenges and needs of individuals with PI and caregivers on social media. The information gathered by SML can generate important longitudinal insights into whether prevailing issues are being adequately addressed and what critical areas of health care improvement health care organizations and policymakers should focus on.

Limitations

While our study is strengthened by a rigorous screening and coding process to analyze and interpret social media data, we acknowledge several limitations. First, since researchers hand-coded all social media posts, it is possible that some codes were missed due to human error affecting the accuracy of the reported frequencies. However, this process was carried out through a meticulous process (ie, researchers held iterative discussions to determine codes and code definitions, with coding performed in duplicate), so this is not a notable concern. Additionally, we broadened search terms in our Boolean search strings to widen our search and maximize the number of relevant posts. However, we acknowledge that some of these search strings could be streamlined and improved upon in the future for optimal retrieval and efficiency in our analysis. Furthermore, only English-language posts were included, which largely limited our search results to content from English-speaking countries. Moreover, most of the posts included were created by users in the United States, meaning our findings may predominantly reflect the health care experiences of individuals living in this country. Another potential limitation is that we only searched YouTube, Reddit, and X; a broader range of social media platforms may have yielded more posts for inclusion. Additionally, we found that X yielded a disproportionately large number of posts compared with YouTube and Reddit, which may have skewed our analysis toward findings from X. Also, the format of posts across the selected platforms was very different and thus difficult to collectively analyze. As such, future research using SML may benefit from selecting more comparable social media platforms to facilitate analysis or analyzing posts from each platform separately. Additionally, we were unable to obtain a complete demographic profile of social media users (eg, age, gender, or sex) due to the limited demographic information available on social media user accounts. Since we analyzed posts from observers, caregivers, and individuals with lived experience as a homogeneous group, differences between the experiences, challenges, or needs of each subgroup were not elucidated. Future studies may benefit from subgroup analysis to identify unique findings for caregivers and individuals with lived experience. Finally, we acknowledge that there may be sampling bias present in our study, as social media users are often younger and more technologically literate [98]. Also, while we did not restrict our sample in terms of geographic location, we acknowledge that there are inequities in technology access and thus ability to use social media platforms [99]. These limitations may reduce the generalizability of our findings.

Conclusions

Our study used SML to explore the unfiltered experiences shared by individuals affected by PIs (ie, individuals with lived PI experience, caregivers, and observers) across social media platforms X, YouTube, and Reddit. We found that many users described challenges associated with PIs and PI care related to the negative personal impacts of living with a PI and receiving inadequate care from HCPs, highlighting the need to address these issues and identify barriers to appropriate PI care. Additionally, social media was often used to request or express a need for health care support and education, suggesting a

XSL•FO
RenderX

critical need for supportive and educational resources for those affected by PIs. Finally, the most frequently observed emotion across posts was anger, consistent with previous findings that individuals with lived PI experience and caregivers often experience negative emotions about the PI experience. Our findings provide valuable insights into the candid experiences of individuals living with PIs, uncovering several gaps in care and research that require intervention to improve the well-being of these individuals. Clinicians, policymakers, and researchers should prioritize addressing the prominent challenges and needs identified from this work to optimize PI care delivery and patient outcomes.

Acknowledgments

We would like to acknowledge Pilar Tabuenca and Uchechukwu Akunna for their help with coding of the data. Generative artificial intelligence was not used in the analysis or writing of the manuscript.

Funding

This study was funded by the University Health Network Foundation. The funders had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Keywords, phrases, and Boolean operators used on social media listening software Awario and X Pro to identify relevant Reddit, YouTube, and X posts.
[DOCX File, 16 KB - nursing_v9i1e76682_app1.docx]

Checklist 1
SRQR checklist.
[DOCX File, 112 KB - nursing_v9i1e76682_app2.docx]

References

1. Al Aboud AM, Manna B. Wound pressure injury management (archived). In: StatPearls: StatPearls Publishing; 2025. [Medline: 30422492]
2. Gefen A, Brienza DM, Cuddigan J, Haesler E, Kottner J. Our contemporary understanding of the aetiology of pressure ulcers/pressure injuries. Int Wound J 2022 Mar;19(3):692-704. [doi: 10.1111/iwj.13667] [Medline: 34382331]
3. Mondragon N, Zito PM. Pressure injury. In: StatPearls: StatPearls Publishing; 2025. [Medline: 32491791]
4. Afzali Borojeny L, Albatineh AN, Hasanpour Dehkordi A, Ghanei Gheshlagh R. The incidence of pressure ulcers and its associations in different wards of the hospital: a systematic review and meta-analysis. Int J Prev Med 2020;11:171. [doi: 10.4103/ijpvm.IJPVM_182_19] [Medline: 33312480]
5. Siotos C, Bonett AM, Damoulakis G, et al. Burden of pressure injuries: findings from the Global Burden of Disease Study. Eplasty 2022;22:e19. [Medline: 35873067]
6. Edsberg LE, Langemo D, Baharestani MM, Posthauer ME, Goldberg M. Unavoidable pressure injury: state of the science and consensus outcomes. J Wound Ostomy Continence Nurs 2014;41(4):313-334. [doi: 10.1097/WON.0000000000000050] [Medline: 24901936]
7. Burston A, Miles SJ, Fulbrook P. Patient and carer experience of living with a pressure injury: a meta-synthesis of qualitative studies. J Clin Nurs 2023 Jul;32(13-14):3233-3247. [doi: 10.1111/jocn.16431] [Medline: 35768933]
8. Spilsbury K, Nelson A, Cullum N, Iglesias C, Nixon J, Mason S. Pressure ulcers and their treatment and effects on quality of life: hospital inpatient perspectives. J Adv Nurs 2007 Mar;57(5):494-504. [doi: 10.1111/j.1365-2648.2006.04140.x] [Medline: 17284276]
9. García-Sánchez FJ, Martínez-Vizcaíno V, Rodríguez-Martín B. Conceptualisations on home care for pressure ulcers in Spain: perspectives of patients and their caregivers. Scand J Caring Sci 2019 Sep;33(3):592-599. [doi: 10.1111/scs.12652] [Medline: 30735262]
10. Shiferaw WS, Akalu TY, Mulugeta H, Aynalem YA. The global burden of pressure ulcers among patients with spinal cord injury: a systematic review and meta-analysis. BMC Musculoskelet Disord 2020 May 29;21(1):334. [doi: 10.1186/s12891-020-03369-0] [Medline: 32471497]

XSL•FO
RenderX

11. Hopkins A, Dealey C, Bale S, Defloor T, Worboys F. Patient stories of living with a pressure ulcer. J Adv Nurs 2006 Nov;56(4):345-353. [doi: 10.1111/j.1365-2648.2006.04007.x] [Medline: 17042814]
12. Ramos P, Borges C, Azevedo I, et al. Burden of informal caregivers of patients with pressure ulcers in a primary care setting. J Wound Care 2022 Oct 2;31(10):864-871. [doi: 10.12968/jowc.2022.31.10.864] [Medline: 36240796]
13. Rodrigues AM, Ferreira PL, Ferré-Grau C. Providing informal home care for pressure ulcer patients: how it affects carers' quality of life and burden. J Clin Nurs 2016 Oct;25(19-20):3026-3035. [doi: 10.1111/jocn.13356] [Medline: 27577533]
14. Padula WV, Delarmente BA. The national cost of hospital-acquired pressure injuries in the United States. Int Wound J 2019 Jun;16(3):634-640. [doi: 10.1111/iwj.13071] [Medline: 30693644]
15. Haesler E, Pittman J, Cuddigan J, et al. An exploration of the perspectives of individuals and their caregivers on pressure ulcer/injury prevention and management to inform the development of a clinical guideline. J Tissue Viability 2022 Feb;31(1):1-10. [doi: 10.1016/j.jtv.2021.10.008] [Medline: 34776327]
16. McInnes E, Chaboyer W, Murray E, Allen T, Jones P. The role of patients in pressure injury prevention: a survey of acute care patients. BMC Nurs 2014;13(1):41. [doi: 10.1186/s12912-014-0041-y] [Medline: 25516720]
17. Hyrkas K, Wiggins M. A comparison of usual care, a patient-centred education intervention and motivational interviewing to improve medication adherence and readmissions of adults in an acute-care setting. J Nurs Manag 2014 Apr;22(3):350-361. [doi: 10.1111/jonm.12221] [Medline: 24612391]
18. Chien LJ, Slade D, Dahm MR, et al. Improving patient-centred care through a tailored intervention addressing nursing clinical handover communication in its organizational and cultural context. J Adv Nurs 2022 May;78(5):1413-1430. [doi: 10.1111/jan.15110] [Medline: 35038346]
19. Schmidt AL, Rodriguez-Esteban R, Gottowik J, Leddin M. Applications of quantitative social media listening to patient-centric drug development. Drug Discov Today 2022 May;27(5):1523-1530. [doi: 10.1016/j.drudis.2022.01.015] [Medline: 35114364]
20. Gupta P, Khan A, Kumar A. Social media use by patients in health care: a scoping review. Int J Healthc Manag 2022 Apr 3;15(2):121-131. [doi: 10.1080/20479700.2020.1860563]
21. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021 May 12;23(5):e17917. [doi: 10.2196/17917] [Medline: 33978589]
22. Huang X, Li Z, Jiang Y, Li X, Porter D. Twitter reveals human mobility dynamics during the COVID-19 pandemic. PLoS ONE 2020;15(11):e0241957. [doi: 10.1371/journal.pone.0241957]
23. Perić Z, Basak G, Koenecke C, et al. Understanding the needs and lived experiences of patients with graft-versus-host disease: real-world European public social media listening study. JMIR Cancer 2023 Nov 10;9:e42905. [doi: 10.2196/42905] [Medline: 37948101]
24. Syntosi A, Felizzi F, Bouchet C. A social media listening study to understand the unmet needs and quality of life in adult and pediatric amblyopia patients. Ophthalmol Ther 2022 Dec;11(6):2183-2196. [doi: 10.1007/s40123-022-00571-2] [Medline: 36175822]
25. Picone M, Inoue S, DeFelice C, et al. Social listening as a rapid approach to collecting and analyzing COVID-19 symptoms and disease natural histories reported by large numbers of individuals. Popul Health Manag 2020 Oct;23(5):350-360. [doi: 10.1089/pop.2020.0189] [Medline: 32897820]
26. Powell GE, Seifert HA, Reblin T, et al. Social media listening for routine post-marketing safety surveillance. Drug Saf 2016 May;39(5):443-454. [doi: 10.1007/s40264-015-0385-6] [Medline: 26798054]
27. Voillot P, Riche B, Portafax M, et al. Social media platforms listening study on atopic dermatitis: quantitative and qualitative findings. J Med Internet Res 2022 Jan 28;24(1):e31140. [doi: 10.2196/31140] [Medline: 35089160]
28. Wolffsohn JS, Leteneux-Pantais C, Chiva-Razavi S, et al. Social media listening to understand the lived experience of presbyopia: systematic search and content analysis study. J Med Internet Res 2020 Sep 21;22(9):e18306. [doi: 10.2196/18306] [Medline: 32955443]
29. Cook NS, Kostikas K, Gruenberger JB, et al. Patients' perspectives on COPD: findings from a social media listening study. ERJ Open Res 2019 Feb;5(1):00128-02018. [doi: 10.1183/23120541.00128-2018] [Medline: 30775374]
30. Rodrigues A, Chauhan J, Sagkriotis A, Aasaithambi S, Montrone M. Understanding the lived experience of lung cancer: a European social media listening study. BMC Cancer 2022 Apr 30;22(1):475. [doi: 10.1186/s12885-022-09505-4] [Medline: 35490223]
31. Gourlan M, Pellechia A, Robineau S, et al. "What pressure ulcers mean to me?" Representations of pressure ulcer in persons with spinal cord injury: a qualitative study. J Tissue Viability 2020 Nov;29(4):324-330. [doi: 10.1016/j.jtv.2020.07.002] [Medline: 32830010]
32. Latimer S, Chaboyer W, Gillespie B. Patient participation in pressure injury prevention: giving patient's a voice. Scand J Caring Sci 2014 Dec;28(4):648-656. [doi: 10.1111/scs.12088] [Medline: 24117711]
33. Tharu NS, Alam M, Bajracharya S, Chaudhary GP, Pandey J, Kabir MA. Caregivers' knowledge, attitude, and practice towards pressure injuries in spinal cord injury at rehabilitation center in Bangladesh. Adv Orthop 2022;2022:8642900. [doi: 10.1155/2022/8642900] [Medline: 35747167]
34. Antony L, Thelly AS. Knowledge on prevention of pressure ulcers among caregivers of patients receiving home-based palliative care. Indian J Palliat Care 2022;28(1):75-79. [doi: 10.25259/IJPC_84_2021] [Medline: 35673375]

XSL•FO
RenderX

35. Schubart JR, Hilgart M, Lyder C. Pressure ulcer prevention and management in spinal cord-injured adults: analysis of educational needs. Adv Skin Wound Care 2008 Jul;21(7):322-329. [doi: 10.1097/01.ASW.0000323521.93058.47] [Medline: 18600074]
36. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
37. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for Reporting Qualitative Research: a synthesis of recommendations. Equator Network. URL: https://www.equator-network.org/reporting-guidelines/srqr/ [accessed 2026-04-29]
38. Awario. 2025. URL: https://awario.com/ [accessed 2024-04-14]
39. X Pro. 2025. URL: https://pro.x.com [accessed 2024-04-14]
40. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
41. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
42. Plutchik R. The nature of emotions. Am Sci 2001;89(4):344. [doi: 10.1511/2001.4.344]
43. FAQ. UHN Research. URL: https://www.uhnresearch.ca/content/faq [accessed 2026-04-29]
44. Activities exempt from human ethics review. University of Toronto: Division of the Vice-President, Research & Innovation. URL: https://research.utoronto.ca/ethics-human-research/activities-exempt-human-ethics-review [accessed 2026-04-29]
45. Song YP, Shen HW, Cai JY, Zha ML, Chen HL. The relationship between pressure injury complication and mortality risk of older patients in follow-up: a systematic review and meta-analysis. Int Wound J 2019 Dec;16(6):1533-1544. [doi: 10.1111/iwj.13243] [Medline: 31606950]
46. Braga IA, Pirett CCNS, Ribas RM, Gontijo Filho PP, Diogo Filho A. Bacterial colonization of pressure ulcers: assessment of risk for bloodstream infection and impact on patient outcomes. J Hosp Infect 2013 Apr;83(4):314-320. [doi: 10.1016/j.jhin.2012.11.008] [Medline: 23313027]
47. Esson L. Inpatients felt that pressure ulcers had emotional, mental, physical, and social effects on quality of life because nurses did not adequately treat or manage their pain or discomfort pressure ulcers and their treatment and effects on quality of life: hospital inpatient perspectives. Evid Based Nurs 2007 Oct;10(4):128. [doi: 10.1136/ebn.10.4.128]
48. Rodrigues AM, Ferré-Grau C, Ferreira PL. Being an informal caregiver of a person with a pressure ulcer in the Azores Islands. Adv Skin Wound Care 2015 Oct;28(10):452-459. [doi: 10.1097/01.ASW.0000471191.11548.dd] [Medline: 26375948]
49. de Souza DMST, Veiga DF, Santos IDDAO, Abla LEF, Juliano Y, Ferreira LM. Health-related quality of life in elderly patients with pressure ulcers in different care settings. J Wound Ostomy Continence Nurs 2015;42(4):352-359. [doi: 10.1097/WON.0000000000000142] [Medline: 26135820]
50. Roussou E, Fasoi G, Stavropoulou A, et al. Quality of life of patients with pressure ulcers: a systematic review. Med Pharm Rep 2023 Apr;96(2):123-130. [doi: 10.15386/mpr-2531] [Medline: 37197280]
51. Ebi WE, Hirko GF, Mijena DA. Nurses' knowledge to pressure ulcer prevention in public hospitals in Wollega: a cross-sectional study design. BMC Nurs 2019;18(1):20. [doi: 10.1186/s12912-019-0346-y] [Medline: 31139012]
52. Gedamu H, Abate T, Ayalew E, Tegenaw A, Birhanu M, Tafere Y. Level of nurses' knowledge on pressure ulcer prevention: a systematic review and meta-analysis study in Ethiopia. Heliyon 2021 Jul;7(7):e07648. [doi: 10.1016/j.heliyon.2021.e07648] [Medline: 34381901]
53. Grešš Halász B, Bérešová A, Tkáčová Ľ, Magurová D, Lizáková Ľ. Nurses' Knowledge and attitudes towards prevention of pressure ulcers. Int J Environ Res Public Health 2021 Feb 10;18(4):1705. [doi: 10.3390/ijerph18041705] [Medline: 33578837]
54. Muhammed EM, Bifftu BB, Temachu YZ, Walle TA. Nurses' knowledge of pressure ulcer and its associated factors at Hawassa University comprehensive specialized hospital Hawassa, Ethiopia, 2018. BMC Nurs 2020;19(1):51. [doi: 10.1186/s12912-020-00446-6] [Medline: 32549784]
55. Saleh MYN, Papanikolaou P, Nassar OS, Shahin A, Anthony D. Nurses' knowledge and practice of pressure ulcer prevention and treatment: an observational study. J Tissue Viability 2019 Nov;28(4):210-217. [doi: 10.1016/j.jtv.2019.10.005] [Medline: 31672404]
56. Coventry L, Towell-Barnard A, Winderbaum J, Walsh N, Jenkins M, Beeckman D. Nurse knowledge, attitudes, and barriers to pressure injuries: a cross-sectional study in an Australian metropolitan teaching hospital. J Tissue Viability 2024 Nov;33(4):792-801. [doi: 10.1016/j.jtv.2024.10.003] [Medline: 39448363]
57. Dalvand S, Ebadi A, Gheshlagh RG. Nurses' knowledge on pressure injury prevention: a systematic review and meta-analysis based on the Pressure Ulcer Knowledge Assessment Tool. Clin Cosmet Investig Dermatol 2018;11:613-620. [doi: 10.2147/CCID.S186381] [Medline: 30538522]
58. Fulbrook P, Lawrence P, Miles S. Australian nurses' knowledge of pressure injury prevention and management: a cross-sectional survey. J Wound Ostomy Continence Nurs 2019;46(2):106-112. [doi: 10.1097/WON.0000000000000508] [Medline: 30801563]
59. Klaas N, Serebro RL. Intensive care nurses' knowledge of pressure injury prevention. BMC Nurs 2024 Dec 2;23(1):876. [doi: 10.1186/s12912-024-02533-4] [Medline: 39623421]

XSL•FO
RenderX

60. Li Z, Marshall AP, Lin F, Ding Y, Chaboyer W. Knowledge of pressure injury in medical and surgical nurses in a tertiary level hospital: a cross-sectional study. J Tissue Viability 2022 Feb;31(1):24-29. [doi: 10.1016/j.jtv.2021.12.003] [Medline: 34887153]
61. Alsaleh FM, Elzain M, Alsairafi ZK, Naser AY. Perceived knowledge, attitude, and practices (KAP) and fear toward COVID-19 among patients with diabetes attending primary healthcare centers in Kuwait. Int J Environ Res Public Health 2023 Jan 29;20(3):2369. [doi: 10.3390/ijerph20032369] [Medline: 36767736]
62. Bettinghaus EP. Health promotion and the knowledge-attitude-behavior continuum. Prev Med 1986 Sep;15(5):475-491. [doi: 10.1016/0091-7435(86)90025-3] [Medline: 3774779]
63. Etafa W, Argaw Z, Gemechu E, Melese B. Nurses' attitude and perceived barriers to pressure ulcer prevention. BMC Nurs 2018;17:14. [doi: 10.1186/s12912-018-0282-2] [Medline: 29686535]
64. Kurtgöz A, Koç Z. Difficulties and obstacles experienced by intensive care unit nurses during the prevention and care of pressure injuries: a qualitative study. J Tissue Viability 2024 May;33(2):269-274. [doi: 10.1016/j.jtv.2024.04.004] [Medline: 38604926]
65. Cesca N, Szczepanski A, Malik W, et al. Facilitators and barriers to pressure injury prevention, management and education: perspectives from healthcare professionals—a qualitative study. Int Wound J 2024 Jan;21(1):e14371. [doi: 10.1111/iwj.14371] [Medline: 37658675]
66. Moore Z, Price P. Nurses' attitudes, behaviours and perceived barriers towards pressure ulcer prevention. J Clin Nurs 2004 Nov;13(8):942-951. [doi: 10.1111/j.1365-2702.2004.00972.x] [Medline: 15533100]
67. Tomas N, Mandume AM. Nurses' barriers to the pressure ulcer risk assessment scales implementation: a phenomenological study. Nurs Open 2024 Jan;11(1):e2079. [doi: 10.1002/nop2.2079] [Medline: 38268250]
68. Heerschap C, Woo KY. Formal caregiver experiences of caring for individuals at risk for or with a pressure injury: a metasynthesis. Adv Skin Wound Care 2022 Dec 1;35(12):680-687. [doi: 10.1097/01.ASW.0000855044.92545.af] [Medline: 35895590]
69. McGraw CA. Nurses' perceptions of the root causes of community-acquired pressure ulcers: application of the model for examining safety and quality concerns in home healthcare. J Clin Nurs 2019 Feb;28(3-4):575-588. [doi: 10.1111/jocn.14652] [Medline: 30129137]
70. Mwebaza I, Katende G, Groves S, Nankumbi J. Nurses' knowledge, practices, and barriers in care of patients with pressure ulcers in a Ugandan teaching hospital. Nurs Res Pract 2014;2014:973602. [doi: 10.1155/2014/973602] [Medline: 24707398]
71. Panagiotopoulou K, Kerr SM. Pressure area care: an exploration of Greek nurses' knowledge and practice. J Adv Nurs 2002 Nov;40(3):285-296. [doi: 10.1046/j.1365-2648.2002.02370.x] [Medline: 12383180]
72. Li Z, Marshall AP, Lin F, Ding Y, Chaboyer W. Registered nurses' approach to pressure injury prevention: a descriptive qualitative study. J Adv Nurs 2022 Aug;78(8):2575-2585. [doi: 10.1111/jan.15218] [Medline: 35307866]
73. Pressure injury management: risk assessment, prevention and treatment. : Registered Nurses' Association of Ontario; 2024 URL: https://rnao.ca/bpg/guidelines/pressure-injuries [accessed 2025-04-02]
74. Kuhnke J. Best practice recommendations for skin health and wound management. : Wounds Canada URL: https://www.woundscanada.ca/doclink/bpr-cover-2025/eyJ0eXAiOiJKV1QiLCJhbGciOiJIUzI1NiJ9.eyJzdWIiOiJicHItY292ZXItMjAyNSIsImlhdCI6MTc0MDg0NzA3OSwiZXhwIjoxNzQwOTMzNDc5fQ.d89UeRRwWC-_GRJtYIazcFjaU6V3pVEG7p26hY3RDgE [accessed 2025-04-02]
75. Liddy C, Moroz I, Affleck E, et al. How long are Canadians waiting to access specialty care? Retrospective study from a primary care perspective. Can Fam Physician 2020 Jun;66(6):434-444. [Medline: 32532727]
76. Schuldt R, Jinnett K. Barriers accessing specialty care in the United States: a patient perspective. BMC Health Serv Res 2024 Dec 5;24(1):1549. [doi: 10.1186/s12913-024-11921-0] [Medline: 39633337]
77. Mazanderani F, O'Neill B, Powell J. "People power" or "pester power"? YouTube as a forum for the generation of evidence and patient advocacy. Patient Educ Couns 2013 Dec;93(3):420-425. [doi: 10.1016/j.pec.2013.06.006] [Medline: 23830239]
78. Naslund JA, Grande SW, Aschbrenner KA, Elwyn G. Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube. PLoS ONE 2014;9(10):e110171. [doi: 10.1371/journal.pone.0110171] [Medline: 25333470]
79. Chua GP, Ng QS, Tan HK, Ong WS. Caregivers of cancer patients: what are their information-seeking behaviours and resource preferences? Ecancermedicalscience 2020;14:1068. [doi: 10.3332/ecancer.2020.1068] [Medline: 32728384]
80. Cole J, Watkins C, Kleine D. Health advice from internet discussion forums: how bad is dangerous? J Med Internet Res 2016 Jan 6;18(1):e4. [doi: 10.2196/jmir.5051] [Medline: 26740148]
81. Bang M, Jang CW, Cho HE, Kim HS. YouTube as a source of information for pressure injury: quality assessment and content analysis. Int Wound J 2024 Mar;21(3):e14729. [doi: 10.1111/iwj.14729] [Medline: 38445793]
82. Knudsen B, Høeg TB, Prasad V. Analysis of tweets discussing the risk of Mpox among children and young people in school (May-October 2022): a retrospective observational study. BMJ Paediatr Open 2024 Jan 31;8(1):e002236. [doi: 10.1136/bmjpo-2023-002236] [Medline: 38296602]
83. Swetland SB, Rothrock AN, Andris H, et al. Accuracy of health-related information regarding COVID-19 on Twitter during a global pandemic. World Med Health Policy 2021 Sep;13(3):503-517. [doi: 10.1002/wmh3.468] [Medline: 34540337]

XSL•FO
RenderX

84. Bloom RD, Beck S, Chou WYS, Reblin M, Ellington L. In their own words: experiences of caregivers of adults with cancer as expressed on social media. Oncol Nurs Forum 2019 Sep 1;46(5):617-630. [doi: 10.1188/19.ONF.617-630] [Medline: 31424447]
85. Ibeh C, Hambridge K. A narrative review of patients' lived experience of having a pressure injury. Br J Nurs 2024 Feb 22;33(4):S10-S23. [doi: 10.12968/bjon.2024.33.4.S10] [Medline: 38386534]
86. Rosen AO, Holmes AL, Balluerka N, et al. Is social media a new type of social support? social media use in Spain during the COVID-19 pandemic: a mixed methods study. Int J Environ Res Public Health 2022 Mar 26;19(7):3952. [doi: 10.3390/ijerph19073952] [Medline: 35409634]
87. Gorecki C, Lamping DL, Brown JM, Madill A, Firth J, Nixon J. Development of a conceptual framework of health-related quality of life in pressure ulcers: a patient-focused approach. Int J Nurs Stud 2010 Dec;47(12):1525-1534. [doi: 10.1016/j.ijnurstu.2010.05.014] [Medline: 20554279]
88. Jackson DE, Durrant LA, Hutchinson M, Ballard CA, Neville S, Usher K. Living with multiple losses: insights from patients living with pressure injury. Collegian 2018 Aug;25(4):409-414. [doi: 10.1016/j.colegn.2017.10.008]
89. Langemo DK, Melland H, Hanson D, Olson B, Hunter S. The lived experience of having a pressure ulcer: a qualitative analysis. Adv Skin Wound Care 2000;13(5):225-235. [Medline: 11075022]
90. Brady WJ, Wills JA, Jost JT, Tucker JA, Van Bavel JJ. Emotion shapes the diffusion of moralized content in social networks. Proc Natl Acad Sci U S A 2017 Jul 11;114(28):7313-7318. [doi: 10.1073/pnas.1618923114] [Medline: 28652356]
91. Choi J, Lee SY, Ji SW. Engagement in emotional news on social media: intensity and type of emotions. Journal Mass Commun Q 2021 Dec;98(4):1017-1040. [doi: 10.1177/1077699020959718]
92. Masciantonio A, Bourguignon D. Too positive to be tweeted? An experimental investigation of emotional expression on Twitter and Instagram. Media Psychol 2024 Mar 3;27(2):243-270. [doi: 10.1080/15213269.2023.2236935]
93. Rathje S, Van Bavel JJ, van der Linden S. Out-group animosity drives engagement on social media. Proc Natl Acad Sci U S A 2021 Jun 29;118(26):e2024292118. [doi: 10.1073/pnas.2024292118] [Medline: 34162706]
94. Schöne JP, Parkinson B, Goldenberg A. Negativity spreads more than positivity on Twitter after both positive and negative political situations. Affect Sci 2021 Dec;2(4):379-390. [doi: 10.1007/s42761-021-00057-7] [Medline: 36043036]
95. Watson J, van der Linden S, Watson M, Stillwell D. Negative online news articles are shared more to social media. Sci Rep 2024 Sep 16;14(1):21592. [doi: 10.1038/s41598-024-71263-z] [Medline: 39285221]
96. Chen L, Baird A, Straub D. A linguistic signaling model of social support exchange in online health communities. Decis Support Syst 2020 Mar;130:113233. [doi: 10.1016/j.dss.2019.113233]
97. Robertson CE, Pröllochs N, Schwarzenegger K, Pärnamets P, Van Bavel JJ, Feuerriegel S. Negativity drives online news consumption. Nat Hum Behav 2023 May;7(5):812-822. [doi: 10.1038/s41562-023-01538-4] [Medline: 36928780]
98. Cotten SR, Schuster AM, Seifert A. Social media use and well-being among older adults. Curr Opin Psychol 2022 Jun;45:101293. [doi: 10.1016/j.copsyc.2021.12.005] [Medline: 35065352]
99. Silva M, Anaba U, Jani Tulsani N, Sripad P, Walker J, Aisiri A. Gender-based violence narratives in internet-based conversations in Nigeria: social listening study. J Med Internet Res 2023 Sep 15;25:e46814. [doi: 10.2196/46814] [Medline: 37713260]

Abbreviations

HCP: health care professional
PI: pressure injury
SML: social media listening
SRQR: Standards for Reporting Qualitative Research

Edited by E Borycki; submitted 28.Apr.2025; peer-reviewed by C Ni, Q Xu; revised version received 09.Feb.2026; accepted 17.Feb.2026; published 16.Jun.2026.

Please cite as:
Song JZ, Tsang I, Mouratidis P, Saeed M, Ornelas L AA, Islam A, Gabison S
Using Social Media Listening to Understand the Pressure Injury Experience: A Qualitative Descriptive Study
JMIR Nursing 2026;9:e76682
URL: https://nursing.jmir.org/2026/1/e76682
doi:10.2196/76682

© Jessica Zihui Song, Isabel Tsang, Parthena Mouratidis, Manahil Saeed, Alma A Ornelas L, Alaul Islam, Sharon Gabison. Originally published in JMIR Nursing (https://nursing.jmir.org), 16.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted

XSL•FO
RenderX

use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Health Care Professionals' Perspectives on the Use of a Wearable Device for Early Detection and Continuous Vital Signs Monitoring of Acute Respiratory Infections in Nursing Homes: Qualitative Study

Laura C Grootegoed[1,2,3,4], MSc; Margriet C Pol[3,5,6], PhD; Cees MPM Hertogh[3,6], MD, PhD; Lisa M Kolodziej[1,2,3], MD; Menno D de Jong[1,2], MD, PhD; Janke Schinkel[1,2], MD, PhD; Laura W van Buul[3,6], PhD

[1]Department of Medical Microbiology and Infection Prevention, Amsterdam UMC location University of Amsterdam, Meibergdreef 9, Amsterdam, The Netherlands

[2]Infectious diseases, Amsterdam institute for Immunology and Infectious Diseases, Amsterdam, The Netherlands

[3]Aging & Later Life, Amsterdam Public Health, Amsterdam, The Netherlands

[4]Digital Health, Amsterdam Public Health, Amsterdam, The Netherlands

[5]Faculty of Health, Center of Expertise Urban Vitality, Research Group Occupational Therapy: Participation and Environment, Amsterdam University of Applied Sciences, Amsterdam, The Netherlands

[6]Department of Medicine for Older People, Amsterdam UMC location Vrije Unversiteit Amsterdam, Amsterdam, The Netherlands

Corresponding Author:
Laura C Grootegoed, MSc
Department of Medical Microbiology and Infection Prevention, Amsterdam UMC location University of Amsterdam, Meibergdreef 9, Amsterdam, The Netherlands

Abstract

Background: The growing aging population and staff shortages are placing pressure on Dutch nursing homes (NHs). These challenges have led to an increased interest in digital health technologies. Among these are wearable devices that allow for remote continuous monitoring of vital signs. An example is the Healthdot (smartQare), a wearable electronic device that continuously monitors heart rate, respiratory rate, and physical activity. In the context of acute respiratory infections (ARIs) in NHs, where initial symptoms can go unnoticed, continuous monitoring may aid in early recognition, timely intervention, and reduce staff workloads. However, little is known about how health care professionals perceive the use of continuous vital signs monitoring devices, such as the Healthdot, for this cause in NHs.

Objective: This study aims to explore the perspectives of healthcare professionals on the use of the Healthdot for early detection and monitoring of ARIs in NHs, to inform potential future implementation.

Methods: Semistructured interviews were conducted with 20 physicians, nurses, and certified nursing assistants from 4 NHs and 1 acute geriatric community hospital located in a NH. Interview transcripts were thematically analyzed to identify themes regarding their perspectives on the use of the Healthdot for monitoring ARIs in this setting.

Results: Five main themes were identified that related to the appropriate use of the Healthdot for NH clients and health care professionals: alignment of Healthdot use and NH clients' treatment policies, balancing safety and freedom, impact of the Healthdot on work processes, supporting rather than replacing care, and possible use during pandemics and in the future. Additionally, several preconditions for the use of the Healthdot were identified, including its usability, a support base among care staff, adequate training and guidance, communication with NH clients and their relatives, and a clear policy regarding its use.

Conclusions: Given the complexity of care in NHs, where clinical care is typically balanced against quality of life and a homelike environment, physicians generally expressed reserved attitudes toward the Healthdot, highlighting the need to consider multiple factors in its implementation. Care staff were generally positive about the device. Nevertheless, tailored assessment for each individual NH client remains essential, balancing treatment goals, safety, autonomy, and person-centered care. Additionally, clear communication and alignment between health care professionals in this setting are crucial, specifically regarding their expectations of the Healthdot's role in care processes. This study offers practical guidance that may inform future implementation efforts of continuous vital sign monitoring devices in NHs.

(JMIR Nursing 2026;9:e84436) doi:10.2196/84436

KEYWORDS

wearable electronic devices; acute respiratory infections; nursing homes; continuous vital signs monitoring; health care professionals

XSL•FO
RenderX

Introduction

Nations across the globe are experiencing a demographic shift toward an older population [1]. The growth in aging populations, combined with a shortage of health care professionals [1,2], is placing a significant burden on Dutch nursing homes (NHs) [3]. In light of these challenges, digital health technologies, such as surveillance or vital signs monitoring devices, are increasingly being explored as promising tools to support care delivery in NH settings [4,5]. According to the Dutch central government, providing care digitally can not only provide older adults with more autonomy but can also support health care professionals by relieving their workload [6]. Implementing these technologies in NHs cannot only help address the issue of staff shortages but also has the potential to improve the quality of care [4,6,7].

One recurring health care problem frequently encountered in NHs involves acute respiratory infections (ARIs) [8]. Early recognition of ARIs and associated clinical deterioration in frail older adults remains challenging, as initial symptoms may go unnoticed, and NH clients with cognitive impairment often struggle to communicate their symptoms, leading to diagnostic and treatment delays [8,9]. Measurement of objective parameters possibly associated with ARIs, such as temperature, heart rate, respiratory rate, oxygen saturation, and reduced physical activity, could potentially be helpful for early recognition of ARIs and associated clinical deterioration [10]. One promising digital health device for remote vital signs monitoring is the Healthdot (smartQare), a wearable electronic patch that measures heart rate, respiratory rate, and physical activity (scored from 0 to 15, ranging from lying still to intensive movement such as sports) [11]. In the future, this application may be used by health care professionals in NHs to support them in making appropriate care choices regarding suspected ARIs, which can be focused on treatment and comfort. Additionally, automated monitoring of heart rate and respiratory rate by the Healthdot may reduce the workload for NH staff by obviating the need for these manual measurements.

While these prospects seem promising, little is known about how health care professionals in NHs perceive the use of wearable devices for continuous vital signs monitoring of ARIs. Insight into their views is essential to facilitate and inform potential future implementation and may help overcome both known and unknown challenges. This qualitative interview study, therefore, explores the perspectives of health care professionals on the use of the Healthdot for monitoring ARIs in NHs.

Methods

This interview study is part of the research project PRIMA (Plaster-Based Respiratory Infection Monitoring Assistant), which focuses on validating the Healthdot for recognizing clinical deterioration in NH clients with ARIs [12]. We used a qualitative research approach to gain a rich understanding of the perspectives of health care professionals on this topic. This study adhered to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines (Checklist 1) [13].

Recruitment and Study Population

Four NHs and 1 acute geriatric community hospital (AGCH) located in a NH were invited to participate in this interview study. These organizations were previously approached during the recruitment phase of the PRIMA project, in which only one of the NHs and the AGCH participated. Hereafter, the term “NH client” refers to both NH residents and AGCH patients located in a NH. In this interview study, all 5 organizations agreed to participate and invited health care professionals within their organizations to take part in the interviews. The invitation was broadly distributed by email, explicitly requesting representation from different professional groups, including physicians, nurses, and certified nursing assistants. A contact person from each organization provided the researchers with the contact details of health care professionals who were interested in participation.

Participants were included if they provided written informed consent and spoke the Dutch language. Participants were not expected to have knowledge of the Healthdot. Participants were purposively selected to create variation in the client population they cared for (ie, individuals with somatic and psychogeriatric conditions, individuals in short-term residential care, individuals undergoing geriatric rehabilitation, and individuals receiving acute geriatric care).

Data Collection

A semistructured interview guide (Multimedia Appendix 1) was developed by the research team (LCG, MCP, CMPMH, LMK, JS, and LWvB) and addressed the following topics: demographics, current practices regarding ARIs, digital health technologies in NH care, and the Healthdot. Prior to the discussion of the topic “Healthdot,” participants received a short presentation explaining how the Healthdot is placed, how it is used for monitoring, and its intended use in nursing homes (ie, for early detection of ARIs and detection of clinical deterioration during an ongoing ARI). More information about the Healthdot can be found at the Philips smartQare website [14]. Formally trained PhD student LG performed a pilot interview with a nurse practitioner who was formerly employed in an NH. Adjustments were made to increase understandability, which were validated by the nurse practitioner.

Interviews were conducted by LG between May 2024 and December 2024 until data saturation was reached, defined as the point during data analysis at which no new information emerged. Depending on the participant’s preference, interviews were held either via Microsoft Teams or in person at the NH. All interviews were audio-recorded. Verbatim transcripts were made, accompanied by field notes that included contextual information such as the setting, the relationship with the participant, the behavior of the participant, and reflections. The transcripts of the first 3 interviews were reviewed by senior researcher LWvB, who provided feedback to enhance the quality and consistency of subsequent interviews.

Data Analysis

Thematic analysis was used to analyze the data in accordance with the 6-phase approach described by Braun and Clarke [15]. Transcripts were read by researchers LG, MP, and LWvB themselves to familiarize with the data. Inductive coding was

XSL•FO
RenderX

performed in MAXQDA 24 (VERBI Software) [16], using a bottom-up approach to allow themes to emerge directly from the data. Two researchers independently open coded nine interviews (LG and MP or LWvB), with an equal distribution of interviews with physicians and care staff. Discrepancies were discussed until agreement was reached, contributing to a shared understanding of coding and interpretation among the researchers. The remaining 11 interviews were open-coded by LG, guided by the coding decisions and interpretations agreed upon during the initial coding phase, after which the codes were reviewed and critically appraised by either MP or LWvB. Coding was conducted iteratively, with regular discussions between researchers LG, MP, and LWvB to review coding decisions, resolve uncertainties, and ensure consistency with earlier coding. Subsequently, axial coding was conducted to identify relationships between codes and organize them into categories. This process led to the identification of overarching themes that captured the main patterns in the data. Throughout the analysis, reflexive discussions among the research team (LG, MP, CH, JS, and LWvB) helped to validate interpretations. Illustrative verbatim quotes were translated by LG and subsequently reviewed and corrected by 2 native English speakers.

Ethical Considerations

Upon review of the study protocol, the Medical Ethics Review Committee of the Amsterdam University Medical Center concluded that this study does not fall within the scope of the Medical Research Involving Human Subjects Act [17] (WMO; case 2024.0480). Participants provided written informed consent prior to the interview.

Results

A total of 20 participants were interviewed. Most participants were aged 36 to 50 years, and the sample included physicians (n=7), nurses (n=9), and other health care professionals (Table 1). Participants provided care for various NH client populations. The interviews lasted between 27 and 72 minutes (mean 47 , SD 9 min).

Table . Characteristics of participants (N=20).

Characteristics	Values, n (%)
Age (y)	
18 - 35	6 (30)
36 - 50	10 (50)
>50	4 (20)
Profession	
Physician	7 (35)
Nurse practitioner	1 (5)
Quality nurse	3 (15)
Nurse	6 (30)
Certified nursing assistant	3 (15)
Type of nursing home department[a]	
Psychogeriatric	9 (45)
Somatic	6 (30)
Short-term residential care	6 (30)
Geriatric rehabilitation	5 (25)
Acute geriatric care	5 (25)
Duration of employment at organization (y)	
<1	4 (20)
1 - 5	6 (30)
6 - 10	4 (20)
>10	6 (30)

[a]Multiple options possible.

Contextual Background

Current Practices in Monitoring ARIs

Participants stated that monitoring of ARIs and the frequency thereof are dependent on the treatment policies of NH clients. NH clients with a curative treatment policy (ie, full medical care including life-prolonging treatment) are monitored regularly, while NH clients with a palliative treatment policy (ie, care primarily focused on comfort, with optional life-prolonging treatment) or a symptomatic treatment policy

XSL•FO
RenderX

(ie, care limited to symptom relief, no life-prolonging treatment) are monitored sporadically or not at all. Monitoring in cases of suspected ARI is performed by nurses or certified nursing assistants, either in consultation with or under the instruction of a physician, and consists of regular measurement of vital signs (ie, heart rate, respiratory rate, temperature, blood pressure, and oxygen saturation) and clinical observations (eg, clinical condition, bed confinement, and food and drink intake). For clinical observations, particularly in clients with psychogeriatric conditions, the continuity of care staff and their familiarity with NH clients were identified as critical factors, which are currently often lacking due to staffing shortages.

Organizational Perspectives on the Use of Digital Health Technologies

Differences were identified in how the participants described the perspective of their organization on digital health technologies. Some participants stated that their organization had no focus on digital health technologies at all, while other participants stated that their organization was actively implementing new technologies. In these latter organizations, digital health technology was considered a key focus area by management, with innovation teams and designated staff members embedded. Participants from some organizations mentioned that digital health technologies were being implemented, yet a clear and structured strategy was lacking, which limited sustainable implementation. Several organizations were already using digital health technologies to monitor vital signs, including bed sensors measuring respiratory and heart rate, vital signs monitors, or standalone devices.

Individual Perspectives on Digital Health Technologies

Participants' individual perspectives on digital health technologies were generally positive, provided that their implementation serves a clear goal that is beneficial for NH clients, health care professionals, or both. Most participants emphasized the need for technology to ensure future-proof health care, with fewer health care workers and more older adults. They noted that digital health technologies can offer possibilities to this end, and can help reduce workload. However, participants stressed that technology should not replace face-to-face care, but rather support it.

Perspectives on the Use of the Healthdot

Regarding the use of the Healthdot for monitoring ARIs in NHs, 5 main themes were identified: alignment of Healthdot use and treatment policy, balancing safety and freedom, impact of the Healthdot on work processes, supporting rather than replacing care, and possible use during pandemics and in the future.

Theme 1: Alignment of Healthdot Use and Treatment Policy

Participants emphasized the importance of having a clearly defined goal when using the Healthdot, aligned with the treatment policy of NH client. The Healthdot was generally considered suitable for NH clients with an active treatment goal, such as those receiving intermediate care after hospital admission or those with a curative treatment policy. It was also highlighted that the Healthdot could be of added value when monitoring ARIs in clients with psychogeriatric conditions, as this group often has difficulties expressing their symptoms. However, these clients tend to be fidgety and may intentionally or unintentionally pull off the Healthdot:

> *Yes, in residential care for people who still want a lot [in life]. And, for example, I think it [the Healthdot] would be very good for young people with dementia living in a group. That is often a group with quite a strong desire for treatment, but who are not eager for a monitoring visit.* [R2: physician]

Some physicians were hesitant about the Healthdot's actual impact on medical management, considering the limited treatment options for viral ARIs:

> *That's where my doubt lies with viral respiratory infections. Usually, you don't do very much, except for influenza. As you said, then you prescribe Oseltamivir. But if it's just a regular respiratory virus, it's a matter of support, symptom management, but beyond that I think: what more can you do?* [R20: physician]

Most participants perceived the Healthdot as less suitable for NH clients receiving palliative or symptomatic care, as vital signs are typically no longer routinely monitored, and deviations identified by the Healthdot would have limited or no consequences for clinical treatment. Physicians, in particular, voiced stronger reservations regarding its use in this context:

> *For example, someone who says: I don't really want any life-prolonging interventions anymore. If I don't wake up tomorrow, that's fine. What do you do with that in your system in terms of signaling? It is nice to know if someone is deteriorating, but you're not going to treat them because of that, you know.* [R8: physician]

Opinions varied concerning the potential added value of the Healthdot in providing comfort-focused care. While most physicians questioned whether earlier detection by the Healthdot would result in the earlier provision of care, care staff—including nurses and certified nursing assistants—mentioned that earlier identification of distress might enable more timely comfort measures:

> *It is useful to notice if someone is restless or becomes restless. Then you can go there and check to see: what does this person need? Really offering comfort.* [R12: quality nurse]

Several ethical dilemmas were raised regarding the potential contribution of the Healthdot to a focus on life-prolonging interventions, whereas it might be more appropriate in the NH population to focus on acceptance of the finiteness of life. In the latter case, one may not want to know and control everything regarding the clinical situation, as this is also accompanied by the burden of difficult decisions for professionals, NH clients, and their relatives:

> *On the one hand, I think: no, we really need to look [at possibilities regarding the Healthdot]. On the other hand, I think, is this truly the direction in which we should be heading, or should we also, at some*

XSL•FO
RenderX

> *point, simply accept that life is finite? And that sometimes, it's just the flu [you die of], so to speak. Of course, that's not a very politically correct thing to say, and it's very nuanced. I'm putting it quite bluntly now, but of course we're talking about people. There is also value in being ill in a nursing home, you know. Life is still meaningful then.* [R4: nurse practitioner]

> *Sometimes it is good to not always know, right? If someone significantly deteriorates without anyone noticing and then perhaps passes away, that can sometimes be a good thing. That not everyone knows, because if you know something, you also have responsibilities. […] For families it is often very difficult to have an opinion about this and it can be better if someone suddenly deteriorates rapidly and dies. Then it just is what it is.* [R6: physician]

Some participants mentioned the shift in Dutch NHs from a medical model toward a well-being and homelike model, which they perceived as conflicting with the implementation of the Healthdot. Nevertheless, other participants recognized that medical elements can still be necessary to support the well-being of NH clients:

> *The most important point is, then we're investing more in curative care, while I think that it should be more about well-being. So, I just wouldn't allocate more budget to further increase the intensity of curative care. We really need to realize that the average length of stay in a nursing home is only a few months, so what are we talking about? It's about making sure they have a good day, period.* [R1: physician]

Overall, a clear distinction emerged between physicians and care staff. Physicians tended to evaluate the Healthdot in terms of its relevance for treatment decisions and were, therefore, more critical in situations where monitoring was unlikely to influence clinical management. In contrast, care staff were generally more positive about its use, as they emphasized its potential to support the early recognition of changes in NH clients' conditions and to guide timely treatment- and comfort-focused care.

Theme 2: Balancing Safety and Freedom

Participants acknowledged that the Healthdot can influence both the safety and freedom of NH clients in positive and negative ways. They mentioned that the Healthdot could enhance safety by allowing earlier detection of deterioration and providing NH clients and their families with a sense of safety. At the same time, some NH clients may perceive this as a form of surveillance and a breach of their privacy. Additionally, restrictive measures taken in response to Healthdot alerts, such as isolating NH clients when an infection is suspected based on an elevated heart or respiratory rate, could be experienced as stressful and limit their freedom:

> *Yes, there is a strong view that no [restrictive] measures should be taken, because freedom is prioritized. That's really the tendency, more freedom, so we generally don't want restrictive measures.* [R6: physician]

Simultaneously, care staff identified potential benefits of the Healthdot in supporting NH clients' sense of freedom by reducing interference in daily life. Participants mentioned that the Healthdot could decrease the need for manual measurements, which are often experienced as burdensome, and eliminate nightly visits, thereby positively impacting sleep quality and recovery. For NH clients who are aggressive or refuse manual measurements, the Healthdot was seen as a less intrusive and more practical alternative.

> *You can't just keep opening doors all the time. People want to sleep, so how useful would it be if you just have a screen that allowed you to continuously monitor vital signs. That if there is an acute situation, you'd already be there before that acute situation.* [R17: nurse]

Altogether, both physicians and care staff acknowledged the potential impact of the Healthdot on safety and freedom. However, the perceived benefits in terms of reducing burden and supporting NH clients' freedom were primarily emphasized by care staff, whereas physicians did not explicitly highlight these aspects.

Theme 3: Impact of the Healthdot on Work Processes

Participants mentioned that the Healthdot could reduce the workload for care staff, by limiting the need for frequent manual measurements. This would also allow care staff to visit upon indication, for example, when an early warning signal is triggered by a prolonged elevated heart rate:

> *I think that it's efficient. I think that maybe you can say that you don't need to do checks 3 times a day, unless, for example, the dashboard of the Healthdot system shows abnormalities.* [R18: nurse]

Care staff expected that this would allow more time for other tasks and increase opportunities for personal interaction with NH clients. The Healthdot was particularly seen as valuable in intermediate care settings, where most monitoring takes place:

> *I think it [use of the Healthdot] also means more time for the clients. Sometimes you feel like you're just going from one client to the next, with multiple tasks. […] I think it would be really nice to have more time for a real conversation with someone, or to actually dive a bit deeper.* [R11: nurse]

Contrarily, some physicians expressed concerns that the Healthdot would increase their workload. They emphasized the large number of measurements and early warning signals generated by the device, along with the responsibility associated with interpreting and acting on these signals. They pointed out that, in medical practice, "measuring is knowing," and with that knowledge comes the responsibility to act. Additionally, concerns were raised about the potential for false-positive signals, particularly in this population where vital signs often deviate from the norm:

> *The more data you collect, the more you have to account for. Because all the data you receive, requires*

XSL•FO
RenderX

> *a response. As a physician, you have to relate to it. You're responsible for it. So, I don't want to know.* [R1: physician]

At the same time, participants acknowledged that the Healthdot could provide consistent, timely, and objective measures. These are often lacking, as measurements are regularly not or not adequately performed in practice. Respiratory rate was particularly regarded as a valuable addition, considering that there is currently no measurement equipment for this:

> *At least you're somewhat less dependent on whether care staff is going to do it [measuring vital signs] or not. So, that could be an advantage. Yes, so especially that you get more consistent and reliable input, whereas now it is often not consistent and also not always reliable.* [R8: Physician]

Nevertheless, participants highlighted that manual measurements would remain necessary, as the Healthdot does not measure all required vital signs, such as temperature, blood pressure, and oxygen saturation. It was stated that these parameters could be measured based on indications from the Healthdot, such as elevated heart or respiratory rates, instead of on a routine basis. Some physicians mentioned that the current level of monitoring by care staff is sufficient, suggesting that the Healthdot is not necessary.

The main difference between physicians and care staff within this theme concerned the perceived workload. While both groups agreed that the Healthdot would likely reduce the workload for care staff, physicians expressed concerns that their workload could increase. Care staff generally did not consider this potential impact on physicians.

Theme 4: Supporting Rather Than Replacing Care

Although participants recognized that the Healthdot could reduce workload, they emphasized that the device cannot replace care staff. Physicians, in particular, noted that observations made by care staff provide more valuable insights than measurements, especially when providing care focused on comfort:

> *Yes, then the question really is what … are they in pain at that moment? Is that why their respiratory rate is going up? Or is it actually the final phase [of their life]? Usually, the respiratory rate doesn't increase just like that. People will become restless, start moving, and that's often already a signal.* [R6: physician]

They also stressed the importance of ensuring that the Healthdot does not lead to a reduction in personal contact with NH clients, as this is often perceived as highly valuable. Care staff themselves generally did not share these concerns. Finally, costs were identified as an important factor. Participants noted that while the Healthdot is expensive, it cannot replace care staff. Therefore, the added value of the Healthdot must be substantial before an organization will consider its implementation.

Theme 5: Possible Use During Pandemics and in the Future

The possible use of the Healthdot during future pandemics or outbreaks was discussed in a few interviews. Although some participants viewed this as a promising application, others were hesitant, referring back to the balance between safety and freedom:

> *During COVID, it would have been nice if you could have seen in advance that someone's heart rate was rising. He develops fever, so he might have COVID, so you could have put him in quarantine earlier. But it's also a bit scary, because as soon as you, say, sneeze once, you might already have to go into quarantine, because that thing measures it all.* [R3: physician]

Moreover, some participants emphasized the potential value of the Healthdot in light of an anticipated future with fewer health care professionals available and a growing aging population. In this context, the Healthdot could support sustainable care delivery in NHs:

> *It's a necessity. It's an absolute necessity. We simply have no other choice, so I definitely think we should move in that direction. We shouldn't wait too long either, because if the statistics are correct, the staffing shortages will become very severe in the coming years.* [R14: quality nurse]

Preconditions for Using the Healthdot

One important precondition stated was the usability of the Healthdot. Participants mentioned that the device should be easy to use, allow monitoring from devices they currently use, be integrated into the electronic health records of NH clients, and be able to measure additional vital signs (ie, temperature, oxygen saturation, and blood pressure). It also needs to be trustworthy, and the early warning signal must be adjustable so that receiving too many false signals can be avoided. Moreover, participants pointed out that the Healthdot is single-use, suggesting that more sustainable alternatives should be explored.

Participants emphasized that successful implementation of the Healthdot requires a support base among care staff. They stated that this can be established by including them during the implementation process and incorporating their feedback. Several participants stated that it is essential to have enough care staff who are knowledgeable and skilled in using the Healthdot, suggesting that a dedicated team member could take the lead. Additionally, adequate training and guidance during the use of the Healthdot were emphasized as important, ensuring confidence for care staff.

Another precondition mentioned by several participants was the need for clear communication with NH clients and their families about expectations regarding the use of the Healthdot. Physicians stressed that it is crucial to prevent unrealistic expectations, for example, regarding how frequently the physician will review the data. They mentioned that some relatives may, at times, demand more than is feasible or preferable. Therefore, several participants recommended limiting NH clients' and relatives' access to the Healthdot data:

> *What I'm a bit worried about is that family members have unrealistic expectations. That you'll be at the*

XSL•FO
RenderX

> *bedside within 5 minutes. That you'll be held accountable for that, in a way.* [R3: physician]

Finally, a clear policy regarding the use of the Healthdot is needed, according to the participants. They noted that it should be clear for whom, when, and for what purpose using the Healthdot is appropriate. Moreover, clear instructions should be available that specify by whom early warning signals are followed up and which actions need to be taken. Participants also mentioned the need for clarifying roles and responsibilities, suggesting that care staff should be responsible for monitoring, while physicians provide clinical guidance and assess the progression of the infection.

Discussion

Principal Findings

This qualitative interview study provides valuable insights into health care professionals' perspectives on the use of the Healthdot for monitoring ARIs in NHs. While care staff were generally positive about the Healthdot, physicians' attitudes were mostly reserved, highlighting the need for careful consideration of the NH context and its associated complexities during potential implementation. Our findings show that implementing wearable continuous vital signs monitoring devices in this setting extends beyond clinical purposes and necessitates careful weighing of the NH client's treatment goals, ethical considerations, the balance of safety and freedom, its impact on care processes, and the importance of supporting rather than replacing personalized care.

Although perspectives of health care professionals on wearable continuous vital signs monitoring have been studied before in hospital settings, to our knowledge, we are the first to study these perspectives in the NH setting. A previous study by Weenk et al [18] found that hospital-based health care professionals were generally positive about continuous vital signs monitoring, whereas more reserved attitudes toward this were identified in our study. This difference might be explained by the context of NHs, where health care professionals must balance clinical care with NH clients' quality of life and a home-like environment. Participants raised ethical concerns regarding the use of the Healthdot, particularly in palliative care situations, as they expressed fears that the Healthdot might lead to life-prolonging care in these cases, instead of promoting comfort and well-being. Nevertheless, van der Steen et al. [19] emphasize that careful and timely recognition and management of symptoms are crucial for enhancing well-being in palliative care [19]. From this perspective, the Healthdot could potentially contribute to person-centered palliative care by supporting the timely recognition and relief of symptoms.

The balance between safety and freedom is a recurring topic in NH care that was also reflected in our findings. Participants noted that the Healthdot could enhance safety by enabling early detection of deterioration caused by ARIs and offering a sense of security, which is in line with findings from another study focusing on monitoring devices in NHs [20]. However, participants in our study also raised concerns about the impact of the Healthdot on NH clients' freedom, including feelings of being watched, privacy intrusion, and restrictive infection control measures. The COVID-19 pandemic demonstrated how such measures affected NH clients' autonomy and mental well-being [21]. Still, the relationship between freedom and care in NHs remains complex. It has been argued that some degree of intervention is always required in institutional care, making autonomy and freedom in this context not straightforward [22]. Haslam-Larmer et al [23] highlighted similar tensions in the context of location monitoring in dementia care, where privacy is often considered secondary to safety concerns [23]. Interestingly, participants also saw potential for the Healthdot to enhance NH clients' sense of freedom by reducing the need for manual measurements and nighttime disruptions. This aligns with findings from Emilsson et al [24], where NH staff reported that camera surveillance monitoring reduced unnecessary night-time visits [24]. To effectively balance safety and freedom, it is essential to consider the individual needs of each NH client, as the value placed on safety or freedom may differ for everyone.

A potential consequence of reducing manual measurements is a decrease in contact moments, which both physicians and care staff in our study considered essential for maintaining clinical insight and relationships. This concern is widely recognized across studies on both surveillance and vital signs monitoring in several care settings [18,24-26]. For instance, it was reported that continuous vital sign monitoring in hospitals may reduce bedside contact and compromise clinical judgment, as such systems cannot assess subjective experiences like pain [18,25,26]. Similarly, research on surveillance monitoring in NHs highlights the irreplaceable value of physical presence for both clinical observation and emotional reassurance [24]. However, Niemeijer et al [22] suggest that monitoring technologies do not necessarily reduce human contact, as during their study on surveillance technologies in residential facilities, nurses continued their rounds and maintained strong relationships [22]. In our study, participants also expected contact to remain necessary, as the Healthdot does not measure all vital signs needed for monitoring of ARIs and manual measurements are still required when indicated by the early warning signal. Moreover, both our findings and previous research suggest that time saved through monitoring may be reinvested in client-related tasks [18,22]. Nevertheless, with the expected increase in older adults in NHs and staff shortages [2,3], there is a risk that any time gained will be used to manage higher caseloads, potentially limiting opportunities for personal interaction.

Despite expectations that the Healthdot could save time for person-centered care, physicians in our study raised concerns about a possible increase in workload. Similar concerns have been reported in studies conducted in hospital settings, where the high frequency of measurements was found to burden physicians [18]. In line with our findings, several studies identified false-positive alerts as a contributor to "alarm fatigue," further adding to the workload [18,22]. Physicians in our study also expressed concerns about the increased burden of responsibility that comes with more measurements. NH clients and their relatives might have unrealistic expectations, assuming immediate intervention when measurements deviate. Previous research has shown that unrealistic demands from family

XSL•FO
RenderX

members are a recurring challenge in NHs and can contribute to staff stress [27]. Managing expectations is, therefore, crucial for sustainable use.

Notably, our findings revealed differing perspectives between physicians and care staff regarding the use of the Healthdot. Across themes, we found that care staff were generally enthusiastic, viewing the device as a way to reduce workload, increase time for personal interaction, and enhance the freedom and well-being of NH clients. Physicians, however, expressed more critical views, raising concerns about the Healthdot's impact on clinical management, increased workload and responsibility, and potential implications for person-centered care. These differences may be explained by the distinct roles and responsibilities that physicians and care staff have within interprofessional collaboration, as described by Schot et al [28]. Whereas physicians may see the Healthdot primarily as a diagnostic tool requiring clinical action and interpretation, nurses may view it mainly as a support system in their caregiving routine, evaluating it in terms of practicality, efficiency, and NH clients' comfort. Recognizing these divergent perspectives is essential, as they may influence acceptance, integration into care routines, and overall effectiveness. This suggests that the successful implementation of the Healthdot will require clear communication, alignment on its intended purpose, shared understanding of its role in care practices, and active collaboration between health care professionals. Additionally, acceptance is influenced by the nursing home's organizational culture and leadership supporting innovation [29].

One limitation of this study is that the Healthdot is not currently used for monitoring ARIs in NHs, meaning that participants had to speculate about its potential use. This hypothetical context limited the extent to which their responses reflect real-world experiences or challenges and led some participants to reflect on the general use of the Healthdot in NHs rather than specifically for monitoring ARIs. Nonetheless, a strength of this study lies in its early involvement of end users, reflecting a growing emphasis in implementation research [30] and providing early insights, considerations, and key prerequisites for potential implementation. Additionally, our participants stressed the importance of including their perspectives to enhance acceptance and practical relevance, which aligns with evidence that such engagement fosters organizational readiness through an innovative culture and shared commitment [31,32]. Although some participants did have prior experience with the Healthdot in a research context due to their participation in the quantitative PRIMA study, this experience was limited to the placement of the Healthdot and did not involve monitoring with the Healthdot. No differences in responses were observed between participants with and without this prior experience.

Conclusions and Implications

Our findings demonstrate that implementing continuous vital signs monitoring devices, such as the Healthdot in NHs, involves multiple considerations and conditions, shaped by the complexity of this care setting. A tailored assessment is essential for each NH client, taking into account factors such as treatment goals, safety, autonomy, and person-centered care. Clear communication and alignment between physicians and care staff, as well as between NH staff and relatives of NH clients are crucial, particularly regarding expectations about the Healthdot's impact on care processes. Although this study focused on ARIs, many insights may be relevant to the broader application of monitoring devices in NHs. This study offers practical guidance and key considerations that may inform future implementation efforts of continuous vital signs monitoring devices in this setting.

Acknowledgments

The authors thank the nursing home health care professionals who participated in the interviews. Additionally, the authors thank Jessica Williams and Tony Williams for their assistance in translating the interview quotes, as well as Kelly Paap for participating in the pilot interview. Generative artificial intelligence tools were used to assist with proofreading and language editing of the text.

Funding

This work was funded by ZonMW (project 10430362220010). The funder had no involvement in the study design, data collection, analysis, interpretation, or writing of the manuscript.

Data Availability

Due to the sensitivity of the data, the dataset will not be shared publicly.

Authors' Contributions

Conceptualization: LCG, LWvB, MCP, JS, CMPMH, LMK, MDdJ
Data curation: LCG
Formal analysis: LCG, LWvB, MCP
Funding acquisition: JS, MDdJ
Investigation: LCG
Methodology: LCG, LWvB, MCP, LMK
Project administration: LCG

XSL•FO
RenderX

Supervision: LWvB, MCP, JS, CMPMH
Validation: LWvB, MCP
Visualization: LCG
Writing – original draft: LCG
Writing – review & editing: LCG, LWvB, MCP, JS, CMPMH, LMK, MDdJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File, 42 KB - nursing_v9i1e84436_app1.docx]

Checklist 1
COREQ checklist.
[PDF File, 432 KB - nursing_v9i1e84436_app2.pdf]

References

1. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2026-05-14]
2. Health workforce. World Health Organization. URL: https://www.who.int/health-topics/health-workforce#tab=tab_1 [accessed 2026-05-14]
3. Wachtlijsten verpleeghuizen: waar wachten we op? [Web page in Dutch]. Actiz. URL: https://www.actiz.nl/wachtlijsten-verpleeghuizen-waar-wachten-we-op [accessed 2026-05-14]
4. Geschat potentieel digitale zorg [Web page in Dutch]. Rijksoverheid. URL: https://www.rijksoverheid.nl/documenten/rapporten/2023/12/22/rapport-ministerie-van-vws-geschat-potentieel-digitale-zorg [accessed 2026-05-14]
5. Digital health. World Health Organization. URL: https://www.who.int/europe/health-topics/digital-health#tab=tab_1 [accessed 2026-05-14]
6. Programma wonen, ondersteuning en zorg voor ouderen (WOZO) [Report in Dutch]. : Ministerie van Volksgezondheid, Welzijn en Sport; 2022 URL: https://open.overheid.nl/documenten/ronl-63e851f8e45b8662a4d04b1d5832b98231d40670/pdf [accessed 2026-05-14]
7. Effectieve inzet van e-health in de verpleeghuiszorg vraagt om professionalisering bij zorgaanbieders [Report in Dutch]. : Inspectie Gezondheidszorg en Jeugd (IGI); 2021 URL: https://www.igj.nl/site/binaries/site-content/collections/documents/2021/09/29/e-health-in-de-verpleeghuiszorg/e-health+in+de+verpleeghuiszorg.pdf [accessed 2026-05-14]
8. Michener A, Heath B, Crnich CJ, et al. Infections in older adults: a case-based discussion series emphasizing antibiotic stewardship. MedEdPORTAL 2018 Sep 21;14:10754. [doi: 10.15766/mep_2374-8265.10754] [Medline: 30800954]
9. D'Agata E, Loeb MB, Mitchell SL. Challenges in assessing nursing home residents with advanced dementia for suspected urinary tract infections. J Am Geriatr Soc 2013 Jan;61(1):62-66. [doi: 10.1111/jgs.12070] [Medline: 23311553]
10. Radin JM, Wineinger NE, Topol EJ, Steinhubl SR. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the USA: a population-based study. Lancet Digit Health 2020 Feb;2(2):e85-e93. [doi: 10.1016/S2589-7500(19)30222-5] [Medline: 33334565]
11. van der Stam JA, Mestrom EHJ, Scheerhoorn J, et al. Accuracy of vital parameters measured by a wearable patch following major abdominal cancer surgery. Eur J Surg Oncol 2022 Apr;48(4):917-923. [doi: 10.1016/j.ejso.2021.10.034] [Medline: 34753618]
12. Using a smart wearable as an early warning system and as a tool to monitor disease severity in COVID-19 real-time in a susceptible population: a proof-of-concept study for infection with SARS-cov2 variants or other respiratory viruses with pandemic poten [Web page in Dutch]. ZonMW Projecten. URL: https://projecten.zonmw.nl/nl/project/using-smart-wearable-early-warning-system-and-tool-monitor-disease-severity-covid-19-real [accessed 2026-05-14]
13. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
14. Ontdek healthdot 3.1 [Webpage in Dutch]. smartQare. URL: https://www.smartqare.nl/producten-en-oplossingen/ontdek-healthdot-3-1/ [accessed 2026-05-15]
15. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
16. MAXQDA. URL: https://www.maxqda.com/ [accessed 2026-05-14]

XSL•FO
RenderX

17. Toetsingskader medisch-wetenschappelijk niet-WMO-plichtig onderzoek [Report in Dutch]. : Ministerie van Volksgezondheid, Welzijn en Sport; 2024 URL: https://www.health-ri.nl/sites/healthri/files/2024-02/VWS%20Toetsingskader%20nWMO-onderzoek_DEF.pdf [accessed 2026-04-24]
18. Weenk M, Koeneman M, van de Belt TH, Engelen LJLPG, van Goor H, Bredie SJH. Wireless and continuous monitoring of vital signs in patients at the general ward. Resuscitation 2019 Mar;136:47-53. [doi: 10.1016/j.resuscitation.2019.01.017] [Medline: 30685546]
19. van der Steen JT, Radbruch L, Hertogh CMPM, et al. White paper defining optimal palliative care in older people with dementia: a Delphi study and recommendations from the European Association for Palliative Care. Palliat Med 2014 Mar;28(3):197-209. [doi: 10.1177/0269216313493685] [Medline: 23828874]
20. Hall A, Brown Wilson C, Stanmore E, Todd C. Moving beyond 'safety' versus 'autonomy': a qualitative exploration of the ethics of using monitoring technologies in long-term dementia care. BMC Geriatr 2019 May 24;19(1):145. [doi: 10.1186/s12877-019-1155-6] [Medline: 31126240]
21. Ah H, Amm R, S T, Sg J, B L, S E. Nurses' ethical challenges when providing care in nursing homes during the COVID-19 pandemic. Nurs Ethics 2023 Feb;30(1):32-45. [doi: 10.1177/09697330221105631] [Medline: 36250534]
22. Niemeijer AR, Depla M, Frederiks B, Francke AL, Hertogh C. CE: original research: the use of surveillance technology in residential facilities for people with dementia or intellectual disabilities: a study among nurses and support staff. Am J Nurs 2014 Dec;114(12):28-37. [doi: 10.1097/01.NAJ.0000457408.38222.d0] [Medline: 25396755]
23. Haslam-Larmer L, Grigorovich A, Shum L, et al. Factors that influence successful adoption of real-time location systems for use in a dementia care setting: mixed methods study. JMIR Aging 2024 Apr 8;7:e45978. [doi: 10.2196/45978] [Medline: 38587884]
24. Emilsson M, Karlsson C, Svensson A. Experiences of using surveillance cameras as a monitoring solution at nursing homes: the eldercare personnel's perspectives. BMC Health Serv Res 2023 Feb 10;23(1):144. [doi: 10.1186/s12913-023-09130-2] [Medline: 36765321]
25. Leenen JPL, Dijkman EM, van Hout A, Kalkman CJ, Schoonhoven L, Patijn GA. Nurses' experiences with continuous vital sign monitoring on the general surgical ward: a qualitative study based on the Behaviour Change Wheel. BMC Nurs 2022 Mar 14;21(1):60. [doi: 10.1186/s12912-022-00837-x] [Medline: 35287678]
26. Areia C, King E, Ede J, et al. Experiences of current vital signs monitoring practices and views of wearable monitoring: a qualitative study in patients and nurses. J Adv Nurs 2022 Mar;78(3):810-822. [doi: 10.1111/jan.15055] [Medline: 34655093]
27. Muldrew Née Preshaw DH, McLaughlin D, Brazil K. Ethical issues experienced during palliative care provision in nursing homes. Nurs Ethics 2019 Sep;26(6):1848-1860. [doi: 10.1177/0969733018779218] [Medline: 29945483]
28. Schot E, Tummers L, Noordegraaf M. Working on working together. a systematic review on how healthcare professionals contribute to interprofessional collaboration. J Interprof Care 2020;34(3):332-342. [doi: 10.1080/13561820.2019.1636007] [Medline: 31329469]
29. Van Puyvelde S, Vandenhaute ML, Bruning L. Artifacts of innovation-oriented organizational culture in nonprofit nursing homes: an explorative study. J Health Organ Manag 2026 Jan 14;40(1):150-163. [doi: 10.1108/JHOM-12-2023-0377] [Medline: 40770840]
30. Ahouah M, Rothan-Tondeur M. End-users and caregivers' involvement in health interventional research carried out in geriatric facilities: a systematic review. Int J Environ Res Public Health 2019 Aug 7;16(16):16. [doi: 10.3390/ijerph16162812] [Medline: 31394752]
31. von Treuer K, Karantzas G, McCabe M, et al. Organizational factors associated with readiness for change in residential aged care settings. BMC Health Serv Res 2018 Feb 1;18(1):77. [doi: 10.1186/s12913-018-2832-4] [Medline: 29390999]
32. Pimentel CB, Mills WL, Snow AL, et al. Adapting strategies for optimal intervention implementation in nursing homes: a formative evaluation. Gerontologist 2020 Nov 23;60(8):1555-1565. [doi: 10.1093/geront/gnaa025] [Medline: 32449764]

Abbreviations

AGCH: acute geriatric community hospital
ARI: acute respiratory infection
COREQ: Consolidated Criteria for Reporting Qualitative Research
NH : nursing home
PRIMA: Plaster-Based Respiratory Infection Monitoring Assistant

XSL•FO
RenderX

Edited by E Borycki, J Sarvestan; submitted 25.Sep.2025; peer-reviewed by A Pullyblank, J Edwards; revised version received 01.Apr.2026; accepted 02.Apr.2026; published 25.May.2026.

Please cite as:
Grootegoed LC, Pol MC, Hertogh CMPM, Kolodziej LM, de Jong MD, Schinkel J, van Buul LW
Health Care Professionals' Perspectives on the Use of a Wearable Device for Early Detection and Continuous Vital Signs Monitoring of Acute Respiratory Infections in Nursing Homes: Qualitative Study
JMIR Nursing 2026;9:e84436
URL: https://nursing.jmir.org/2026/1/e84436
doi:10.2196/84436

© Laura Cornelle Grootegoed, Margriet C Pol, Cees MPM Hertogh, Lisa M Kolodziej, Menno D de Jong, Janke Schinkel, Laura W van Buul. Originally published in JMIR Nursing (https://nursing.jmir.org), 25.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Training Programs for Dementia Care Staff and Aggression in Nursing Homes: Secondary Analysis of a Cross-Sectional Online Survey

Reed Bratches[1], PhD; Tamara Backhouse[2], PhD, MA, BSc(Hons); Shaoqing Ge[3], PhD; Megan Hebdon[4], PhD, DNP, FNP-C, APHN-BC, NC-BC; Dalton Riegel[5]; Noah Stein[5]; Katie Trainum[6], PhD; Rita Jablonski[1], PhD

[1]School of Nursing, University of Alabama at Birmingham, 1720 University Boulevard, Birmingham, AL, United States

[2]School of Health Sciences, University of East Anglia, Norwich, England, United Kingdom

[3]The University of Texas at Austin, Austin, TX, United States

[4]University of Utah, Salt Lake City, UT, United States

[5]Franklin & Marshall College, Lancaster, PA, United States

[6]Center for healthcare Evaluation, Research, and Promotion, Corporal Michael J. Crescenz VA Medical Center, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Reed Bratches, PhD
School of Nursing, University of Alabama at Birmingham, 1720 University Boulevard, Birmingham, AL, United States

Abstract

Background: Training programs help dementia care staff handle residents' behavioral symptoms in nursing homes and assisted living facilities. However, it is unclear how feeling well-prepared from such training relates to experiences of physical and verbal aggression from residents.

Objective: This study aimed to determine whether perceived training preparedness was related to experiences of physical and verbal aggression from residents in nursing homes and assisted living facilities using data from a large, nationally representative dementia care workforce sample.

Methods: This study was an analysis of the 2024 National Dementia Workforce Study, a cross-sectional online survey of the national dementia care workforce. Weighted regression modeling was used to identify associations between feeling well-prepared by training programs and experiencing physical or verbal aggression.

Results: Dementia care staff in nursing homes (weighted n=896,320) commonly reported experiencing physical (365,452/896,320, 40.8%) and verbal (470,536/896,320, 52.5%) aggression from residents with dementia, whereas more than two-thirds (614,561/896,320, 68.6%) reported feeling that their training prepared them well to manage resident behaviors. Staff in assisted living facilities (n=738,856) less commonly reported experiencing physical (198,964/738,856, 26.9%) and verbal (223,343/738,856, 30.2%) aggression, whereas almost three-quarters (527,621/738,856, 71.4%) reported feeling that their training prepared them well to manage resident behaviors. In regression modeling in nursing homes, feeling well-prepared by training for managing resident behaviors was associated with a lower likelihood of experiencing physical aggression (odds ratio [OR] 0.4, 95% CI 0.223-0.711; *P*=.01) and verbal aggression (OR 0.5, 95% CI 0.299-0.844; *P*=.02). In assisted living facilities, feeling well-prepared by training for managing resident behaviors was associated with a lower likelihood of experiencing physical (OR 0.14, 95% CI 0.04-0.55; *P*=.01) and verbal (OR 0.07, 95% CI 0.0134-0.354; *P*=.001) aggression.

Conclusions: While physical and verbal aggression from residents with dementia is common in nursing homes and assisted living facilities, nationally representative weighted regression modeling highlights an association between feeling well-prepared to handle resident behaviors and a lower likelihood of experiencing aggression. As the data used were derived from a cross-sectional survey, the causal link between these 2 factors should be explored in interventional studies. Our results indicate that effective training programs are associated with worker-reported reduced experiences of physical and verbal aggression in both nursing homes and assisted living facilities. The specific components of training programs that may support effectiveness remain unknown.

(JMIR Nursing 2026;9:e94077) doi:10.2196/94077

KEYWORDS

dementia care workforce; nursing home; assisted living; behavioral symptoms of dementia; care-resistant behavior

XSL•FO
RenderX

Introduction

Aggression is a behavioral symptom of dementia that manifests as physically or verbally hostile or violent behavior [1]. Physical aggression includes striking, kicking, hair pulling, spitting, and throwing things, whereas verbal aggression includes swearing, insulting, threatening, or verbal sexual harassment [2]. Physically and verbally aggressive behaviors are common in residential facilities such as nursing homes and assisted living facilities, with estimates indicating aggression in 51% of patients with dementia in these facilities [3]. Staff exposure to aggression is associated with a higher likelihood of burnout and physical harm, including bruising and lacerations, and contributes to workforce turnover [4].

General nursing training programs rarely include focused training on managing the behavioral symptoms of dementia, including aggression [5]. While some licensed professional roles (eg, registered nurses and licensed practical nurses) require prerequisites and standardized competencies, the unlicensed direct care workforce receives limited formal education and highly variable on-the-job training. Studies and workforce reviews describe unlicensed direct care workers as underprepared for dementia-related agitation and aggression despite providing hands-on care to people with dementia [6,7]. In nursing education, dementia-specific content is inconsistently integrated into prelicensure curricula and taught under broad geriatric instruction rather than as a distinct competency, although proposed dementia-specific training programs have identified the management of aggression and agitation as a core competency for future training development [8].

There is a gap in the evidence on the effectiveness of training on the experience of physical or verbal aggression in nursing home or assisted living facility staff. Demonstrated differences in the effect of training programs are unclear, and studies are limited by small, nongeneralizable samples [9]. There is an opportunity to determine whether there is an association between training programs and the experience of physical or verbal aggression by staff in nursing homes and assisted living facilities using a survey that provides a nationally representative sample. Understanding whether training programs can mitigate occupational exposures to aggression may support interventions to reduce future workforce turnover [4].

The objective of this study was to leverage a newly released, nationally representative sample of the US paid dementia care workforce to determine the prevalence of physical and verbal aggression and whether there was an association between training programs and the frequency of physical or verbal aggression. We hypothesized that effective training would be associated with a reduction in the experience of physical and verbal aggression from residents. By identifying the association using a large, weighted, nationally representative dataset, we can inform policy, training standards, and workforce development to support the paid dementia care workforce in the United States.

Methods

Data

This study used wave 1 (2024) data from the nursing home and assisted living surveys collected by the National Dementia Workforce Study (NDWS) [10]. The NDWS is a multiyear, federally funded data infrastructure initiative designed to generate nationally representative data on the US dementia care workforce. The study conducted a 2-stage cluster sampling design of establishment- and individual-level participants from nursing homes and assisted living facilities to produce probability-based estimates of the national dementia care workforce population [11]. Survey weights were developed by NDWS statisticians and survey designers to permit nationally representative inference at the organization and staff levels, accounting for complex survey design, stratification, and clustering.

Ethical Considerations

This study was approved by the University of Alabama at Birmingham Institutional Review Board (IRB-300015390). This analysis and reporting adhered to the Reporting of Studies Conducted Using Observational Routinely-Collected Data (RECORD) statement for the reporting of studies conducted using observational data.

Dependent Variables

Our primary dependent variable was the presence or absence of physical or verbal aggression. To classify physical or verbal aggression, we used question 42 on the nursing home and assisted living surveys, which reads “In your job at this community over the past year, how often have you experienced the following?” We assessed the presence of physical aggression using subitem I, which reads “Hitting or other physical aggression from residents,” and verbal aggression using subitem J, which reads “Yelling or other verbal aggression from residents.” Potential responses were “never,” “rarely,” “sometimes,” and “often.” We dichotomized these variables as “not frequently” (encompassing “never” or “rarely”) and “frequently” (encompassing “sometimes” or “often”) because the original distribution did not meet the proportional odds assumptions of the Brant test, precluding analysis via linear or ordinal logistic regression.

Training was operationalized to reflect respondents’ perceived preparedness to manage aggressive behaviors in residents with dementia rather than the presence or absence of specific training modalities. Although the NDWS wave 1 surveys include items assessing exposure to formal and informal dementia-specific training (items NH4 and NH5 for nursing homes and AL4 and AL5 for assisted living facilities), these items ask broadly about training. Because the primary focus of this analysis was preparedness to handle aggressive behaviors, we used the global preparedness items (NH6 and AL6), which ask respondents whether their training prepared them to care for residents with dementia. These items capture respondents’ self-assessment of readiness rather than training format and were therefore better aligned with our analytic objective of examining perceived preparedness for challenging behavioral situations.

XSL•FO
RenderX

Independent Variables

Independent variables included sex, age, race, ethnicity, feeling well-prepared by training to manage participant behaviors, rurality, and licensure. Licensure was determined by currently holding a registered nurse or licensed practical nurse license, whereas "nonlicensed" encompassed all other nonlicensed staff, including certified nursing assistants, home health aides, and medication assistants. We also included variables for directly working with residents and being a double-duty caregiver, where a staff member also provided care for an adult child at home.

Table 1 shows the variables and their item numbers and operationalizations.

Table . Variables, item numbers, and operationalizations.

Analytic variable	NDWS[a] item number	Item wording (abbreviated)
Licensure	NH1 or AL1	Do you currently hold any professional license(s)?
Direct care	NH35 or AL35	Do you provide direct care to residents with Alzheimer's disease or other dementias?
Educational level	NH3 or AL3	What is the highest level of education you have completed?
Sex	NH56 or AL56	What is your sex? (Male or female; masked in PUF[b])
Age (mean, SD, and range)	NH48 or AL48	What year were you born? (Used to calculate age; masked in PUF)
Race	NH50 or AL50	What is your race? (Categories masked or collapsed in PUF)
Ethnicity	NH49 or AL49	Are you Hispanic, Latino/a, or of Spanish origin?
Family caregiver	NH37a-f or AL37a-f	Are you currently or have you ever been a family caregiver for a person with Alzheimer's disease or other dementias? (Item series)
Physical aggression	NH38a-i or AL38a-i	How often have residents you care for done the following... (eg, hit, kicked, or physically hurt someone)?
Verbal aggression	NH38a-i or AL38a-i	How often have residents you care for yelled at, threatened, or verbally abused staff?
Training preparedness	NH6 or AL6	Overall, how well did your training prepare you to care for residents with Alzheimer's disease or other dementias?
Rurality	Derived (facility level)	Rural vs nonrural classification derived from facility geographic indicators described in the NDWS wave 1 survey report (not a single staff-reported item)

[a]NDWS: National Dementia Workforce Study.

[b]PUF: Public Use File.

Statistical Analysis

Overview

Descriptive statistics were used to summarize the characteristics of the included formal caregivers. Means and SDs were reported for continuous variables, whereas frequencies and percentages were reported for categorical variables. All descriptive estimates account for the complex survey design using the NDWS-provided sampling weights. All analyses were conducted in R (3.4.2; R Foundation for Statistical Computing). Missing data were handled via listwise deletion and were not imputed given the large, weighted sample size and the potential for weighted imputation to bias results.

Univariate Modeling

Chi-square and 2-tailed *t* tests were used to examine the association between the independent variables and a dichotomous high vs low frequency of experiencing physical or verbal aggression from residents.

Regression Modeling

Penalized logistic regression models using the Firth bias-reduced likelihood were fit to determine the strength and direction of the effect of feeling well-prepared by training programs on the dichotomous high vs low frequency of experiencing physical or verbal aggression from residents in nursing homes or assisted living facilities. NDWS-provided survey weights were used to represent the national dementia care workforce population. Clustering was accounted for in the survey object by specifying the primary sampling unit rather than a random effect included

XSL•FO
RenderX

in the regression model as the cluster is a product of the NDWS sampling scheme rather than a multilevel structure of interest. Variance inflation factors were used to identify potentially problematic collinearity, with a variance inflation factor threshold of 10 or higher identified as problematic collinearity.

Supplementary Analysis

We compared respondents included in the regression models (complete cases) with the full survey sample for our variables of interest, incorporating survey weights. We also conducted an additional supplementary regression model that included ethnicity, despite ethnicity exhibiting potentially problematic collinearity, due to the established importance of ethnicity in the dementia care workforce [12,13].

Results

Overview

The dementia care workforce in nursing homes (n=896,320) was primarily female (828,558/896,320, 92.4%), White (440,158/896,320, 49.1%), and non-Hispanic (778,249/896,320, 86.8%), with a mean age of 41.49 (SD 12.65) years. Less than half (365,452/896,320, 40.8%) reported frequently experiencing physical aggression from residents, whereas more than half (470,536/896,320, 52.5%) reported frequently experiencing verbal aggression from residents. More than two-thirds (614,561/896,320, 68.6%) reported feeling that their training prepared them well to manage resident behaviors.

The dementia care workforce in assisted living facilities (n=738,856) was primarily female (631,070/738,856, 85.4%), White (490,181/738,856, 66.3%), and non-Hispanic (650,020/738,856, 88%), with a mean age of 41.36 (SD 13.94) years. More than one-quarter (198,964/738,856, 26.9%) reported frequently experiencing physical aggression from residents, whereas just under one-third (223,343/738,856, 30.2%) reported frequently experiencing verbal aggression from residents. Almost three-quarters (527,621/738,856, 71.4%) reported feeling that their training prepared them well to manage resident behaviors.

Table 2 shows the participant characteristics.

XSL•FO
RenderX

Table . Participant characteristics.

	Assisted living facilities (weighted n=738,856)	Nursing homes (weighted n=896,320)	Assisted living facilities (unweighted n=447)	Nursing homes (unweighted n=394)
Licensure, n (%)				
Licensed	74,140 (18.9)	295,052 (34.3)	60 (13.4)	126 (41.9)
Nonlicensed	318,778 (81.1)	563,909 (65.7)	387 (86.6)	233 (59.1)
Direct care, n (%)				
Direct care to residents	663,615 (89.9)	838,753 (93.6)	388 (88.8)	365 (92.6)
Not providing direct care	75,241 (10.2)	57,566 (6.4)	59 (11.2)	29 (7.4)
Educational level, n (%)				
Associate degree	76,224 (10.3)	96,090 (10.7)	45 (10.1)	47 (11.9)
Bachelor's degree or higher	23,025 (3.1)	99,805 (11.1)	47 (10.5)	50 (12.7)
Lower than college	601,914 (81.5)	525,309 (58.6)	312 (70.0)	229 (58.1)
Vocational diploma or certificate	37,633 (5.1)	175,116 (19.5)	42 (9.4)	68 (17.3)
Sex, n (%)				
Female	631,070 (86.5)	828,558 (93.5)	388 (88.8)	353 (91.9)
Male	98,639 (13.5)	58,064 (6.5)	49 (11.2)	31 (8.1)
Age (y)				
Mean (SD)	41.36 (13.94)	41.49 (12.65)	40.596 (14.104)	40.222 (12.721)
Range	20.0-80.0	20.0-73.0	20-80	20-73
Race, n (%)				
Black or African American	126,423 (17.1)	367,459 (41)	83 (18.6)	104 (26.4)
Other	122,251 (16.5)	88,703 (9.9)	113 (25.3)	38 (9.6)
White	490,181 (66.3)	440,158 (49.1)	251 (56.2)	252 (64)
Ethnicity, n (%)				
Not Hispanic	650,020 (88.8)	778,249 (87.4)	356 (80)	351 (89.5)
Hispanic	82,350 (11.2)	112,085 (12.6)	89 (20)	41 (10.5)
Family caregiver, n (%)				
No	564,091 (76.4)	690,171 (77.3)	340 (76.2)	309 (78.6)
Yes	174,723 (23.6)	202,720 (22.7)	106 (23.8)	84 (21.4)
Physical aggression, n (%)				
Rarely or never	539,892 (73.1)	530,868 (59.2)	297 (66.4)	257 (65.2)
Sometimes or often	198,964 (26.9)	365,452 (40.8)	150 (33.6)	137 (34.8)
Verbal aggression, n (%)				
Rarely or never	515,513 (69.8)	425,784 (47.5)	265 (59.3)	201 (51)
Sometimes or often	223,343 (30.2)	470,536 (52.5)	182 (40.7)	193 (49)
Training preparedness, n (%)				
Not well-prepared	211,235 (28.6)	281,758 (31.4)	131 (29.3)	121 (30.7)
Well-prepared	527,621 (71.4)	614,561 (68.6)	316 (70.7)	273 (69.3)
Rurality, n (%)				
Not rural	484,502 (65.6)	675,002 (75.3)	429 (96)	313 (79.4)
Rural	254,354 (34.4)	221,318 (24.7)	18 (4)	81 (20.6)

XSL•FO
RenderX

Univariate Modeling

In univariate modeling in nursing homes, race (*P*=.03) and training preparedness (*P*=.002) were associated with experiencing physical aggression. Family caregiver status (*P*=.04) and training preparedness (*P*=.01) were associated with experiencing verbal aggression.

In univariate modeling in assisted living facilities, race (*P*=.002) and training preparedness (*P*<.001) were associated with experiencing physical aggression. Sex (*P*=.02), race (*P*=.002), and training preparedness (*P*<.001) were associated with experiencing verbal aggression.

Table 3 shows the univariate modeling results.

Table . Univariate modeling results.

	Assisted living facilities, *P* value		Nursing homes, *P* value	
	Physical aggression	Verbal aggression	Physical aggression	Verbal aggression
Age (mean and SD)	.49	.36	.18	.35
Licensure	.29	.31	.21	.30
Direct care role	.01	.01	.67	.35
Sex	.05	.02	.49	.27
Race	.002	.002	.03	.14
Ethnicity	.08	.33	.23	.36
Family caregiver	.38	.31	.30	.04
Training preparedness	<.001	<.001	.002	.01
Rural	.79	.50	.33	.82

Regression Modeling

In the fully adjusted model estimating the likelihood of reporting frequent physical aggression in nursing homes, feeling well-prepared by training for managing resident behaviors (odds ratio [OR] 0.40, 95% CI 0.223-0.711; *P*=.01) was associated with a lower likelihood of reporting frequent physical aggression.

In the fully adjusted model estimating the likelihood of reporting frequent verbal aggression in nursing homes, feeling well-prepared by training for managing resident behaviors (OR 0.50, 95% CI 0.299-0.844; *P*=.02) was associated with a lower likelihood of reporting frequent verbal aggression.

In the fully adjusted model estimating the likelihood of reporting physical aggression in assisted living facilities, feeling well-prepared by training for managing resident behaviors (OR 0.14, 95% CI 0.036-0.556; *P*=.01) and being a family caregiver (OR 0.19, 95% CI 0.043-0.778; *P*=.03) were associated with a lower likelihood of reporting frequent physical aggression.

In the fully adjusted model estimating the likelihood of reporting verbal aggression in assisted living facilities, working in a direct care role (OR 5.09, 95% CI 1.223-21.194; *P*=.03) was associated with a higher likelihood of reporting frequent verbal aggression, and feeling well-prepared by training for managing resident behaviors (OR 0.07, 95% CI 0.014-0.354; *P*<.001) and being a family caregiver (OR 0.18, 95% CI 0.061-0.516; *P*<.001) were associated with a lower likelihood of verbal aggression.

Table 4 shows the regression modeling results.

XSL•FO
RenderX

Table . Regression modeling results.

Term	Assisted living facilities				Nursing homes			
	Physical aggression		Verbal aggression		Physical aggression		Verbal aggression	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Intercept	0.52 (0.066-4.052)	.54	1.31 (0.153-11.254)	.81	3.99 (0.973-16.384)	.07	3.50 (1.041-11.776)	.06
Licensed	0.34 (0.067-1.72)	.20	0.33 (0.044-2.425)	.29	1.32 (0.612-2.864)	.48	1.18 (0.605-2.286)	.64
Direct care role	4.86 (0.761-31.079)	.11	5.09 (1.223-21.194)	.03	0.97 (0.282-3.365)	.97	1.22 (0.287-5.207)	.79
Male sex	2.79 (0.8-9.723)	.12	2.68 (0.872-8.24)	.10	0.57 (0.182-1.77)	.34	0.45 (0.166-1.122)	.13
Age	0.99 (0.954-1.034)	.76	0.99 (0.938-1.035)	.56	0.98 (0.951-1.005)	.13	0.98 (0.956-1.01)	.23
Non-White race	1.85 (0.452-7.582)	.40	2.54 (0.416-15.378)	.32	0.63 (0.345-1.142)	.14	0.62 (0.37-1.042)	.09
Family caregiver	0.19 (0.043-0.778)	.03	0.18 (0.061-0.516)	<.001	1.52 (0.857-2.709)	.17	2.33 (0.857-6.315)	.11
Training preparedness (well-prepared)	0.14 (0.036-0.556)	.01	0.07 (0.014-0.354)	<.001	0.40 (0.223-0.711)	.01	0.50 (0.299-0.844)	.02
Rurality	1.24 (0.264-5.817)	.79	0.89 (0.176-4.546)	.89	0.84 (0.317-2.206)	.72	1.13 (0.763-1.659)	.56

Supplementary Modeling

Sensitivity Analysis

In the sensitivity analysis, we did not find statistically substantial differences between the complete-case sample included in the regression modeling and those who were missing data for key demographic variables (Table 5).

Table . Sensitivity analysis comparing complete cases.

Variable	Nursing homes, *P* value	Assisted living facilities, *P* value
Age	.14	.64
Sex	.23	.13
Race	.62	.23
Training preparedness	.55	.91
Ethnicity	.24	.32

Regression Modeling

In the supplementary analysis of assisted living facilities, feeling that training prepared them well for managing resident behaviors (OR 0.13, 95% CI 0.04-0.56; *P*=.01) and family caregiver status (OR 0.20, 95% CI 0.05-0.75; *P*=.03) were associated with a lower likelihood of reporting frequent physical aggression. Hispanic ethnicity (OR 14.03, 95% CI 2.77-71.01; *P*<.001) was associated with a higher likelihood of reporting frequent physical aggression.

In the supplementary analysis of assisted living facilities, a direct care role (OR 4.74, 95% CI 1.21-18.68; *P*=.04) was associated with a higher likelihood of reporting frequent verbal aggression. Family caregiver status (OR 0.18, 95% CI 0.06-0.51; *P*<.001) and feeling prepared by training (OR 0.07, 95% CI 0.01-0.37; *P*=.01) were associated with a lower likelihood of reporting frequent verbal aggression.

In the supplementary analysis of nursing homes, feeling prepared by training (OR 0.48, 95% CI 0.26-0.89; *P*=.03) was associated with a lower likelihood of reporting frequent verbal aggression.

Table 6 shows the supplementary regression modeling results.

XSL•FO
RenderX

Table . Supplementary regression modeling results.

Term	Assisted living facilities				Nursing homes			
	Physical aggression		Verbal aggression		Physical aggression		Verbal aggression	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Intercept	0.05 (0.00 - 0.48)	.02	0.58 (0.04 - 8.78)	.70	2.05 (0.60 - 7.02)	.27	2.14 (0.58 - 7.95)	.27
Licensed	0.27 (0.04 - 1.76)	.19	0.3 (0.03 - 2.69)	.30	1.39 (0.57 - 3.38)	.48	1.2 (0.59 - 2.44)	.62
Direct care role	4.25 (0.90 - 20.12)	.08	4.74 (1.21 - 18.68)	.04	1.06 (0.33 - 3.34)	.93	1.3 (0.33 - 5.14)	.71
Male sex	2.44 (0.67 - 8.84)	.19	2.51 (0.81 - 7.79)	.12	0.54 (0.16 - 1.75)	.31	0.43 (0.16 - 1.16)	.12
Age	1 (0.96 - 1.04)	.91	0.99 (0.94 - 1.04)	.65	0.98 (0.95 - 1.00)	.10	0.98 (0.96 - 1.01)	.21
Non-White race	2.23 (0.52 - 9.51)	.29	2.84 (0.44 - 18.32)	.28	0.65 (0.32 - 1.30)	.24	0.63 (0.36 - 1.13)	.14
Hispanic ethnicity	14.03 (2.77 - 71.01)	<.001	2.4 (0.47 - 12.25)	.30	2.18 (0.52 - 9.12)	.30	1.74 (0.62 - 4.85)	.31
Family caregiver	0.20 (0.05 - 0.75)	.03	0.18 (0.06 - 0.51)	<.001	1.48 (0.86 - 2.56)	.18	2.26 (0.85 - 6.00)	.12
Training preparedness (well-prepared)	0.13 (0.04 - 0.56)	.01	0.07 (0.01 - 0.37)	.01	0.38 (0.18 - 0.78)	.17	0.48 (0.26 - 0.89)	*.03*
Rurality	0.98 (0.21 - 4.46)	.98	0.81 (0.15 - 4.36)	.81	0.83 (0.32 - 2.18)	.71	1.14 (0.78 - 1.65)	.51

Discussion

Overview of Findings

In both nursing homes and assisted living facilities, training programs for handling resident behaviors were associated with a lower likelihood of experiencing physical aggression, and in assisted living facilities, these training programs were associated with a lower likelihood of experiencing verbal aggression.

Comparison With Prior Literature

Prior work emphasizes the importance of dementia-specific training to manage resident behaviors. Livingston et al [14] demonstrated in a systematic review of 33 studies that aggression can be mitigated through caregiver activities, especially person-centered approaches to care that systematically support resident behaviors. Other reviews have found that aggression prevention in residential facilities requires staff to be equipped with knowledge and skills to intervene early to prevent situational escalation into physically or verbally aggressive behaviors, although Seitz et al [15] concluded that large-scale studies of staff training in nursing homes and residential facilities should examine whether these results are robust beyond protocol-driven interventions. Our results indicate that, on a nationally representative scale, there is an association between a reduced frequency of aggressive behaviors and perception that trainings prepared staff to manage participant behaviors. While the cross-sectional nature of the NDWS precludes causal speculation, other work has hypothesized that training that includes early recognition and intervention for problematic behaviors can reduce aggression in residents with dementia [16,17].

Our finding that being a double-duty caregiver, or caring for a person at home, was associated with a reduced likelihood of experiencing aggressive behaviors in assisted living facilities was initially surprising. However, this correlation could be due to their experience in understanding and recognizing escalating cascades of behaviors when caring for another person in the home [18]. Limited literature exists to determine the specific skills that double-duty caregivers apply in both home and long-term care settings, although a future study could determine the influence of home-based caregiving on long-term care strategies to manage aggression.

Limitations

The large, well-conducted survey methodology and robust weighting procedures provide strength to our findings. However, care is needed in the interpretation of results due to the cross-sectional nature of the study; causality cannot be assumed. Our findings indicate that perceived training efficacy for managing resident behaviors was linked to a decrease in the frequency of aggressive incidents; however, the specific elements of the training contributing to this effect remain undetermined. On the other hand, staff experiencing lower aggression may create the perception that their training sufficiently prepared them to manage resident behaviors. Future interventional work should determine the causal path.

XSL•FO
RenderX

Additionally, the study power was limited by the initial sampling frame rather than the weighted sample. There could have been aspects of the study sample that biased the results, although the robust weighting procedures and clustering in the survey design should mitigate this effect. We treated training preparedness as dichotomous due to small cell sizes in the "not at all prepared" group. While this speaks to the level of training received at nursing homes and assisted living facilities, future work should determine whether there is a stepwise change between similar categories as there could be a meaningful difference between feeling not at all prepared and somewhat prepared.

Conclusions

In a nationally representative sample of the dementia care workforce, effective training programs were associated with worker-reported reduced experiences of physical and verbal aggression in both nursing homes and assisted living facilities. The specific components of training programs that may support effectiveness remain unknown.

Acknowledgments

The authors acknowledge the considerable efforts of the National Dementia Workforce Study (NDWS; U54AG084520) in the collection of the data.

Funding

This research was supported by the Advancing Workforce Analysis and Research for Dementia (AWARD) Network, a National Institute on Aging (Grant: 1R24AG077014). The views expressed in written materials do not necessarily reflect the official policies of the Department of Health and Human Services (HHS) or the National Institute on Aging.

Data Availability

Data access is restricted by National Dementia Workforce Study (NDWS) policy. To access data, a request should be submitted through the NDWS. The analytic code is available subsequent to data access. The datasets generated or analyzed during this study are available in the National Dementia Workforce Study repository via LINKAGE [19].

Conflicts of Interest

None declared.

References

1. Lachs MS, Rosen T, Teresi JA, et al. Verbal and physical aggression directed at nursing home staff by residents. J Gen Intern Med 2013 May;28(5):660-667. [doi: 10.1007/s11606-012-2284-1] [Medline: 23225256]
2. Zeller A, Hahn S, Needham I, Kok G, Dassen T, Halfens RJ. Aggressive behavior of nursing home residents toward caregivers: a systematic literature review. Geriatr Nurs 2009;30(3):174-187. [doi: 10.1016/j.gerinurse.2008.09.002]
3. Whall AL, Colling KB, Kolanowski A, et al. Factors associated with aggressive behavior among nursing home residents with dementia. Gerontologist 2008 Dec;48(6):721-731. [doi: 10.1093/geront/48.6.721] [Medline: 19139246]
4. Yan E, Wan D, To L, et al. Staff turnover intention at long-term care facilities: implications of resident aggression, burnout, and fatigue. J Am Med Dir Assoc 2024 Mar;25(3):396-402. [doi: 10.1016/j.jamda.2023.10.008] [Medline: 37972647]
5. Searby A, Snipe J, Maude P. Aggression management training in undergraduate nursing students: a scoping review. Issues Ment Health Nurs 2019 Jun;40(6):503-510. [doi: 10.1080/01612840.2019.1565874] [Medline: 30958085]
6. Gaugler JE, Hobday JV, Robbins JC, Barclay MP. Direct care worker training to respond to the behavior of individuals with dementia: the CARES® Dementia-Related Behavior™ online program. Gerontol Geriatr Med 2016;2:2333721415626888. [doi: 10.1177/2333721415626888] [Medline: 26894209]
7. Yeh J, Beld M, Pond B, et al. Competency-based training boosts dementia knowledge and skills in home care workers. Alzheimers Dement 2025 Jun;21(6):e70323. [doi: 10.1002/alz.70323] [Medline: 40538020]
8. Mayo AM. Dementia-specific nursing care competencies for nursing education and long-term care practice. Nurs Clin North Am 2022 Jun;57(2):217-232. [doi: 10.1016/j.cnur.2022.02.004] [Medline: 35659984]
9. Jutkowitz E, Brasure M, Fuchs E, et al. Care-delivery interventions to manage agitation and aggression in dementia nursing home and assisted living residents: a systematic review and meta-analysis. J Am Geriatr Soc 2016 Mar;64(3):477-488. [doi: 10.1111/jgs.13936] [Medline: 27000321]
10. Novak P. The National Dementia Workforce Study: perspective from the National Institute on Aging. J Am Geriatr Soc 2025 Nov;73(11):3354-3358. [doi: 10.1111/jgs.70065] [Medline: 40874428]
11. Wagner J, Wagner LM, Zimmerman S, et al. The National Dementia Workforce Study: the plan for organization sample frames and data collection. J Am Geriatr Soc 2025 Nov;73(11):3336-3343. [doi: 10.1111/jgs.70036] [Medline: 40906878]
12. Estrada LV, Agarwal M, Stone PW. Racial/ethnic disparities in nursing home end-of-life care: a systematic review. J Am Med Dir Assoc 2021 Feb;22(2):279-290.e1. [doi: 10.1016/j.jamda.2020.12.005] [Medline: 33428892]

XSL•FO
RenderX

13. Martinez IL, Acosta Gonzalez E, Black L. Services for Latino family dementia caregivers: is the workforce prepared? J Appl Gerontol 2024 Sep;43(9):1251-1258. [doi: 10.1177/07334648241234982] [Medline: 38454617]
14. Livingston G, Kelly L, Lewis-Holmes E, et al. Non-pharmacological interventions for agitation in dementia: systematic review of randomised controlled trials. Br J Psychiatry 2014 Dec;205(6):436-442. [doi: 10.1192/bjp.bp.113.141119] [Medline: 25452601]
15. Seitz DP, Brisbin S, Herrmann N, et al. Efficacy and feasibility of nonpharmacological interventions for neuropsychiatric symptoms of dementia in long term care: a systematic review. J Am Med Dir Assoc 2012 Jul;13(6):503-506.e2. [doi: 10.1016/j.jamda.2011.12.059] [Medline: 22342481]
16. Chrzescijanski D, Moyle W, Creedy D. Reducing dementia-related aggression through a staff education intervention. Dementia 2007;6(2):271-286. [doi: 10.1177/1471301207080369]
17. Kunik ME, Stanley MA, Shrestha S, et al. Aggression prevention training for individuals with dementia and their caregivers: a randomized controlled trial. Am J Geriatr Psychiatry 2020 Jun;28(6):662-672. [doi: 10.1016/j.jagp.2020.01.190] [Medline: 32115311]
18. Gaugler JE, Reese M, Mittelman MS. Effects of the NYU caregiver intervention-adult child on residential care placement. Gerontologist 2013 Dec;53(6):985-997. [doi: 10.1093/geront/gns193] [Medline: 23339050]
19. Maust DT, Wagner J, Zimmerman S, et al. The National Dementia Workforce Study: overview. J Am Geriatr Soc 2025 Nov;73(11):3318-3326. [doi: 10.1111/jgs.70061] [Medline: 40888611]

Abbreviations

NDWS: National Dementia Workforce Study
OR: odds ratio
RECORD: Reporting of Studies Conducted Using Observational Routinely-Collected Data

Edited by E Borycki; submitted 24.Feb.2026; peer-reviewed by E Yan, G Davlyatov; revised version received 26.May.2026; accepted 09.Jun.2026; published 10.Jul.2026.

Please cite as:
Bratches R, Backhouse T, Ge S, Hebdon M, Riegel D, Stein N, Trainum K, Jablonski R
Training Programs for Dementia Care Staff and Aggression in Nursing Homes: Secondary Analysis of a Cross-Sectional Online Survey
JMIR Nursing 2026;9:e94077
URL: https://nursing.jmir.org/2026/1/e94077
doi:10.2196/94077

© Reed Bratches, Tamara Backhouse, Shaoqing Ge, Megan Hebdon, Dalton Riegel, Noah Stein, Katie Trainum, Rita Jablonski. Originally published in JMIR Nursing (https://nursing.jmir.org), 10.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

A One-Year Study Using Digital Biomarkers From Sensing Technologies to Assess Changes in Physical Activity Levels and Sleep Quality in Nursing Home Residents With Dementia: Observational Study

Lydia D Boyle[1]; Monica Patrascu[1]; Bettina S Husebo[1]; Kristoffer Haugarvoll[2]; Ole Martin Steihaug[3]; Brice Marty[1]

[1]Center for Elderly and Nursing Home Medicine, University of Bergen, Årstadveien 17, Bergen, Vestland, Norway

[2]Neuro SysMed, Haukeland University Hospital, Bergen, Vestland, Norway

[3]Haraldsplass Diakonale Sykehus, Bergen, Vestland, Norway

Corresponding Author:
Lydia D Boyle
Center for Elderly and Nursing Home Medicine, University of Bergen, Årstadveien 17, Bergen, Vestland, Norway

Abstract

Background: Proxy-rated questionnaires remain the standard for the assessment of activity and sleep for people with dementia living in nursing homes. Sensing technologies, such as wearables, can generate continuous data that provide quantitative insights into daily activities and behavioral and psychological symptoms, such as sleep disturbance. This study explores the use of sensing technologies in the detection of changes in physical activity levels and sleep behaviors over time.

Objective: This study aims to explore the long-term capabilities of multimodal sensing technologies for assessing physical activity levels and sleep quality using selected digital biomarkers for nursing home residents with dementia. Objectives were for observation to be aligned with real-world conditions in which such sensing technologies would be applied within a nursing home environment and to assess whether distinct differences in selected digital biomarkers can be observed accurately and reliably longitudinally.

Methods: This study included 11 participants (79 - 93 y) recruited from 2 dementia care units in Norway. A smartwatch (Garmin Vivoactive 5 or Garmin Venu 3) and a radar-based system (Vital Things, Somnofy) were used to collect 7 days and 6 nights of data on physical activity levels and sleep quality at baseline, 6 months, and 1 year. The Personal Self-Maintenance Score and Neuropsychiatric Inventory–Nursing Home Version (nighttime behaviors section K) were also administered. Digital biomarkers included Euclidean norm minus one (ENMO), sleep efficiency (SE), wake after sleep onset (WASO), sleep regulatory index (SRI), sleep fragmentation index (SFI), total sleep time (TST), and time out of bed (no presence).

Results: A total of 9 participants were included in the final analysis. Differences were found in nighttime ENMO (P=.01) and in 4 of the sleep biomarkers: TST (P=.02), SE (P=.02), WASO (P=.01), and SRI (P=.01). The long-term reliability of the group ENMO was poor (intraclass correlation coefficient 0.00 - 0.02); however, it was moderate to strong (0.58 - 0.79) between individuals at each time point. The adherence and acceptability of the technologies was high (88%-96%), and the application of the devices was well tolerated by the participants, with no adverse events.

Conclusions: The use of sensing technologies could enable more objective, data-driven future care models for people with dementia residing in nursing homes; however, the results emphasized in this study require the recommendation for cautious, well-designed use of digital biomarkers for clinical decision-making.

(JMIR Nursing 2026;9:e95194) doi:10.2196/95194

KEYWORDS

sensors; sensing technology; wearables; sleep disturbance; physical activity; nursing homes; dementia

Introduction

Background

The global prevalence of dementia is expected to reach 139 million by 2050 [1], and in countries such as Norway, more than two-thirds of nursing home residents live with dementia [2]. Monitoring neuropsychiatric symptoms (NPSs) and changes in physical activity typically relies on proxy-rated assessments, which are limited by subjectivity, retrospective bias, and low sensitivity to subtle or short-term change [3]. As health care systems face increasing workforce shortages, there is a growing need for scalable, objective approaches that can support continuous monitoring and individualized care [4]. Recent

XSL•FO
RenderX

advances in sensing technologies, including wrist-worn wearables and contactless radar systems, offer minimally obtrusive, real-time data collection in naturalistic environments and may help address these challenges.

Physical inactivity is widespread in nursing home environments [5,6], with daily routines dominated by sedentary behaviors such as sitting, resting, and sleeping [7]. Despite the importance of physical activity for maintaining function, there are no standardized recommendations for activity dosage in people with advanced dementia [8]. Although the relationship between daytime activity and sleep quality remains insufficiently understood, several studies suggest that higher daytime activity may be associated with improved sleep [9,10]. Functional ability is typically assessed using proxy-rated activities of daily living (ADL) questionnaires, but these tools are retrospective and infrequently administered, reducing their use for detecting early or subtle decline, which is an important limitation given the strong association between functional loss and increased morbidity and mortality [11].

Accelerometry provides several methods for quantifying physical activity, including step counts, intensity, frequency, and vector magnitude [12], and common metrics derived from raw acceleration data include Euclidean norm minus one (ENMO), mean amplitude deviation (MAD), Monitor-Independent Movement Summary (MIMS), activity index, and Rate of Change Acceleration Movement (ROCAM) [12,13]. A recent feasibility study in nursing home residents with dementia demonstrated that commercial wearables could accurately classify physical activity over a 24-hour period [14], supporting the promising potential of sensing technologies in dementia care [15,16]. However, research validating activity classification and intensity thresholds in older adults with dementia remains limited, as demonstrated in a systematic review by Gorman et al [13] exploring physical activity and sedentary behaviors in older adults, finding that there are no standardized accelerometry-based thresholds for moderate-to-vigorous or sedentary activity, with most cut-offs validated in healthy adults rather than cognitively impaired populations. Research using wrist-worn accelerometers and actigraphy in a cohort of 415 nursing home residents in Australia demonstrated that these devices could capture patterns of inactivity and sleep duration in this population. The study also highlighted practical challenges, particularly around device accuracy and participant adherence, which was particularly low (<50%), concluding that additional work is needed to refine sensor placement, device characteristics, and methodological approaches for use in a nursing home setting [7].

Sleep disturbances are prevalent in nursing home residents and affect up to 90% of those with dementia [17], with common issues including insomnia, sundowning, restless legs, and fragmented sleep. Sleep quality is particularly compromised in people with dementia residing in nursing homes [18,19], and although nonpharmacological interventions are recommended [20], psychotropic medications remain widely used and may worsen sleep or increase the risk of injury [21], contributing to further NPSs such as agitation, anxiety, and nighttime restlessness [18]. Key sleep metrics, including efficiency, latency, wakefulness after sleep onset, fragmentation, regulation, and total sleep, decline with age and are further impaired for people with dementia [22]. Autonomic regulation of sleep also changes with age and neurodegeneration, particularly regarding parasympathetic nervous system involvement in the dampening of the sympathetic nervous system during deep sleep [23]. Sleep durations of 6 to 9 hours have been associated with improved parasympathetic function in older adults with neurodegenerative disease [24].

Polysomnography (PSG) remains the gold standard for sleep assessment but is often impractical and burdensome for people with dementia in nursing homes [25]. Consumer wearables provide sleep estimates using accelerometry and photoplethysmography, but they frequently overestimate sleep in individuals with sleep dysfunction [26] and vary in accuracy [27]. Newer commercial wearable models offer sleep staging, and although sleep sensitivity tends to be adequate, the poor detection of awake periods reduces accuracy, making these devices not yet adequate for research or clinical standards and not validated in an older adult population [28]. Radar-based systems [29] offer a contact-free alternative capable of monitoring sleep in free-living environments. These systems produce digital biomarkers comparable to PSG and show promising accuracy [30], though further research is needed to confirm their reliability and clinical use for longitudinal monitoring in real-world dementia care settings.

Study Aim and Objectives

This study aims to explore the long-term capabilities of multimodal sensing technologies for assessing physical activity levels and sleep quality using selected digital biomarkers for nursing home residents with dementia. Objectives were for observation to be aligned with real-world conditions in which such sensing technologies would be applied within a nursing home environment, and to assess whether distinct differences in selected digital biomarkers can be observed accurately and reliably longitudinally.

Methods

Design and Setting

This study is related to an observational pilot study: digital phenotyping for changes in activity at the end of life in people with dementia (DIPH.DEM): an observational trial based on sensing technology. DIPH.DEM is a cornerstone study for a larger observational study titled “Decoding Death and Dying in People With Dementia by Digital Thanotyping” (5-D; REK 657596, NEM 2023/166). DIPH.DEM was designed to improve the overall quality, recruitment strategies, and efficacy of the larger 5-D study. This paper is an extension of a previous article [31], which used the DIPH.DEM baseline data (description in Table 1) to investigate the association between daytime activity and sleep quality in people with dementia residing in a nursing home. The DIPH.DEM cohort and study design have previously been described in detail in 2 previous articles [31,32].

Participants (n=11) 65 years or older were recruited from 2 dementia wards in a Norwegian nursing home. Eligibility for the study included that the participants had lived in the nursing home for more than 6 weeks and had a previous diagnosis of

XSL•FO
RenderX

moderate-to-severe cognitive impairment or dementia, confirmed within the medical record and/or using the Clinical Dementia Rating (CDR) scale administered at baseline. Exclusion criteria included the presence of delirium at the baseline assessment and a current life expectancy of less than six weeks as suggested by the multidisciplinary nursing home team. Data were collected from February 2024 to May 2025. The proxy-rated questionnaires and continuous digital monitoring using sensing technologies, Garmin Vivoactive 5 or Garmin Venu 3 and Vital Things Somnofy, were conducted at 3 distinct time points: baseline, 6 months, and 1 year.

Table . Sample characteristics from the DIPH.DEM[a] study (N=11) [31].

Demographics of study participants	Values
Age (y), mean (SD; range) (n=11)	84.0 (6.1; 77 - 93)
Gender, n	
Female	8
Male	3
Months living in the nursing home, mean (SD; range) (n=11)	17.6 (12.5; 3 - 42)
Dementia as primary diagnosis, n	6
Alzheimer dementia, n	3
Parkinson disease, n	1
Clinical Dementia Rating score, mean (SD; range) (n=11)	2.2 (0.7; 1 - 3)
Mild, n	4
Moderate, n	5
Severe, n	2
Personal Self Maintenance Score, mean (SD; range) (n=11)	15.7 (5.7; 8 - 23)
Comorbidities (GMHR[b]), mean (SD; range) (n=11)	2.6 (0.7; 1 - 3)
Neuropsychiatric Inventory (NPI-NH[c])–12 subsections, mean (SD; range)	
Agitation (n=5)	2.2 (2.2; 1 - 6)
Delusion (n=4)	3.0 (2.4; 1 - 6)
Hallucinations (n=3)	3.0 (2.6; 1 - 6)
Depression (n=4)	2.0 (1.4; 1 - 4)
Anxiety (n=3)	1.3 (0.6; 1 - 2)
Apathy (n=4)	5.8 (4.6; 1 - 12)
Irritability (n=5)	3.6 (2.9; 1 - 8)
Euphoria (n=1)	8.0 (0.0; 8 - 8)
Disinhibition (n=4)	1.5 (0.6; 1 - 2)
Aberrant motor behavior (n=4)	1.0 (0.0; 1 - 1)
Nighttime behavior disturbances (n=6)	4.7 (2.9; 1 - 8)
Appetite or eating (n=4)	4.0 (2.9; 1 - 8)
Total medication use, mean (SD; range) (n=11)	7.6 (3.4; 2 - 13)
Psychotropic drug use, mean (SD; range) (n=8)	1.9 (1.4; 1 - 5)

[a]DIPH.DEM: digital phenotyping for changes in activity at the end of life in people with dementia.

[b]GMHR: General Medical Health Rating.

[c]NPI-NH: Neuropsychiatric Inventory-Nursing Home.

Ethical Considerations

Verbal and written informed consent were obtained in direct conversation with the residents who demonstrated sufficient capacity to consent. For those lacking this ability, we obtained consent from the resident's legal guardian, usually a family member, after explaining the aims and protocol of the study. The DIPH.DEM study was approved by the Regional Committee for Medical and Health Research Ethics (REK) in Norway in October 2023 (approval number 634938).

XSL•FO
RenderX

Traditional Questionnaires

Traditional outcome measures used in this study were the Neuropsychiatric Inventory–Nursing Home version (NPI-NH) [33], assessing symptoms of sleep disturbance (Nighttime Behaviours section K), and the Physical Self Maintenance Scale (PSMS) [34], assessing functional activity levels and adverse effects of inactivity. The NPI-NH (section K) [35] includes questions about difficulty falling asleep, frequent nightly awakenings, wandering, dressing in the middle of the night, and waking earlier than normal to begin the day. The PSMS [36] includes functional activities of ambulation, bathing, grooming, dressing, toileting, and feeding. To define the severity and level of disturbances within this study, we chose to use a global scoring system from 6 to 30 (6 questions and 5 possible answers), with lower scores indicating lower dependency and higher scores indicating higher dependency.

Additional baseline measurements reported in this study were the General Medical Health Rating (GMHR) [37] for the assessment of general health and comorbidities, 4AT (Assessment Test for delirium) [38], and the CDR [39]. The CDR questionnaire (range 0 - 3) is a widely accepted global scale developed to clinically identify the presence of dementia and stage its severity (score of >0.5: cognitive impairment/likely dementia; 1: mild; 2: moderate; 3: severe dementia). Within our study, participants were included in the study with a score equal to or greater than 1 on the CDR. The GMHR (range 1 - 4) assesses the number and impact of comorbidities (1=very good; 2=good; 3=moderate, and 4=bad health). The 4AT Test (range 0 - 12) is a screening tool for delirium and consists of 4 questions evaluating alertness, cognitive function, and acute change, with a score of 4 or above indicating potential delirium. All traditional questionnaires used were previously validated and provided in the Norwegian language.

Sensing Technologies and Digital Biomarkers

The sensing technologies used in the study were Somnofy (Vital Things AS, Trondheim, Norway) [40] version 0.7 with sleep algorithm 1.0 and Garmin (Garmin Ltd., Switzerland) Vivoactive5 smartwatch [26] firmware version 11.14 or Garmin Venu 3 smartwatch (X) firmware version 12.11. Digital biomarkers used in this study were acceleration-based physical activity estimates using ENMO, sleep efficiency (SE), wake after sleep onset (WASO), sleep regulatory index (SRI), total sleep time (TST), sleep fragmentation index (SFI), and no presence (time out of bed) (Table 1). Garmin Vivoactive5 was used for data collection at baseline and 6 months and upgraded to Garmin Venu3 at the 1-year time point for improved battery capacity and data storage capabilities. Both smartwatch models contained accelerometers used to capture raw acceleration signals. The Venu3 includes additional sensors, a barometric altimeter, gyroscope, and the updated Elevate V5 heart-rate sensor, while the earlier model uses the V4 sensor. Only accelerometer-derived raw data were analyzed in this study. Participants wore the device continuously for 7 days at each data collection point (baseline, 6 months, 1 year), including during sleep and bathing.

Garmin wearables, including Vivoactive and Venu models, demonstrate acceptable validity and strong interdevice reliability for step counting and movement assessment in laboratory and free-living settings [41,42]. Prior studies report accurate step detection across walking speeds and age groups, including older adults with slower gait patterns [43,44]. Garmin devices have also shown reliable estimation of physical activity using raw acceleration data [45], and ENMO-based thresholds applied to wrist-worn Garmin sensors provide high sensitivity and specificity for distinguishing sedentary and light-intensity behaviors [46].

Data Extraction and Preprocessing

Physical Activity Levels: Euclidean Norm Minus One—Acceleration Data

Raw data parameters for the smartwatches were synced, paired, and retrieved using a hub through Fitrockr [47], a data broker based in Berlin, Germany, and acceleration was collected at 25 Hz. The extracted data were imported and converted in MATLAB and STATA-compatible types. Based on the Unix time timestamps to Unix time, the data were segmented into 14-hour periods (07:00:00 AM-08:59:59 PM) representing daytime; the first and last partial days were excluded to only consider full day recordings (24 h), resulting in 5 days and 6 nights of total acceleration data per participant. Nighttime physical activity was considered within the remaining timeframe of 09:00:00 PM-06:59:59 AM and based on nightly routines at the nursing home. The raw acceleration data were converted to the universal metric ENMO for physical activity analysis. ENMO is the vector magnitude of the 3 raw signals (x, y, z) minus 1. One-minute periods were used during daytime (07:00:00 AM-08:59:59 PM), nighttime (09:00:00 PM-06:59:59 AM), and daily 24-hour (07:00:00 AM-06:59:59 AM) measurements (mean, median, SD). The thresholds for physical activity levels using ENMO classifications were determined according to previous literature [46] and included sedentary activity to time (<18 milli-g), light physical activity (18 - 22 milli-g), light-to-moderate activity (>22 - 60 milli-g), and moderate-to-vigorous physical activity (MVPA; >60 milli-g). The calculation used for conversion of raw accelerometry data to ENMO (units: milli-g) was as follows:

$$ENMOi=x(t)i2+y(t)i2+ z(t)i2 -1000$$

Digital Sleep Biomarkers

Sleep metrics (SE, SRI, WASO, time in no-presence, and SFI) were exported directly from the Somnofy platform. All Somnofy sleep data were manually reviewed to remove daytime sleep episodes occurring between 07:00 AM and 08:59 PM. These daytime events were documented separately to avoid influencing nighttime sleep estimates. Daytime sleep is not reported here, as the radar system only detects sleep occurring within its sensing range (ie, in bed) and therefore does not holistically capture daytime sleep, which may occur elsewhere than the bed. The nighttime window was selected to align with the routine administration of psychotropic medications for sleep disturbance (typically around 09:00 PM) and standard bedtime procedures at the nursing home. Participants were excluded if fewer than 10 hours of nighttime sleep data (one full night) were not available.

XSL•FO
RenderX

SE was calculated as the proportion of total sleep time relative to time in bed, expressed as a percentage from 0 to 100. SRI, also provided by Somnofy, reflects the consistency of sleep-wake patterns across days and requires at least 4 consecutive nights of data; values are expressed as a percentage and range from 0 to 100. WASO represents the cumulative minutes awake after initial sleep onset. The SFI score quantifies sleep fragmentation as the number of transitions from deep to light sleep per hour of sleep and is expressed as a percentage from 0 to 100. Somnofy additionally reports time in "no presence," defined here as time or minutes spent outside the sensor's effective range (>3 m), which was interpreted as time out of bed during the nighttime period. All wake periods overlapping with "no presence" were manually checked to confirm accuracy.

Statistical Analysis

Descriptive statistics and narrative synthesis are provided including mean, median, minimum, maximum, and SD for all relative variables. The data were tested for normality and analyzed using MATLAB R2023a and STATA SE18.5; outliers were detected using box plots. The data were dependent, nonparametric, and not normally distributed resulting in the use of the Friedman test to assess differences between the 3 data collection time points. The Friedman test is used for one-way repeated measures ANOVA by ranks. A post hoc evaluation of the specific differences between the time points in which the null hypothesis was rejected was performed using the Wilcoxon signed-rank test. Kendall *W* interpretation of effect is as follows: 0.1 - 0.29 small effect size, 0.3 - 0.5 moderate effect size, and >0.5 large effect size. A sensitivity analysis was performed as a subanalysis of the significant results by removing influential extreme outliers, which were defined as totals beyond 1.5 times the IQR. The results were considered statistically significant at a $P<.05$, and at a χ^2 (2 *df*) critical value of ≥5.991. Reliability as temporal stability is reported using an intraclass correlation coefficient (ICC 1,1 and ICC 3,1) defined with scores ranging from 0 to 1 and score interpretation: <0.5 as poor agreement, 0.5 to <0.75 as moderate agreement, 0.75 to <0.90 as good agreement, and 0.9 to 1.0 as excellent agreement.

Results

Eleven participants were recruited from February 2024 to May 2025. Two participants passed away during the data collection period, resulting in final analysis of 9 participants. Data were missing for the following reasons: charging of the smartwatch (1 - 2 hours on day 4) and periodic nonadherence from the participants (removal of the smartwatch for periods ≤2 hours). Adherence was defined as the total wear time (Garmin VivoActive5 or Venu3) calculated after the initial data extraction of the first and last day, and was 96% at baseline, 88% at 6 months, and 94% at 1 year. For the acceleration data, after the removal of partial days within the available data profiles (first and last day), there were a total of 135 acquisition periods available for the final analysis. Additionally, a total of 189 data acquisition periods, consisting of data from a total of 219 nights, were used for analysis of the sleep quality digital variables.

Participant characteristics and demographics at baseline for the DIPH.DEM study (n=11) can be found in Table 1. Participants were mostly female (8/11), age ranging from 77 to 93, living in the nursing home for a mean of 17.6 (SD 12.5) months, and with an average score of 2.2 on the CDR indicating a moderate level of dementia. Participants had a mean score of 2.6 (SD 0.7) on the GMHR, suggesting a moderate level of comorbidities consisting of greater than or equal to 3 chronic illnesses and concomitant medication use. Chronic diagnoses within the group included Alzheimer disease, Parkinson disease, diabetes, and heart disease. At baseline, assessed by the NPI-NH (section K for nighttime behavior), almost half (6/11) of the group was classified as having nighttime sleep behaviors defined within the study as frequent nighttime activity and awakenings. Total daily polypharmacy within the group ranged between 2 and 13 total medications per person, and the use of daily psychotropics due to sleep disturbances within the group profile ranged from 1 to 5 prescriptions per participant, per night.

There were longitudinal differences in nighttime physical activity levels from baseline to 1 year ($P=.01$), with a declining trend (33% decrease from baseline to 1 year) in all ENMO means, most pronounced from 6 months to the 1-year time points (Figures 1-3). The majority of the group's 24-hour activity intensities, as illustrated in Figure 1, were either sedentary threshold or below (<18 milli-g) or between sedentary time and light-to-moderate physical activity thresholds (18 - 60 milli-g). In Figure 2, we observe a loss of day or nighttime rhythm or flattening of ENMO intensities, most distinct at the 1-year time point, indicating that daytime ENMO more closely resembled the nighttime ENMO intensities at the 1-year time point. The temporal stability (ICC) within and between individuals over time was poor (0.00 - 0.02; Table 2), suggesting that the acceleration data (ENMO) was not a reliable measurement longitudinally; however, variance between-persons at a single time point ranged from moderate to good (0.58 - 0.79), suggesting that ENMO can distinguish between individuals cross-sectionally in a stable and meaningful way. After removing outliers and aggregating repeated measurements, there was no statistically significant change in nighttime ENMO across time points; however, the effect size remained moderate, indicating a meaningful degree of within-participant change over time.

Four of the digital sleep quality biomarkers were found to have differences from baseline to 1 year: TST ($P=.02$), SE ($P=.02$), WASO ($P=.01$), and SRI ($P=.01$; Table 2, Figures 4 and 5). This suggests that these sleep quality digital biomarkers are sensitive enough over time to be used as a tool for identifying discrete changes in sleep quality. Temporal stability (ICC) of the sleep digital biomarkers over time was moderate to strong (0.54 - 0.92; Table 2), with the most reliable longitudinal metrics being SE (0.92) and WASO (0.78). This indicates that these sleep biomarkers, derived from the Somnofy radar-based sensors, provided clinically meaningful data over time with minimal residual variance. A sensitivity analysis was performed for subanalysis of TST, SE, WASO, and SRI. After the removal of extreme outliers, SE and SRI showed no significant changes over time but moderate within-participant effect sizes. Both

XSL•FO
RenderX

WASO and TST demonstrated longitudinal change, each with large effect sizes indicating substantial within-participant shifts.

Figure 1. Distribution of 24-hour physical activity levels (Euclidean norm minus one [ENMO]) for one participant at baseline (yellow), 6 months (green), and 1 year (gray). Explanation of thresholds is as follows: space between orange lines=sedentary time (<18 milli-g), between red and blue lines=light physical activity (18 - 22 milli-g), between blue and black lines=light-to-moderate physical activity (>22 - 60 milli-g), all else from black lines=moderate-to-vigorous activity (>60 milli-g).

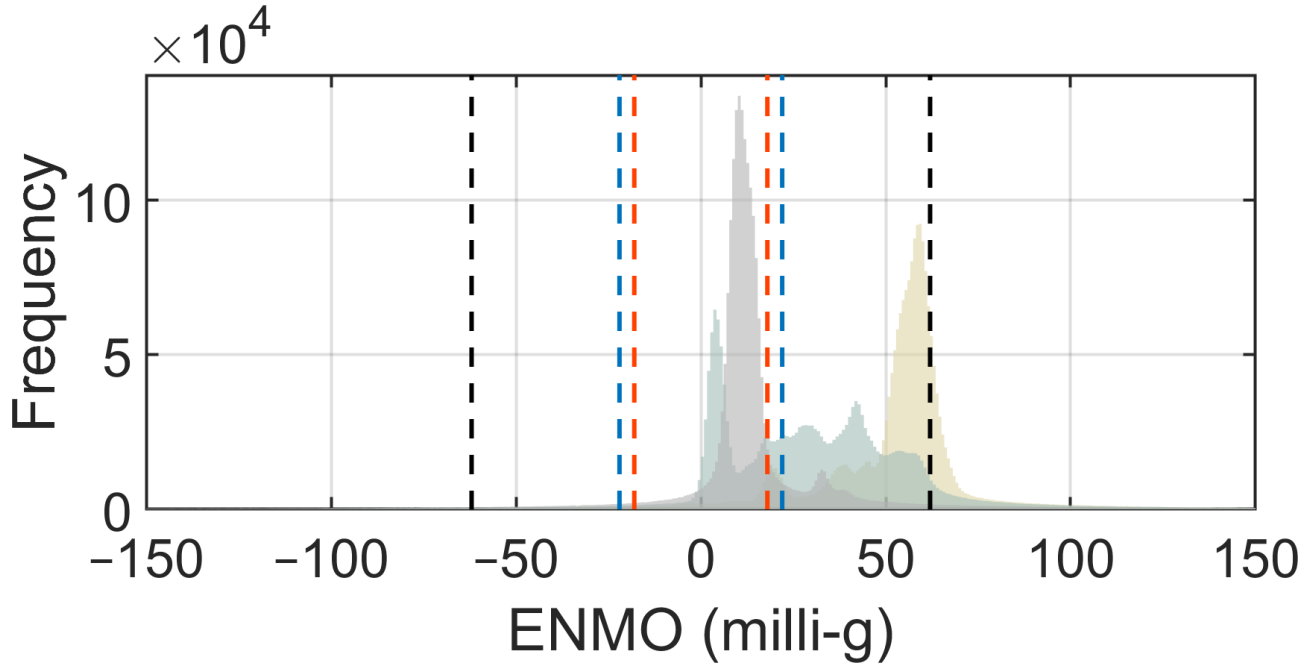
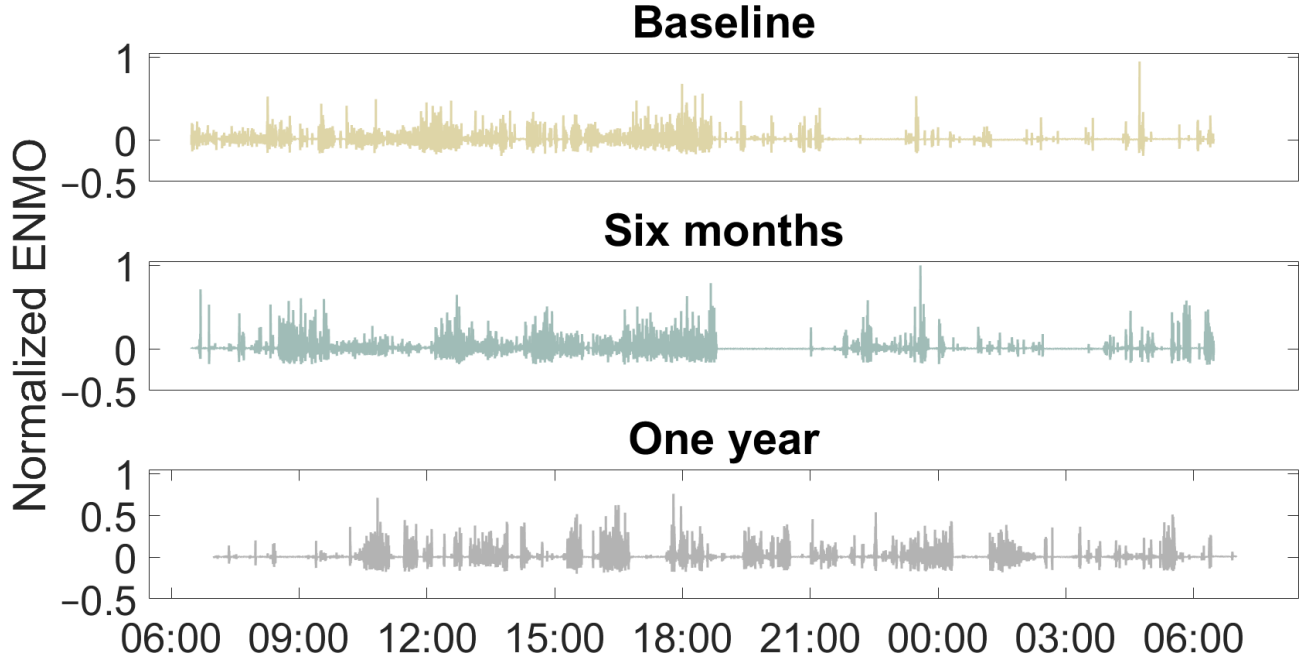
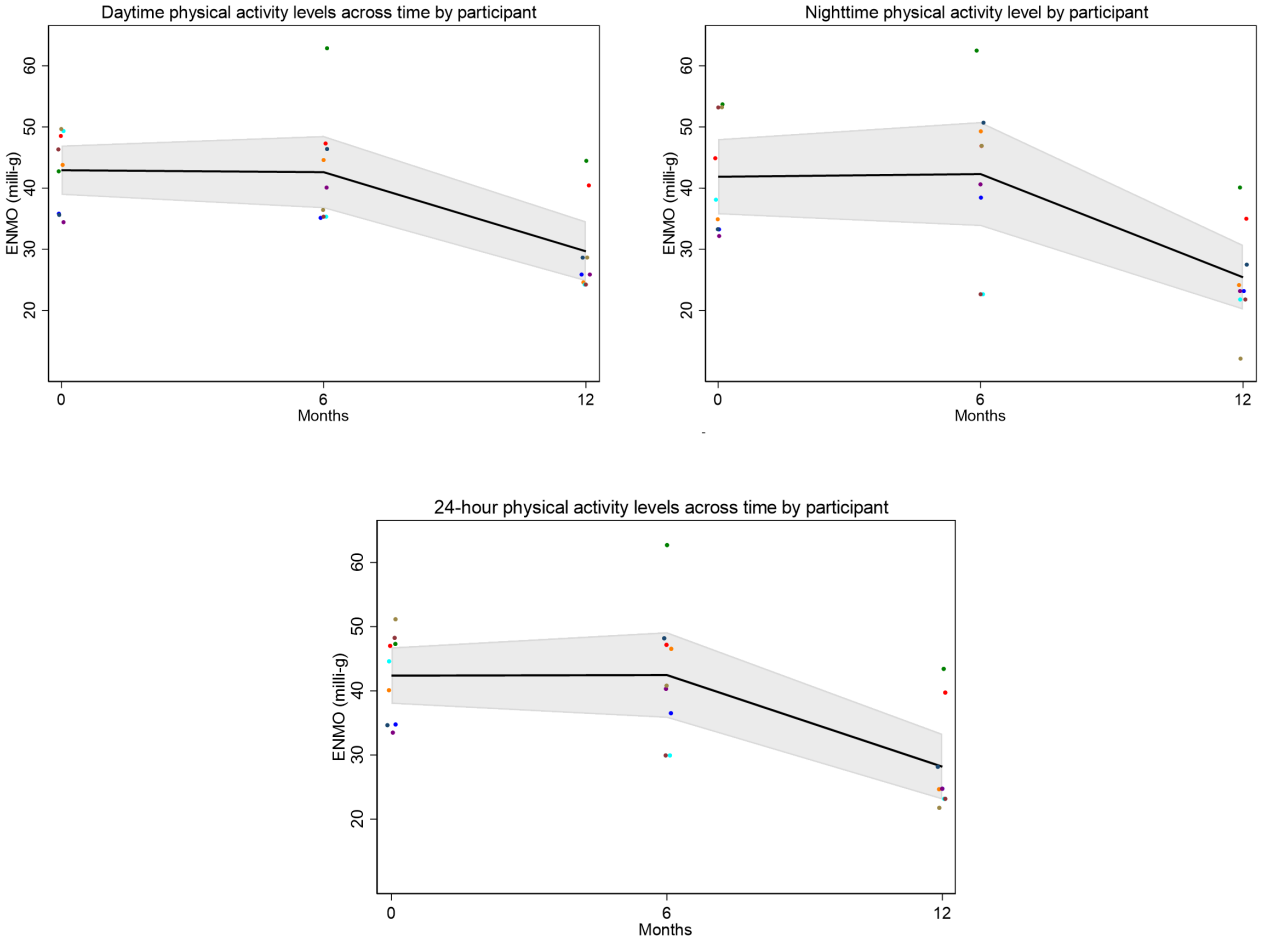


Figure 2. A 24-hour physical activity level. Definition of normalized Euclidean norm minus one (ENMO): the EMNO values were divided by the maximum value of the 3 days represented and are a snapshot for the same participant (Figure 1) at baseline, 6 months, and 1 year.



XSL•FO
RenderX

Figure 3. Mean differences of rank variance in Euclidean norm minus one (ENMO) representing physical activity levels, daytime, nighttime, 24 hours from baseline, 6 months, and 1 year. Each participant is represented by a different color "dot," and the figures include all data points used for analysis for each participant. One "dot" then represents 1 data point. The gray band characterizes the IQR at 25/75%.



XSL•FO
RenderX

Table . Longitudinal changes (baseline, 6 months, and 1 year) from selected sensing and traditional outcome measures (N=9)[a].

Selected variables: longitudinal results at baseline, 6 months, and 1 year	Baseline		Six months		One year		Friedman test			Reliability
	Mean (SD)	Median (range)	Mean (SD)	Median (range)	Mean (SD)	Median (range)	Chi-square (*df*)	Kendall *W*	*P* value	ICC[b]
NPI-NH-K[c] nighttime behaviors	5.4 (2.6)	6 (1 - 8)	9 (0.0)	9 (9)	3 (2.4)	2.5 (1 - 6)	11.0 (2)	0.46	.20	0.20
Personal Self Maintenance Scale—Physical Activity	17.3 (5.0)	17 (10 - 23)	16.4 (4.6)	18 (8 - 21)	17.8 (4.5)	20 (8 - 22)	11.5 (2)	0.48	.18	0.46
Daytime activity (ENMO[d]) Garmin (milli-g)[e] 07:00 AM-08:59 PM	42.9 (6.2)	43.8 (34.4 - 50)	42.6 (9.1)	40.1 (35.1 - 63)	29.7 (7.5)	25.9 (24.2 - 45)	44.9 (2)	0.38	.24	0.01
Nighttime activity (ENMO) Garmin (milli-g) 09:00 PM-06:59 AM	41.9 (9.4)	38.1 (32.2 - 54)	42.3 (13.0)	47.0 (22.7 - 63)	25.4 (8.1)	23.2 (12 - 40)	63.4 (2)	0.56	.01[f]	0.00
24-hour activity (ENMO) Garmin (milli-g) midnight-11:59 PM	42.4 (6.7)	44.6 (33.5 - 51)	42.5 (10.2)	40.8 (30 - 63)	28.2 (7.9)	24.7 (21.8 - 43)	56.2 (2)	0.47	.05	0.02
Total sleep time (h)	10.0 (2.1)	10.6 (7 - 13)	7.1 (2.6)	7.7 (2 - 10)	7.7 (2.9)	8.5 (1 - 11)	18.3[f] (2)	0.76[f]	.02[f]	0.54
Sleep efficiency (%)	65.4 (18.6)	72 (22 - 83)	66.7 (19.3)	76 (21 - 82)	66.9 (19.7)	75 (18 - 81)	18.7[f] (2)	0.78[f]	.02[f]	0.92
Wake after sleep onset (h)	2.4 (1.4)	2.2 (0.2 - 5)	2.3 (1.1)	2.0 (1 - 4.2)	2.2 (1.4)	2.0 (0.5 - 5)	20.9[f] (2)	0.87[f]	.01[f]	0.78
Sleep regularity index (%)	72.6 (11.8)	71.3 (55 - 93)	68.9 (17.3)	75.7 (44 - 86)	71.2 (11.7)	69 (56 - 85)	19.0[f] (2)	0.79[f]	.01[f]	0.55
Sleep fragmentation index (%)	20.2 (13.8)	17.6 (5 - 54)	26.1 (8.6)	26.5 (11 - 37)	36.7 (16.2)	39.3 (12 - 63)	12.4 (2)	0.51	.14	0.17
Time out of bed (min)	12.7 (4.6)	11.2 (9 - 22)	12.8 (14.7)	9.4 (2.5 - 48)	5.0 (2.9)	5.0 (2 - 9)	12.9 (2)	0.54	.12	0.19

[a]No presence refers to the amount of time spent out of bed and/or out of sensor range, ICC ranges from 0 to 1, with 1 being perfect reliability or temporal stability.

[b]ICC: intraclass correlation coefficient.

[c]NPI-NH-K: Neuropsychiatric Inventory–Nursing Home–section K nighttime behaviors.

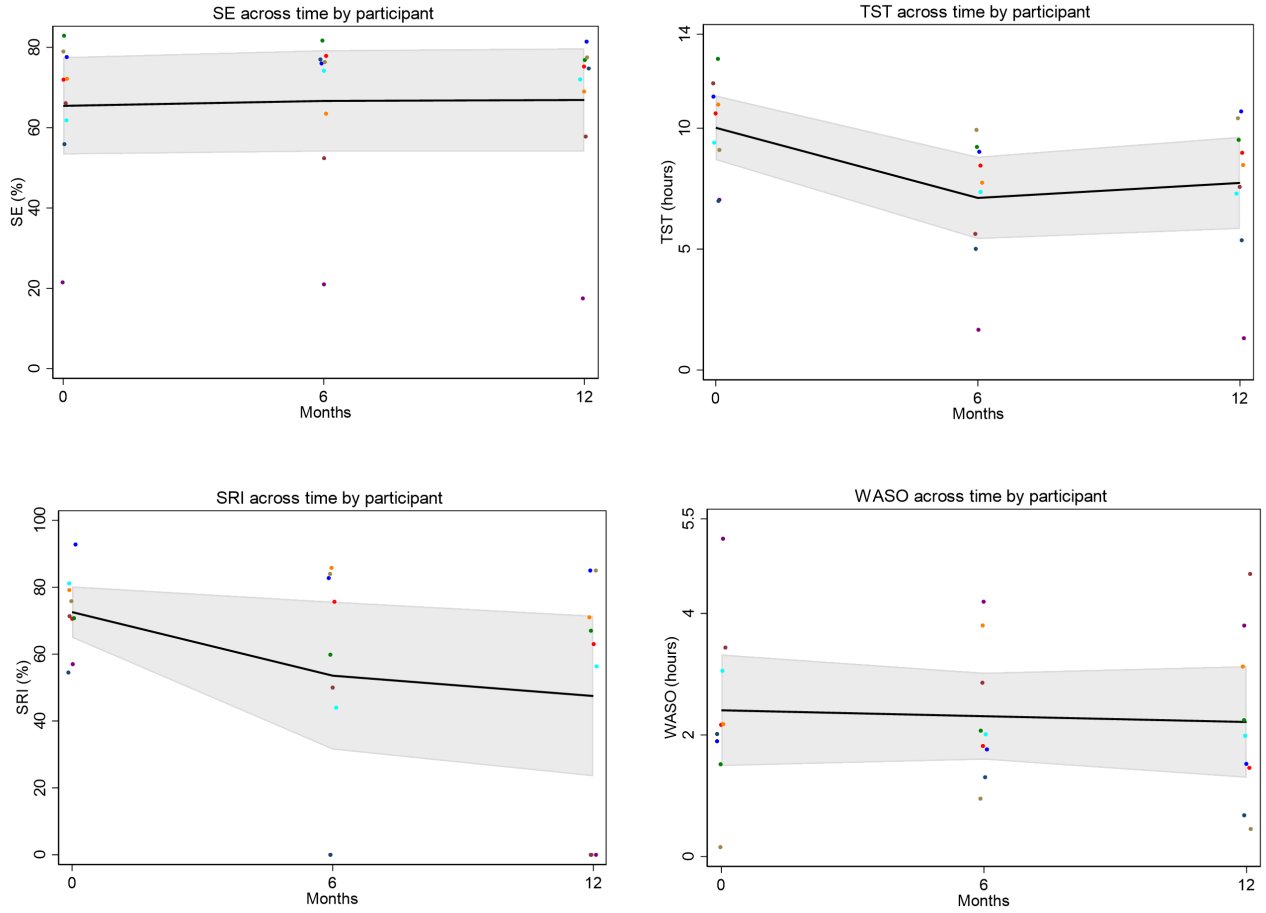
[d]ENMO: Euclidean norm minus one.

XSL•FO
RenderX

[e]milli-g: 1/1000 of the standard acceleration due to Earth's gravity (g).

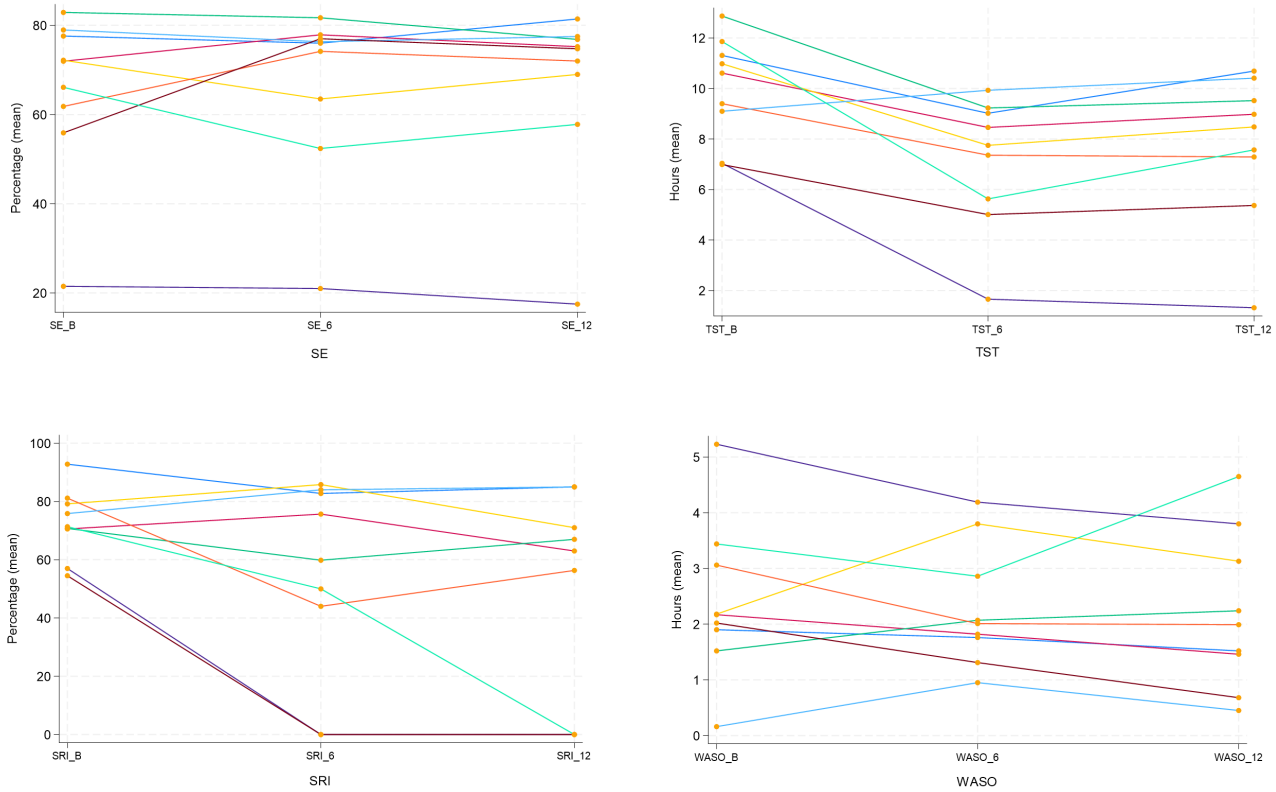
[f]$P<.05$, $\chi^2 \geq 5.991$, Kendall W interpretation: 0.1 - 0.29 small effect size, 0.3 - 0.5 moderate effect size, >0.5 large effect size.

Figure 4. Median differences of rank variance in total sleep time (TST), wake after sleep onset (WASO), sleep efficiency (SE), and sleep regulatory index (SRI), Friedman test from baseline, 6 months, and 1 year. Each participant is represented by a different color "dot," and the figures include all data points used for analysis for each participant. One "dot" then represents one data point. The gray band characterizes the IQR at 25/75%.



XSL•FO
RenderX

Figure 5. Longitudinal mean of individual trajectories at baseline (B), 6 months (6), and 1 year (12) for total sleep time (TST), wake after sleep onset (WASO), sleep efficiency (SE), and sleep regulatory index (SRI); each color represents a separate participant (n=9).



Discussion

Principal Findings

This study aimed to explore the use of sensing technologies in the assessment of longitudinal changes in physical activity levels and sleep quality in people with dementia living in a nursing home. We found that nighttime physical activity levels (ENMO) and 4 distinct digital sleep biomarkers (TST, SE, WASO, SRI) showed discrete changes across the 1-year data collection period, suggesting that these selected digital biomarkers could detect modest variations and declines in activity and sleep quality over time. Temporal stability of the sleep metrics was moderate to excellent; however, further inspection of the reliability of all longitudinal ENMO measured physical activity level data was poor. This suggests that the acceleration metric was unreliable as a longitudinal assessment tool and would be best applied to a single time-point, where reliability remained moderate to good or that for real-life use, additional sensor modalities and positions are required for better quality data. Adherence and acceptability of the selected technologies remained high across the 1-year period, with no adverse events, suggesting that the chosen devices and level of unobtrusiveness were a good match for the cohort. This study's findings are important in the illustration of the capabilities and limitations of sensing technologies and their corresponding digital biomarkers as clinical decision-making tools for people with dementia. It also demonstrates the trade-off between the level of unobtrusiveness and quality of data when choosing sensing technologies for use within more vulnerable groups.

The standardization and clarification of measurement context for digital biomarkers is needed, as emphasized in a recent meta-analysis [48] stating that accelerometry is increasingly being used to quantify physical activity, yet daily activity metrics produced using acceleration data remain heterogeneous and primarily reflect overall movement rather than specific health-related exertion. In our study, ENMO-derived activity levels distinguished long-term group differences; however, the longitudinal reliability of these measures was poor. This is consistent with previous work [49,50] stating that the most accurate results and models occur with thigh or low back placement of sensors for the monitoring of physical activity in older adults. Therefore, device placement and user burden are central considerations; in this study's nursing home population, a wrist-worn device was selected to minimize obtrusiveness and encourage adherence for older, potentially frail residents. While wrist-based wearables may be appropriate for single time point assessments, long-term monitoring may require alternative placements, such as the hip or low back, or integration with additional sensor modalities, to improve long-term reliability.

The World Health Organization (WHO) published the Global Action Plan on Physical Activity (GAPPA) [51,52], providing policy guidance for increased activity levels, reduced sedentary behavior, and tailored interventions across age groups. The report highlights limitations of self- and proxy-reported measures and advocates integrating wearable technologies into global physical activity surveillance. It also calls for standardized protocols to examine dose-response relationships between activity, sedentary behavior, and health outcomes, and for future research that enables data triangulation, refined estimates of total activity volume, and the identification of age-

XSL•FO
RenderX

and disease-specific associations [53]. This is a promising step toward future global clinical use and recommendations for standardized metrics of accelerometry-based data; however, focused research on more vulnerable groups, such as people with dementia, is still lacking.

The landscape of sleep-monitoring technologies for individuals with dementia has evolved rapidly, integrating passive and wearable systems tailored for long-term care and home use. Traditional actigraphy remains widely used to assess sleep-wake patterns and circadian disruptions in people with dementia [48]; recently, more advanced noncontact sensors, such as under-mattress sleep mats leveraging ballistocardiograph and respiration monitoring, and radar technology, are being deployed to track heart rate, breathing, movement, and fragmentation of sleep, reducing the need for wearable devices [27,30]. Due to the rapid development of new sensing technologies that rival gold standard PSG, a recent study emphasized the importance of the evaluation of novel sleep technologies conducted within longitudinal, real-world settings including people with advanced dementia, as a pathway for future device validation [30]. These recommendations reflect a shift toward more nuanced, long-term, and minimally burdensome monitoring modalities that are critical for capturing sleep changes in nursing-home and real-world settings, while addressing the unique compliance challenges of dementia care.

The fundamental differences between traditional proxy-rated questionnaires and digital biomarkers derived from sensing technologies may capture distinct phenomena and often yield divergent conclusions [54]. The NPI-NH and PSMS are established proxy-rated instruments commonly used in nursing home research to assess behavioral symptoms, functional ability, and sleep-related disturbances. However, these questionnaires do not capture the same dimensions of sleep quality represented by the digital measures used in this study. Differences in data capture frequency also limit comparability and validation between digital and traditional tools, as traditional questionnaires are typically completed every 30 to 90 days, while sensing technologies generate continuous, real-time data. Because proxy-rated tools and sensor-based measures assess different constructs and operate across different temporal scales, they are not directly interchangeable. Future research should incorporate assessment instruments that more closely correspond to digital sleep and activity metrics to strengthen construct validity and improve clinical interpretation. We cannot, therefore, conclude that sensing technologies provide superior assessments; rather, they can offer a unique definition of physical activity levels and sleep patterns through digital biomarkers, which can strengthen and fill in the blanks between periodic traditional proxy-rated assessments.

Due to rapid development and consequential obsolescence, the future validation of sensing technologies should involve repeated measures within the same modality to accelerate validation and contextualize digital biomarker results within target populations, such as people with dementia. Global consensus on the use of a standardized metric for accelerometry data and clinical translation of the corresponding thresholds should be prioritized to improve the generalization and applicability of results in real-world settings. Establishing population and context-specific thresholds, which account for age, health status, cultural, and environmental factors, will be essential to improving generalizability, applicability, and interpretability in future studies. Until these challenges are appropriately addressed, integrating proxy-rated assessments, inclusion of human-in-the-loop interpretation for data, and cautious interpretation of digital biomarkers is most appropriate for vulnerable populations such as people with dementia residing in nursing homes.

Although this study demonstrates the use of sensing technologies and digital biomarkers, it is not without limitations. The sample size, though appropriate for a pilot study design, is small, and the results should therefore be interpreted with caution. A sensitivity analysis was conducted to account for the small sample size, the influence of extreme values, and to clarify which effects were most robust. After removing outliers and aggregating repeated measurements, the overall pattern of change remained consistent with the original results, but only WASO and TST retained statistical significance, indicating that these domains showed the strongest and most reliable longitudinal shifts. Clinically, this suggests that while participants exhibited meaningful variability across physical activity metrics and multiple sleep parameters, the most stable and reproducible changes over time occurred in wakefulness and sleep duration, whereas other measures may require larger samples to detect effects with greater confidence. Regardless, further investigation is warranted based on these preliminary results. Further, this study uses the universal metric ENMO to interpret the acceleration data into physical activity levels; however, there is currently no standardized metric established for clinical translation of physical activity digital biomarkers, and another method could produce varied results. This study's objective was not to validate specific sensing technologies, and we would like to note that the nature of rapidly developing technologies creates a challenge for the timely validation of these novel devices and a gap between the use of these devices and digital biomarkers in research versus real-world applications.

Conclusions

Sensing technologies are promising for supporting more objective, data-driven care models in nursing homes. Acceleration-based metrics from wrist-worn devices appear most robust for cross-sectional assessment, whereas contactless radar-derived sleep biomarkers demonstrate stronger suitability and accuracy for both snapshot and longitudinal monitoring. Our findings underscore the need for cautious, context-aware interpretation of digital biomarkers. Further work is required to establish the accuracy and reliability of these sensing methods before they are incorporated into clinical decision-making frameworks.

XSL•FO
RenderX

Acknowledgments

We first need to thank the participants, their loved ones, and nursing home team for their invaluable contributions to this study. We would also like to thank Helse Vest and our partners at Neuro-SysMed and Haraldsplass Deaconess Hospital for their continued support of the project.

Grammarly was used to suggest language improvements in the manuscript.

Funding

Funding was provided by the Western Norway Regional Health Authority (Helse Vest RHF; sponsor protocol code: F-12829-D10484).

Data Availability

All data supporting the results and conclusions of this article will be made available by the authors, without undue reservation.

Authors' Contributions

All authors (LDB, MP, OMS, KH, BSH, BM) contributed to the study conception and design. LDB contributed to the initial writing of the manuscript. LDB and BM performed all data processing and analysis. All authors read, assisted with further editing, and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Emre M. Dementia associated with Parkinson's disease. Lancet Neurol 2003 Apr;2(4):229-237. [doi: 10.1016/s1474-4422(03)00351-x] [Medline: 12849211]
2. GjØra L, Strand BH, Bergh S, et al. Current and future prevalence estimates of mild cognitive impairment, dementia, and its subtypes in a population-based sample of people 70 years and older in Norway: the HUNT study. J Alzheimers Dis 2021;79(3):1213-1226. [doi: 10.3233/JAD-201275] [Medline: 33427745]
3. Lyketsos CG. Neuropsychiatric symptoms in dementia: overview and measurement challenges. J Prev Alzheimers Dis 2015 Sep;2(3):155-156. [doi: 10.14283/jpad.2015.60] [Medline: 26779454]
4. Giebel C, Readman MR, Godfrey A, Gray A, Carton J, Polden M. Geographical inequalities in dementia diagnosis and care: a systematic review. Int Psychogeriatr 2025 Jun;37(3):100051. [doi: 10.1016/j.inpsyc.2025.100051] [Medline: 39986949]
5. den Ouden M, Bleijlevens MHC, Meijers JMM, et al. Daily (in)activities of nursing home residents in their wards: an observation study. J Am Med Dir Assoc 2015 Nov 1;16(11):963-968. [doi: 10.1016/j.jamda.2015.05.016] [Medline: 26155723]
6. Baidya S, Connolly CJ, Petersen JM, et al. Barriers and facilitators to physical activity among older adults in residential aged care facilities: a mixed methods systematic review using the social ecological model. J Aging Health 2025 Dec;37(10):702-720. [doi: 10.1177/08982643241302209] [Medline: 39576885]
7. Moyle W, Jones C, Murfield J, et al. Levels of physical activity and sleep patterns among older people with dementia living in long-term care facilities: a 24-h snapshot. Maturitas 2017 Aug;102:62-68. [doi: 10.1016/j.maturitas.2017.05.015] [Medline: 28610686]
8. Olfermann R, Ebner-Priemer U, Reichert M, Giurgiu M. Comparability of accelerometry outcomes across popular metrics and widespread sensor positions. PLoS One 2025;20(12):e0337897. [doi: 10.1371/journal.pone.0337897] [Medline: 41335607]
9. Atoui S, Chevance G, Romain AJ, Kingsbury C, Lachance JP, Bernard P. Daily associations between sleep and physical activity: a systematic review and meta-analysis. Sleep Med Rev 2021 Jun;57:101426. [doi: 10.1016/j.smrv.2021.101426] [Medline: 33571893]
10. Liang Z, Zhao S, Tian S, et al. Determinants and interrelations of exercise effects on sleep quality: a multimethod meta-analysis. Sleep Med Rev 2026 Apr;86:102239. [doi: 10.1016/j.smrv.2026.102239] [Medline: 41650690]
11. Kim JH, Lee SB. Evaluation of activities of daily living: current insights and future horizons. Ann Geriatr Med Res 2025 Jun;29(2):143-158. [doi: 10.4235/agmr.24.0172] [Medline: 39945130]
12. Arvidsson D, Fridolfsson J, Börjesson M. Measurement of physical activity in clinical practice using accelerometers. J Intern Med 2019 Aug;286(2):137-153. [doi: 10.1111/joim.12908] [Medline: 30993807]
13. Gorman E, Hanson HM, Yang PH, Khan KM, Liu-Ambrose T, Ashe MC. Accelerometry analysis of physical activity and sedentary behavior in older adults: a systematic review and data analysis. Eur Rev Aging Phys Act 2014;11(1):35-49. [doi: 10.1007/s11556-013-0132-x] [Medline: 24765212]

XSL•FO
RenderX

14. Gawehns D, Portegijs S, van Beek APA, van Leeuwen M. Using consumer wearables to estimate physical activity of nursing home residents with dementia. Explor Digit Health Technol 2025;3:101137. [doi: 10.37349/edht.2025.101137]
15. Au-Yeung WTM, Miller L, Beattie Z, et al. Monitoring behaviors of patients with late-stage dementia using passive environmental sensing approaches: a case series. Am J Geriatr Psychiatry 2022 Jan;30(1):1-11. [doi: 10.1016/j.jagp.2021.04.008] [Medline: 34039534]
16. Verhoog S, Gubelmann C, Bano A, Muka T, Franco OH, Marques-Vidal P. Comparison of different software for processing physical activity measurements with accelerometry. Sci Rep 2023 Feb 18;13(1):2879. [doi: 10.1038/s41598-023-29872-7] [Medline: 36806337]
17. López-García A, López-Fernández RM, Martínez-González-Moro I. Analysis of sleep quality in people with dementia: a preliminary study. Gerontol Geriatr Med 2023;9:23337214231151473. [doi: 10.1177/23337214231151473] [Medline: 36726411]
18. Shin HY, Han HJ, Shin DJ, Park HM, Lee YB, Park KH. Sleep problems associated with behavioral and psychological symptoms as well as cognitive functions in Alzheimer's disease. J Clin Neurol 2014 Jul;10(3):203-209. [doi: 10.3988/jcn.2014.10.3.203] [Medline: 25045372]
19. Webster L, Costafreda Gonzalez S, Stringer A, et al. Measuring the prevalence of sleep disturbances in people with dementia living in care homes: a systematic review and meta-analysis. Sleep 2020 Apr 15;43(4):zsz251. [doi: 10.1093/sleep/zsz251] [Medline: 31633188]
20. Wilfling D, Calo S, Dichter MN, Meyer G, Möhler R, Köpke S. Non-pharmacological interventions for sleep disturbances in people with dementia. Cochrane Database Syst Rev 2023 Jan 3;1(1):CD011881. [doi: 10.1002/14651858.CD011881.pub2] [Medline: 36594432]
21. McCleery J, Sharpley AL. Pharmacotherapies for sleep disturbances in dementia. Cochrane Database Syst Rev 2020 Nov 15;11(11):CD009178. [doi: 10.1002/14651858.CD009178.pub4] [Medline: 33189083]
22. Li J, Vitiello MV, Gooneratne NS. Sleep in normal aging. Sleep Med Clin 2018 Mar;13(1):1-11. [doi: 10.1016/j.jsmc.2017.09.001] [Medline: 29412976]
23. Zoccoli G, Amici R. Sleep and autonomic nervous system. Curr Opin Physiol 2020 Jun;15:128-133. [doi: 10.1016/j.cophys.2020.01.002]
24. Liu CR, Yang CY, Sharma D, et al. Associations between sleep duration and autonomic nervous system regulation in patients with probable Alzheimer's disease: a cross-sectional pilot study. Clocks Sleep 2024 Sep 24;6(4):533-545. [doi: 10.3390/clockssleep6040035] [Medline: 39449309]
25. Boulos MI, Jairam T, Kendzerska T, Im J, Mekhael A, Murray BJ. Normal polysomnography parameters in healthy adults: a systematic review and meta-analysis. Lancet Respir Med 2019 Jun;7(6):533-543. [doi: 10.1016/S2213-2600(19)30057-8] [Medline: 31006560]
26. Kastelic K, Dobnik M, Löfler S, Hofer C, Šarabon N. Validity, reliability and sensitivity to change of three consumer-grade activity trackers in controlled and free-living conditions among older adults. Sensors (Basel) 2021 Sep 17;21(18):6245. [doi: 10.3390/s21186245] [Medline: 34577457]
27. Yoon H, Choi SH. Technologies for sleep monitoring at home: wearables and nearables. Biomed Eng Lett 2023 Aug;13(3):313-327. [doi: 10.1007/s13534-023-00305-8] [Medline: 37519880]
28. Cusack NM, Venkatraman PD, Raza U, Faisal A. Review—smart wearable sensors for health and lifestyle monitoring: commercial and emerging solutions. ECS Sens Plus 2024 Mar 1;3(1):017001. [doi: 10.1149/2754-2726/ad3561]
29. Heglum HSA, Kallestad H, Vethe D, Langsrud K, Sand T, Engstrøm M. Distinguishing sleep from wake with a radar sensor: a contact-free real-time sleep monitor. Sleep 2021 Aug 13;44(8):zsab060. [doi: 10.1093/sleep/zsab060] [Medline: 33705555]
30. Yamakawa M, Kang HS, Wang H, Konno R. Sleep quality assessment of adults in care settings using non-wearable sleep trackers: scoping review. Int J Nurs Pract 2024 Dec;30(6):e13240. [doi: 10.1111/ijn.13240] [Medline: 38317567]
31. Boyle LD, Patrascu M, Husebo BS, Steihaug OM, Haugarvoll K, Marty B. Use of digital biomarkers from sensing technologies to explore the relationship between daytime activity levels and sleep quality in nursing home residents with dementia: a proof-of-concept study. Sensors (Basel) 2025 Oct 29;25(21):6635. [doi: 10.3390/s25216635] [Medline: 41228858]
32. Boyle LD, Marty B, Haugarvoll K, Steihaug OM, Husebo BS. Use of sensing technologies to assess sleep quality and physical activity levels in nursing home residents with dementia taking nightly psychotropic drugs for sleep disturbance: a cross-sectional study. Front Aging 2026;7:1768935. [doi: 10.3389/fragi.2026.1768935] [Medline: 41822301]
33. Selbaek G, Kirkevold O, Sommer OH, Engedal K. The reliability and validity of the Norwegian version of the Neuropsychiatric Inventory, nursing home version (NPI-NH). Int Psychogeriatr 2008 Apr;20(2):375-382. [doi: 10.1017/S1041610207005601] [Medline: 17559707]
34. Edwards MM. The reliability and validity of self-report activities of daily living scales. Can J Occup Ther 1990 Dec;57(5):273-278. [doi: 10.1177/000841749005700507]
35. Saari T, Koivisto A, Hintsa T, Hänninen T, Hallikainen I. Psychometric properties of the Neuropsychiatric Inventory: a review. J Alzheimers Dis 2022;86(4):1485-1499. [doi: 10.3233/JAD-200739] [Medline: 32925068]

XSL•FO
RenderX

36. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [doi: 10.1093/geront/9.3_Part_1.179] [Medline: 5349366]
37. Lyketsos CG, Galik E, Steele C, et al. The General Medical Health Rating: a bedside global rating of medical comorbidity in patients with dementia. J Am Geriatr Soc 1999 Apr;47(4):487-491. [doi: 10.1111/j.1532-5415.1999.tb07245.x] [Medline: 10203127]
38. Tieges Z, Maclullich AMJ, Anand A, et al. Diagnostic accuracy of the 4AT for delirium detection in older adults: systematic review and meta-analysis. Age Ageing 2021 May 5;50(3):733-743. [doi: 10.1093/ageing/afaa224] [Medline: 33951145]
39. Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. Int Psychogeriatr 1997;9 Suppl 1:173-176. [doi: 10.1017/s1041610297004870] [Medline: 9447441]
40. Toften S, Pallesen S, Hrozanova M, Moen F, Grønli J. Validation of sleep stage classification using non-contact radar technology and machine learning (Somnofy®). Sleep Med 2020 Nov;75:54-61. [doi: 10.1016/j.sleep.2020.02.022] [Medline: 32853919]
41. Evenson KR, Spade CL. Review of validity and reliability of Garmin activity trackers. J Meas Phys Behav 2020 Jun;3(2):170-185. [doi: 10.1123/jmpb.2019-0035] [Medline: 32601613]
42. Fuller D, Colwell E, Low J, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR mHealth uHealth 2020 Sep 8;8(9):e18694. [doi: 10.2196/18694] [Medline: 32897239]
43. Bianchini E, Rinaldi D, De Carolis L, et al. Reliability of average daily steps measured through a consumer smartwatch in Parkinson disease phenotypes, stages, and severities: cross-sectional study. JMIR Form Res 2025 Mar 18;9:e63153. [doi: 10.2196/63153] [Medline: 40101170]
44. Höchsmann C, Knaier R, Eymann J, Hintermann J, Infanger D, Schmidt-Trucksäss A. Validity of activity trackers, smartphones, and phone applications to measure steps in various walking conditions. Scand J Med Sci Sports 2018 Jul;28(7):1818-1827. [doi: 10.1111/sms.13074] [Medline: 29460319]
45. White JW 3rd, Finnegan OL, Tindall N, et al. Comparison of raw accelerometry data from ActiGraph, Apple Watch, Garmin, and Fitbit using a mechanical shaker table. PLoS One 2024;19(3):e0286898. [doi: 10.1371/journal.pone.0286898] [Medline: 38551940]
46. Migueles JH, Cadenas-Sanchez C, Alcantara JMA, et al. Calibration and cross-validation of accelerometer cut-points to classify sedentary time and physical activity from hip and non-dominant and dominant wrists in older adults. Sensors (Basel) 2021 May 11;21(10):3326. [doi: 10.3390/s21103326] [Medline: 34064790]
47. Khaled M, Sareban M, Kreuzthaler M, Schulz S, Hussein R. Overview of existing tools for extracting health and fitness data from mHealth apps. Stud Health Technol Inform 2022 Jun 29;295:49-50. [doi: 10.3233/SHTI220657] [Medline: 35773803]
48. Cote AC, Phelps RJ, Kabiri NS, Bhangu JS, Thomas KK. Evaluation of wearable technology in dementia: a systematic review and meta-analysis. Front Med (Lausanne) 2020;7:501104. [doi: 10.3389/fmed.2020.501104] [Medline: 33505979]
49. Skjødt M, Brønd JC, Tully MA, et al. Moderate and vigorous physical activity intensity cut-points for hip-, wrist-, thigh-, and lower back worn accelerometer in very old adults. Scand J Med Sci Sports 2025 Jan;35(1):e70009. [doi: 10.1111/sms.70009] [Medline: 39753998]
50. Shi C, Babiker N, Urbanek JK, Grossman RL, Huisingh-Scheetz M, Rzhetsky A. Free-living wrist and hip accelerometry forecast cognitive decline among older adults without dementia over 1- or 5-years in two distinct observational cohorts. NPJ Aging 2022 Jun 6;8(1):7. [doi: 10.1038/s41514-022-00087-w] [Medline: 35927250]
51. Heuvelman F, Birkholz L, Tcymbal A, et al. The impact of public policy on socioeconomic equity in physical activity: a systematic review. Int J Behav Nutr Phys Act 2026 Feb 4;23(1):20. [doi: 10.1186/s12966-026-01880-6] [Medline: 41639714]
52. Physical activity measurement and surveillance in adults: report of a scoping and planning meeting, 27-28 November 2023. : World Health Organization; 2024 URL: https://iris.who.int/server/api/core/bitstreams/03f33b6a-0bae-4140-8087-499b8c3ffaf9/content [accessed 2026-08-14]
53. Troiano RP, Stamatakis E, Bull FC. How can global physical activity surveillance adapt to evolving physical activity guidelines? Needs, challenges and future directions. Br J Sports Med 2020 Dec;54(24):1468-1473. [doi: 10.1136/bjsports-2020-102621] [Medline: 33239352]
54. VandeBunte A, Gontrum E, Goldberger L, et al. Physical activity measurement in older adults: wearables versus self-report. Front Digit Health 2022;4:869790. [doi: 10.3389/fdgth.2022.869790] [Medline: 36120711]

Abbreviations

4AT: Assessment Test for Delirium
5-D: Decoding Death and Dying in people with Dementia by Digital thanotyping
ADL: activities of daily living
CDR: Clinical Dementia Rating
DIPH.DEM: digital phenotyping for changes in activity at the end of life in people with dementia
ENMO: Euclidean norm minus one

XSL•FO
RenderX

GAPPA: Global Action Plan on Physical Activity
GMHR: General Medical Health Rating
ICC: intraclass correlation coefficient
MAD: mean amplitude deviation
MIMS: Monitor-Independent Movement Summary
MVPA: moderate-to-vigorous physical activity
NPI-NH (K): Neuropsychiatric Inventory-Nursing home version, section K (nighttime behaviors)
PSG: polysomnography
PSMS: Physical Self Maintenance Score
REK: Regional Committee for Medical and Healthcare Research Ethics
ROCAM: Rate of Change Acceleration Movement
SE: sleep efficiency
SFI: sleep fragmentation index
SRI: sleep regulatory index
TST: total sleep time
WASO: wake after sleep onset
WHO: World Health Organization

Edited by E Borycki; submitted 12.Mar.2026; peer-reviewed by H Strauven, K Esquivel, X Tao; accepted 16.Jul.2026; published 28.Aug.2026.

Please cite as:
Boyle LD, Patrascu M, Husebo BS, Haugarvoll K, Steihaug OM, Marty B
A One-Year Study Using Digital Biomarkers From Sensing Technologies to Assess Changes in Physical Activity Levels and Sleep Quality in Nursing Home Residents With Dementia: Observational Study
JMIR Nursing 2026;9:e95194
URL: https://nursing.jmir.org/2026/1/e95194
doi:10.2196/95194

© Lydia D Boyle, Monica Patrascu, Bettina S Husebo, Kristoffer Haugarvoll, Ole Martin Steihaug, Brice Marty. Originally published in JMIR Nursing (https://nursing.jmir.org), 28.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Nurse-Led Digital Intervention for Sodium Restriction in Chronic Kidney Disease: Mixed Methods Implementation Study

Yu Shi[1], PhD; Hongmei Peng[1], MD; Youying Zhang[1], MD; Jingyu Zhang[2], PhD; Yang Li[1], BSc; Shi Pu[1], MD; Xiangchun Tang[1], BSc; Jinghong Zhao[1], PhD

[1]Department of Nephrology, Chongqing Key Laboratory of Prevention and Treatment of Kidney Disease, Chongqing Clinical Research Center of Kidney and Urology Diseases, The Second Affiliated Hospital (Xinqiao Hospital), Army Medical University (Third Military Medical University), 83 Xinqiao Zhengjie, Shapingba District, Chongqing, China

[2]State Key Laboratory of Ultrasound in Medicine and Engineering, Chongqing Medical University, Chongqing, China

Corresponding Author:
Jinghong Zhao, PhD
Department of Nephrology, Chongqing Key Laboratory of Prevention and Treatment of Kidney Disease, Chongqing Clinical Research Center of Kidney and Urology Diseases, The Second Affiliated Hospital (Xinqiao Hospital), Army Medical University (Third Military Medical University), 83 Xinqiao Zhengjie, Shapingba District, Chongqing, China

Abstract

Background: Digital interventions can support sodium restriction in chronic kidney disease; yet, implementation and engagement barriers remain poorly understood.

Objective: This mixed methods study evaluated a nurse-led digital program for sodium restriction, focusing on 3 key aspects: implementation, engagement, and patient experience.

Methods: This study was initially designed as a randomized controlled trial but was revised to a mixed methods implementation study due to high rates of missing primary outcome data. Quantitative measures included 24-hour urinary sodium excretion (the primary outcome of the original randomized controlled trial), knowledge tests, blood pressure measurements, and quality-of-life assessments. Concurrently, we conducted in-depth qualitative interviews with a purposive sample of participants. Quantitative and qualitative data were collected in parallel, analyzed separately, and then integrated to provide a comprehensive understanding of participants' experiences with the intervention.

Results: Among 99 enrolled patients (46 in the enhanced digital intervention group and 53 in the minimal digital control group), the program led to a significant improvement in sodium-related knowledge as a secondary exploratory outcome (P=.005). However, the 24-hour urine collection completion rate was only 24.2% (n=24), precluding reliable conclusions regarding the primary outcome. Qualitative findings revealed a "measurement paradox," in which the gold-standard assessment itself posed a major barrier to adherence. Integrated data illustrated a patient journey marked by early enthusiasm, followed by declining engagement due to digital fatigue, family resistance, and poor fit with daily life. This misalignment created a disconnect between acquired knowledge and measurable behavior change.

Conclusions: A nurse-led digital intervention was associated with modest improvements in sodium knowledge in chronic kidney disease. However, success requires reducing the monitoring burden and improving real-world fit. The "measurement paradox" highlights the need for patient-centered outcome measures. We propose a "stage-matched intervention" framework to guide adaptive digital tools. Future research should prioritize multicenter pragmatic trials to bridge the knowledge-action gap.

Trial Registration: Chinese Clinical Trials Registry ChiCTR2200060627; https://tinyurl.com/43sx53j4

(JMIR Nursing 2026;9:e94330) doi:10.2196/94330

KEYWORDS

chronic kidney disease; sodium restriction; digital health; mixed-methods research; nurse-led digital intervention

Introduction

Chronic kidney disease (CKD) affects millions worldwide, and high sodium intake accelerates disease progression [1-3]. Clinical guidelines recommend a daily sodium intake below 2.3 grams or 100 mmol to control blood pressure and preserve kidney function [4-6], but real-world adherence remains poor. Fewer than 30% of patients with CKD meet this target [7-9], revealing a critical knowledge-practice gap in disease management. As frontline providers, nurses are well positioned to leverage digital tools to improve sodium management [10-12]. To systematically address these self-care deficits, this study was guided by Orem's Self-Care Deficit Nursing Theory, which posits that individuals require nursing support when their

XSL•FO
RenderX

self-care capabilities are insufficient [13]. For patients with CKD, inadequate knowledge, poor adherence, and low self-monitoring skills represent key self-care deficits in sodium management. The nurse-led digital intervention was therefore designed to identify these deficits, provide compensatory support, and enhance self-care agency, fully aligning with the core logic of Orem's theory. Accordingly, we developed a theory-driven, nurse-led mobile digital sodium restriction program embedded within our independently designed Su Yi app (Thinmed Medical Technology (Chongqing) Co, Ltd), integrating nursing expertise into patient education to address these self-care deficits [14,15].

While digital therapeutics have demonstrated benefits for chronic conditions such as hypertension [16-19], their application in CKD management remains nascent. Most existing studies rely on idealized randomized controlled trial (RCT) designs (Checklist 1) [7,20] that prioritize internal validity but overlook real-world implementation barriers, including digital fatigue, social resistance, and procedural burden. A prime example is the 24-hour urinary sodium test, the gold standard for estimating sodium intake, which imposes a considerable burden on patients due to its cumbersome nature [21]. Yet, its role as a primary end point is rarely reconsidered, creating a critical measurement paradox. Low adherence to such measures renders study outcomes inconclusive, hindering the identification of why interventions fail. Although some studies indicate that digital interventions can improve dietary knowledge among patients with CKD [22,23], the reasons why knowledge gains often do not translate into improved behavior or physiological outcomes remain unclear. Critically, large-scale failures in objective data collection often overshadow patients' actual experiences with digital tools, obscuring systemic mismatches between intervention design and the realities of daily life.

Existing literature identifies a key paradox in digital health interventions. Patients often show significant improvements in health-related knowledge; yet, objective behavioral metrics, such as physiological assessment completion rates, remain consistently low. In sodium management programs specifically, knowledge acquisition does not reliably translate into adherence to urine collection protocols used to verify dietary changes. This discrepancy raises a critical question: why do theoretically effective interventions struggle to sustain in real-world settings? Possible explanations include intervention design flaws, insufficient patient motivation, or a fundamental misalignment between digital tools and patients' daily experiences. The field currently lacks an in-depth understanding of the full engagement process among patients with CKD in digital sodium management. Without insights into patients' emotional shifts, practical challenges, and evolving needs during intervention, there is a risk of repeatedly developing technically advanced but clinically impractical solutions. To address this gap, advanced practice nurses and clinical nurse researchers conducted a parallel mixed methods study integrating quantitative and qualitative approaches. Led by the nursing team, the intervention was designed and implemented to systematically evaluate both the effectiveness and patient experience of a nurse-developed digital sodium restriction program for patients with CKD.

Initially registered as an RCT (ChiCTR2200060627) with 24-hour urinary sodium as the primary outcome, this study was reconceptualized as an exploratory mixed methods implementation study due to 75.7% (75/99) missing data on the primary outcome. The objectives shifted from efficacy verification to exploring implementation barriers and patient engagement experiences. Guided by Orem's Self-Care Deficit Nursing Theory, we hypothesized 2 core points. First, the nurse-led digital intervention may improve sodium-related knowledge but is likely to face high attrition in burdensome 24-hour urinary sodium collection due to mismatched self-care capabilities. Second, mapping longitudinal engagement trajectories is expected to identify the core implementation barriers, informing human-centered digital tool development and clarifying the knowledge-behavior disconnect. This study addresses a critical gap in digital nursing research by moving beyond traditional efficacy-focused RCTs to prioritize real-world implementation and user experience. Our findings offer practical guidance for designing adaptive, feasible digital interventions for CKD management.

Methods

Study Design and Setting

This study was prospectively registered as a parallel-group RCT in the Chinese Clinical Trial Registry (ChiCTR2200060627; registered on June 6, 2022). The initial design aimed to evaluate intervention efficacy, with 24-hour urinary sodium excretion as the primary outcome and secondary outcomes including sodium-related knowledge, blood pressure, and quality of life (QOL). A 3-month follow-up duration was selected based on previous digital health implementation studies in CKD and hypertension, which have demonstrated that 8 to 12 weeks is sufficient to detect changes in intermediate physiological and behavioral outcomes in real-world settings [7,17]. This time frame balances the need to observe meaningful behavior change while minimizing attrition in longitudinal implementation research.

During the 3-month interim review, only 24.2% (24/99) of participants completed valid 24-hour urine collections, resulting in 75.7% (75/99) missing data for the primary outcome, far below the prespecified acceptable threshold of 60%. Because increasing the sample size could not improve data completeness, the study design was formally revised from an RCT to a single-center, exploratory mixed methods implementation study [24-26].

This design adaptation was consistent with reporting guidelines for pragmatic implementation research. The original sample of 99 participants remained adequate for mixed methods inquiry, which typically requires at least 80 quantitative participants and 15 to 30 qualitative participants [27]. Accordingly, the analytical focus shifted from confirmatory efficacy testing to exploratory implementation outcomes. The core objective was revised to evaluate implementation processes, participant engagement experiences, and barriers to a nurse-led digital sodium restriction intervention in patients with CKD.

XSL•FO
RenderX

During the 3-month interim progress review conducted in November 2022, the completion rate for the prespecified primary outcome measurement was merely 24.2% (24/99), far below the expected level. Concurrently, notable participant attrition was observed in both groups. Given the poor feasibility of continuing the original primary outcome evaluation, our research team comprehensively assessed the actual research progress and field conditions. After obtaining formal approval from the ethics committee, we decided to abandon the planned primary outcome efficacy analysis, redefine this project as an exploratory mixed methods implementation study, and shift our core research focus to exploring intervention implementation processes and real-life patient experiences.

Ethical Considerations

The research design and study protocol were approved by the Ethics Committee of the Second Affiliated Hospital of Army Medical University (approval number 2022-Research-040-01). The revised study design was subsequently reviewed and approved by the same ethics committee (approval number 2022-Research-515-01). All participants provided written informed consent prior to enrollment. All study procedures were conducted in accordance with the Declaration of Helsinki.

Participant Recruitment and Screening

Eligible participants were recruited and screened by trained clinical nurses from inpatients admitted to the nephrology department between June 2022 and August 2022. Inclusion criteria were as follows: age 18 to 65 years; diagnosis of CKD stages 1‐5 [28], including those with hypertension; an estimated glomerular filtration rate stable over the past 3 months; 24-hour urinary sodium excretion >130 mmol/day at baseline; Barthel Index score ≥60; Mini-Mental State Examination (MMSE) score ≥24; proficiency in using smartphones and mobile applications; and willingness to participate. Exclusion criteria included the following: recent acute kidney injury; receipt of renal replacement therapy; major cardiovascular event within 6 months; uncontrolled hypertension (systolic blood pressure >160 mm Hg or diastolic blood pressure >100 mm Hg); prior participation in similar digital health programs; or any condition likely to interfere with study compliance.

Randomization and Blinding

Eligible participants were randomly assigned (1:1) to either the intervention or control group using computer-generated random sequences in SPSS (version 24.0; IBM Corp). Allocation was concealed via sequentially numbered, opaque, and sealed envelopes prepared by an independent statistician. Due to the nature of the intervention, participants and the intervention delivery nurse could not be blinded. However, to minimize ascertainment bias, outcome assessors, laboratory technicians, and data analysts remained blinded to group allocation throughout the study. The 46:53 group size discrepancy resulted from random fluctuation of the computer-generated sequence with no manual modification. Such a degree of imbalance falls within the expected range of chance variation in simple randomization and does not compromise the validity of random allocation.

Interventions

Control Group (Minimal Digital Intervention)

Participants received minimal digital intervention [29], which included a single educational session on sodium restriction and access to the Su Yi app—a self-developed mobile platform providing static educational content (1 video and 1 text document) [30]. They were prompted daily via the app to measure and log their morning blood pressure. Both groups used the Su Yi app; however, the intervention content and intensity differed between the groups.

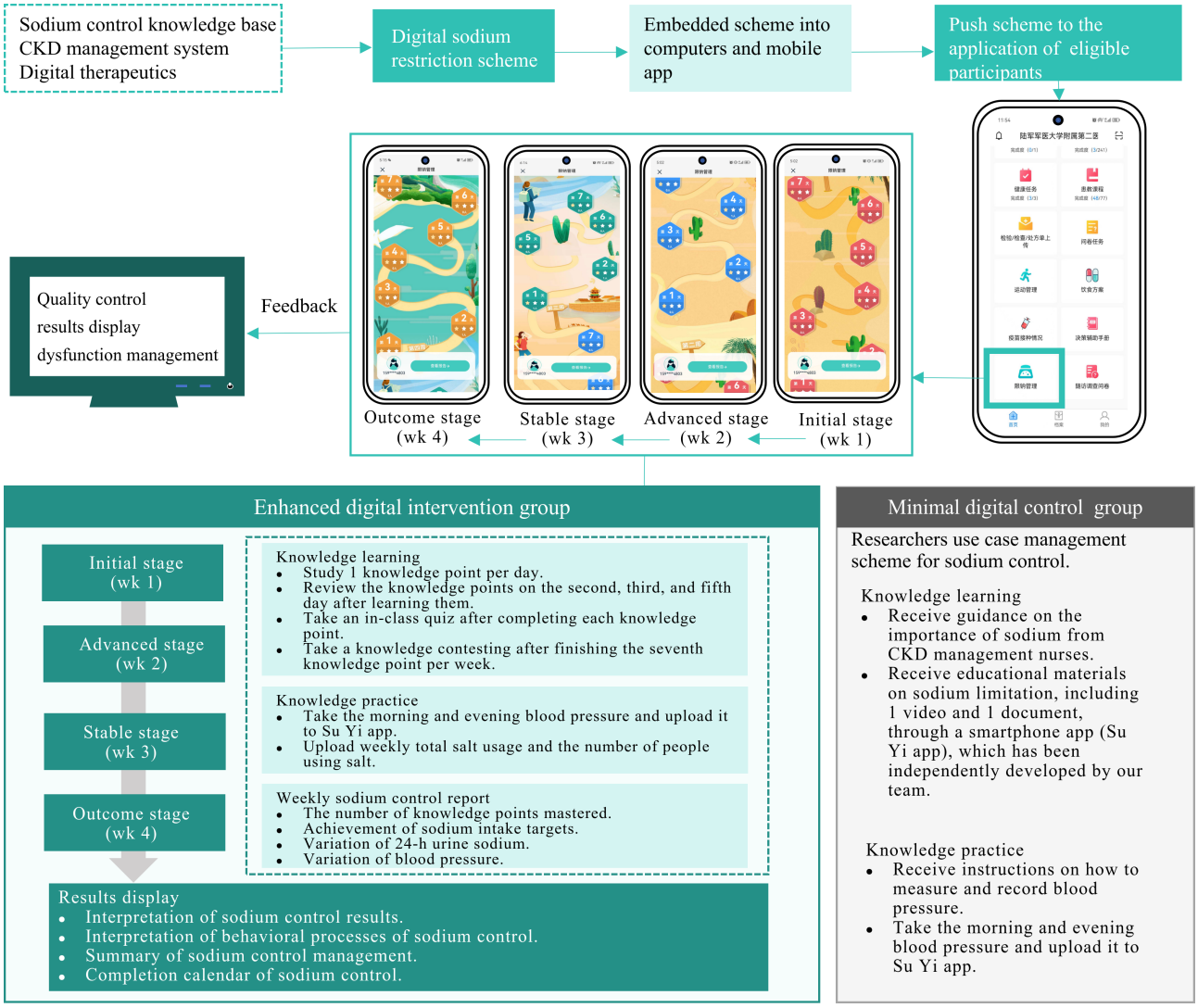
Intervention Group (Enhanced Digital Intervention)

Participants received an enhanced digital intervention that was strictly grounded in Orem's Self-Care Deficit Nursing Theory. The intervention followed three core theoretical components: (1) deficit identification through baseline knowledge assessment, (2) compensatory nursing support via structured education and gamified learning, and (3) self-care agency enhancement using personalized feedback and autonomous practice. The 4-week intervention was delivered in four progressive stages (Figure 1): (1) initiation (knowledge learning), (2) advancement (quizzes and peer competition), (3) stabilization (behavioral practice), and (4) outcome (personalized reporting). This staged structure directly operationalized Orem's theoretical logic of identifying deficits, providing compensatory support, and promoting self-care capacity. Gamified modules, in-app quizzes, peer competition, and standardized behavioral tasks were integrated to enhance engagement. Upon completion of the 4-week intervention, each participant received a personalized report illustrating knowledge mastery, salt intake goal achievement, trends in 24-hour urinary sodium excretion, and blood pressure trajectories. All activities were completed autonomously via the app, minimizing clinical burden while maintaining consistent nursing oversight through remote monitoring.

Both groups attended a follow-up visit 3 months post intervention for outcome assessment.

XSL•FO
RenderX

Figure 1. Formation and application process of the digital sodium control system. CKD: chronic kidney disease.



Sample Size Calculation

The sample size was calculated using PASS (version 15.0; IBM Corp) software based on the anticipated between-group difference in 24-hour urinary sodium excretion at the 3-month follow-up. Data from a previous randomized trial suggested a mean excretion of approximately 173.5 mmol/day in control groups and 147.5 mmol/day in intervention groups, indicating a potential difference of 26 mmol/day [31]. A common SD of 34.0 mmol/day was assumed based on the same study. Using a 2-sample, 2-tailed *t* test with α=.05% and 90% power, 37 participants per group were required. Accounting for an estimated 20% dropout rate, the final target sample size was set at 47 per group (total n=94).

Data Collection Procedures

Quantitative data were collected at baseline (T0) and at the 3-month follow-up (T1). Outcomes of interest included 24-hour urinary sodium excretion [32] and blood pressure [33] as potential indicators of physiological response. Blood pressure was defined as the average of 3 consecutive office blood pressure measurements. Additional outcomes comprised sodium control knowledge, QOL, and laboratory parameters. Sodium control knowledge was assessed using a 10-item questionnaire developed by the research team. The questionnaire was based on CKD sodium restriction guidelines, refined through 3 rounds of expert review, with a content validity index of 0.92, and validated via pilot testing with 20 participants and a Cronbach α of 0.78, and the questionnaire is provided in Multimedia Appendix 1. QOL was measured using KDQOL-SF (Kidney Disease Quality of Life-Short Form) 1.3 [34], and laboratory parameters included serum creatinine, urea, uric acid, electrolytes, hemoglobin, and albumin. Adherence to 24-hour urine collection was anticipated to present challenges; therefore, this measure was analyzed exploratorily and interpreted cautiously, with due consideration of completion rates.

Qualitative data were collected through semistructured in-depth interviews with a purposive subsample of 23 participants. Participants who discontinued the intervention were not included due to practical constraints related to time and research resources. This subsample comprised only intervention completers; accordingly, all qualitative findings were limited to participants who completed the full intervention program.

All interviews were conducted in Mandarin Chinese by 2 nurse researchers with master's degrees who had received systematic training in qualitative research (Checklist 2). These researchers were not involved in intervention delivery to minimize potential

XSL•FO
RenderX

bias. Face-to-face interviews were conducted in the outpatient CKD management clinic, with each session lasting 20 to 40 minutes. All interviews were audio-recorded, transcribed verbatim, and anonymized. The transcripts were then translated into English using dual translation and back-translation to ensure linguistic and conceptual fidelity. Regular debriefing sessions were held throughout data collection to enhance reflexivity and analytic rigor. The interview guide was developed under the core framework of Orem's Self-Care Deficit Nursing Theory, with the integration of key concepts from Empowerment Theory. The guide was designed to explore participants' experiences across multiple dimensions, including task execution, emotional responses, barriers encountered, and evolving needs. Thematic analysis was used to analyze the interview transcripts. To ensure coding reliability, 2 researchers independently coded 20% of the data, yielding strong intercoder agreement with a Cohen κ value of 0.72. Discrepancies were resolved through consensus discussion. The finalized codebook, which included code definitions, thematic hierarchies, representative quotes, and decision rules, was then applied to the full dataset by the primary researcher. The codebook is provided in Multimedia Appendix 2. Sensitivity analyses were conducted to assess the potential influence of selection bias on qualitative interpretations. We hypothesized that participants who discontinued the intervention may hold distinct perspectives related to higher intervention burden, digital fatigue, and competing time demands. Such systematic differences may lead to overestimation of participant engagement and underestimation of real-world implementation barriers if noncompleters are excluded from analysis.

Statistical and Analytical Approaches

Quantitative Analysis

During manuscript revision, we consulted a biostatistician regarding the handling of the substantial missing data in 24-hour urinary sodium excretion. It was determined that traditional inferential statistics (eg, *t* tests, ANOVA) would yield unreliable estimates due to severe attrition (75/99, 75.8% missing) and potential selection bias. Consequently, we have revised the manuscript to reflect a mixed methods evaluation framework, removed causal interpretations of urinary sodium outcomes, and focused the discussion on implementation insights and patient-reported experiences.

Continuous variables were tested for normality using the Shapiro-Wilk test. Normally distributed data were presented as mean (SD) and compared using independent samples *t* tests. Nonnormally distributed variables were reported as median (IQR) and analyzed using Mann-Whitney *U* tests. Categorical variables were expressed as frequencies (%) and compared using chi-square or Fisher exact tests. Intention-to-treat (ITT) analysis was applied to all primary outcomes. Statistical significance was set at $P<.05$. Analyses were performed using SAS software (version 9.4; SAS Institute Inc). Given the substantial missingness in 24-hour urinary sodium excretion (>60%), we carefully evaluated the data quality and implemented the following handling strategies: (1) ITT analysis was applied to all available data to maintain the randomized design's integrity; (2) between-group inferential comparisons were not performed, even for completers, due to the high risk of selection bias; and (3) multiple imputation was not adopted as it would require untestable assumptions about the missing values, for example, missing at random, given the extreme missing rate. We interpret this outcome exploratorily, focusing on implementation insights rather than efficacy conclusions.

Qualitative Analysis

Transcripts were analyzed using thematic analysis guided by Empowerment Theory. A hybrid deductive-inductive approach was used: initial coding was informed by predefined domains (tasks, emotions, pain points, and needs), while emergent themes were iteratively refined. Two researchers independently coded 20% of the transcripts to assess reliability, achieving consensus through discussion. NVivo 12 (QSR International) was used for data management.

Integration of Mixed Methods

A convergent parallel mixed methods design was used in accordance with the GRAMMS (Good Reporting of A Mixed Methods Study) guidelines (Checklist 3) [35]. Quantitative and qualitative data were collected in parallel, analyzed independently, and subsequently integrated using a joint display matrix. Quantitative outcomes provided the primary analytical framework, while qualitative findings were used to explain, supplement, and contextualize the quantitative results. To ensure integration rigor, 2 researchers conducted independent interpretive integration; discrepancies were resolved through consensus discussions, and a third researcher reviewed the entire integration process.

Results

Participant Flow

Of 303 screened patients, 99 were randomized: 46 to the enhanced digital intervention group and 53 to the minimal digital control group. Baseline characteristics were similar between groups (Table 1), confirming balanced randomization. At the 3-month follow-up, 52 (98.1%) in the control group and 45 (97.8%) in the intervention group completed blood pressure and blood component analysis, with the exception of 1 death in the intervention group and 1 withdrawal in the control group for personal reasons. However, only 13 (24.5%) in the control group and 11 (23.9%) in the intervention group completed the 24-hour urine collection (Figure 2), resulting in an overall completion rate of 24.2% (24/99) for this outcome measure.

XSL•FO
RenderX

Table . Baseline demographic and clinical characteristics of participants (N=99).

Variables	Control (minimal digital; n=53)	Intervention (enhanced digital; n=46)	*z*	Chi-square (*df*)	*P* value
Age (y), median (IQR)	48 (37 - 54)	44 (31-52)	−1.24	—[a]	.21
Sex, n (%)			—	1.56 (1)	.21
Male	32 (60.3)	22 (47.8)			
Female	21 (39.6)	24 (52.1)			
Educational level			—	1.47 (4)	.83
Primary education or below	11 (20.7)	7 (15.2)			
Junior middle school	19 (35.8)	15 (32.6)			
Senior high school	7 (13.2)	9 (19.5)			
Associate degree	8 (15.1)	6 (13.0)			
Bachelor's degree or above	8 (15.1)	9 (19.5)			
Hypertension			—	1.51 (1)	.22
No	20 (37.7)	23 (50)			
Yes	33 (62.2)	23 (50)			
CKD[b] stage			—	7.20 (4)	.13
1	13 (24.5)	20 (43.4)			
2	10 (18.8)	12 (26.1)			
3	13 (24.5)	6 (13.0)			
4	8 (15.1)	3 (6.5)			
5	9 (16.9)	5 (10.8)			
BMI (kg/m²), median (IQR)	24.6 (21.5-26.9)	24.0 (22.1-26.8)	−0.15	—	.88
Waist circumference (cm), median (IQR)	85 (78-94)	81 (77-92)	−1.55	—	.12
Urine protein to creatinine ratio (g/g), median (IQR)	1.8 (0.6-2.4)	1.2 (0.3-3.9)	−0.52	—	.60
Serum creatinine (μmol/L), median (IQR)	146.7 (85.3-315.5)	94.3 (94.3-56.6)	−1.90	—	.06
Urea (mmol/L), median (IQR)	9.2 (5.8-12.8)	6.1 (4.6-11.7)	−1.44	—	.15
Uric acid (μmol/L), median (IQR)	392.3 (329.8-482.6)	400.3 (308.0-481.0)	−0.69	—	.49
Glucose (mmol/L), median (IQR)	5.2 (4.7-5.6)	5.2 (4.8-5.9)	0.29	—	.78
Serum sodium (mmol/L), median (IQR)	139.0 (137.6-140.2)	138.6 (137.6-140.0)	−0.85	—	.40
Serum potassium (mmol/L), median (IQR)	4.1 (4.0-4.5)	4.1 (3.9-4.4)	−0.46	—	.64
Serum phosphate (mmol/L), median (IQR)	1.2 (1.0-1.3)	1.2 (1.0-1.3)	−0.29	—	.77

XSL•FO
RenderX

Variables	Control (minimal digital; n=53)	Intervention (enhanced digital; n=46)	*z*	Chi-square (*df*)	*P* value
Serum calcium (mmol/L), median (IQR)	2.2 (2.1-2.3)	2.2 (2.1-2.3)	−1.21	—	.22
Total protein (g/L), median (IQR)	65.1 (59.1-72.4)	65.9 (58.1-73.6)	−0.14	—	.89
Hemoglobin (g/L), median (IQR)	126.0 (113.0-142.0)	129.0 (106.0-138.0)	−0.40	—	.80
Triglycerides (mmol/L), median (IQR)	1.8 (1.2-2.8)	1.6 (1.3-2.2)	−0.40	—	.69
Total cholesterol (mmol/L), median (IQR)	4.9 (4.4-6.0)	5.0 (4.1-6.0)	<0.01	—	>.99
Systolic blood pressure (mm Hg), median (IQR)	130 (116-148)	127 (115-145)	−0.07	—	.95
Diastolic blood pressure (mm Hg), median (IQR)	82 (77-93)	82 (75-90)	−0.24	—	.81
Heart rate (beats/min), median (IQR)	82 (80-87)	81 (78-85)	−0.84	—	.40
Urinary sodium (mmol/24 h), median (IQR)	156.3 (141.4-184.6)	166.3 (140.8-206.7)	−0.84	—	.42
Sodium control knowledge score	3 (2-5)	3 (2-5)	1.13	—	.26
SF-36[c], median (IQR)	2295 (2025-2500)	2247 (2247-1935)	−0.34	—	.73
KDTA[d], median (IQR)	3126 (2673-3440)	3165 (2758-3483)	0.29	—	.77

[a]—: not applicable.

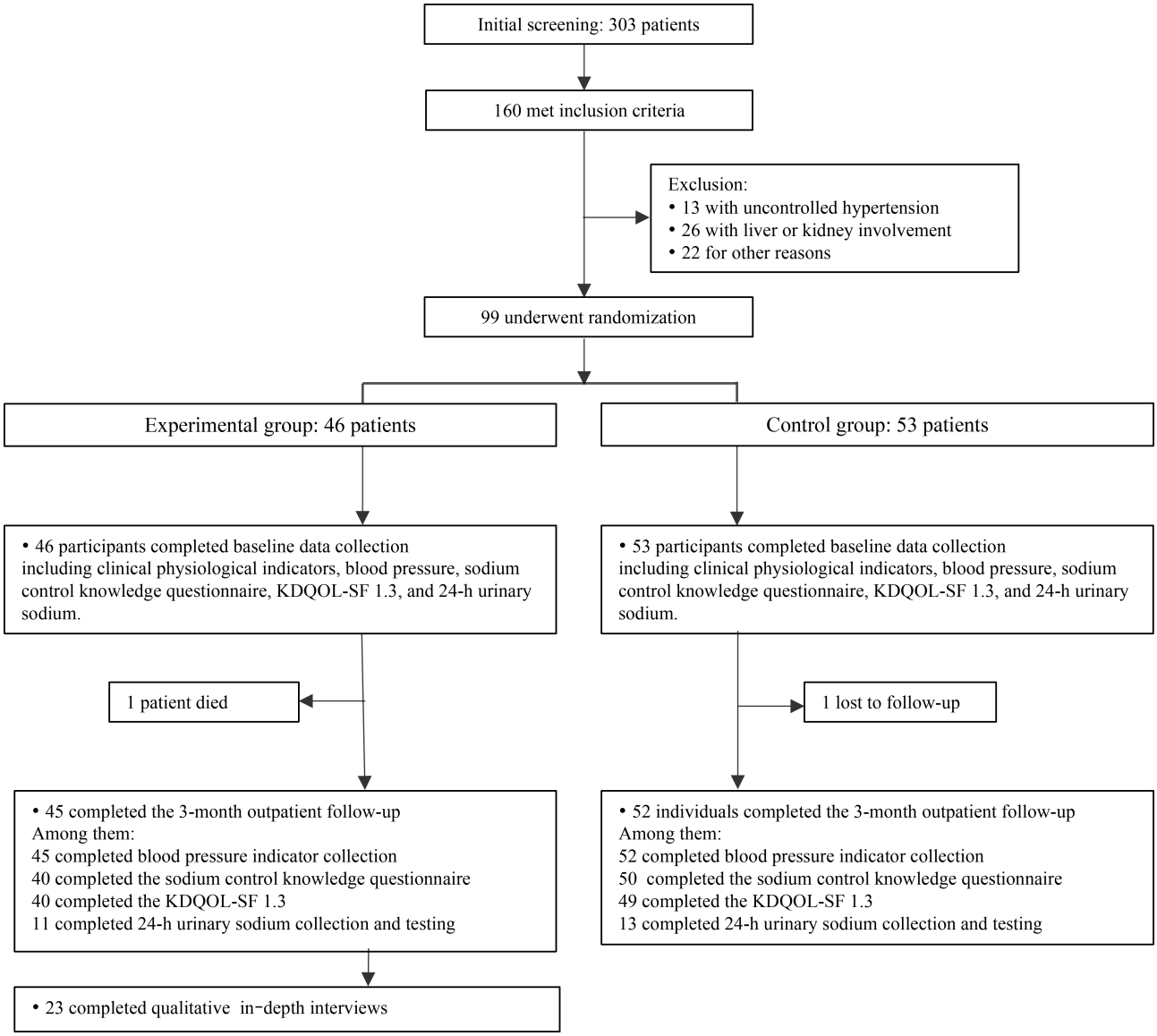
[b]CKD: chronic kidney disease.

[c]SF-36: 36-Item Short Form Health Survey.

[d]KDTA: kidney disease targeted areas.

XSL•FO
RenderX

Figure 2. Flow diagram of participant enrollment. KDQOL-SF: Kidney Disease Quality of Life-Short Form.



Baseline Characteristics

Baseline demographic and clinical characteristics were balanced between groups (Table 1). The median age was 48 (IQR 37 - 54) years in the control group and 44 (IQR 31-52) years in the intervention group (*P*=.21). There were no significant differences in sex distribution, CKD stage, comorbidities, or laboratory parameters, confirming successful randomization.

Quantitative Outcomes

As shown in Table 2, the completion rate for 24-hour urinary sodium excretion was critically low (24/99, 24.2%), precluding reliable estimation of between-group differences or definitive conclusions about behavioral change. Among the small subset of completers (n=24), median sodium excretion appeared lower in the intervention group (97.3, IQR 73.7 - 124.0) compared to controls (154.4 mmol/day, IQR 123.2 - 190.9). However, inferential testing was not conducted due to insufficient power and selection bias concerns.

However, a significant improvement was observed in the secondary outcome of sodium control knowledge: intervention group participants scored 8 (IQR 7-8) versus 6 (IQR 6-8) in controls (z=2.84, *P*=.005). No significant differences were found in blood pressure, QOL (36-Item Short Form Health Survey [SF-36] and kidney disease targeted areas), or other physiological markers.

XSL•FO
RenderX

Table . Analysis of outcomes at 3-month follow-up by study group (N=97)[a].

Outcomes	Control (minimal digital; n=52)		Intervention (enhanced digital; n=45)		*z*	*P* value
	n[b]	Median (IQR)	n[b]	Median (IQR)		
Sodium control knowledge score	50	6 (6-8)	40	8 (7-8)	2.84	.005
SF-36[c]	49	2515 (2355-2640)	40	2535 (2315-2780)	1.06	.30
KDTA[d]	49	3375 (3048-3755)	40	3535 (3082-3939)	0.78	.44
Systolic blood pressure (mm Hg)	52	126 (112-141)	45	125 (112-139)	−0.22	.83
Diastolic blood pressure (mm Hg)	52	78 (72-88)	45	80 (73-88)	0.45	.66
Heart rate (beats/min)	52	80 (78-85)	45	80 (78-86)	0.38	.70

[a]Completion rate for 24-hour urinary sodium excretion was 24.2% (24/99). Between-group comparisons were not performed due to low adherence and risk of selection bias.

[b]Number of participants with available data for each outcome.

[c]SF-36: 36-Item Short Form Health Survey.

[d]KDTA: kidney disease targeted areas.

Qualitative Findings

Overview of the Patient Journey Map

Thematic analysis was conducted among 23 intervention completers; noncompleters were not interviewed due to resource constraints. Sensitivity analyses indicated that the exclusion of noncompleters may reduce the representativeness of the patient journey map. Consequently, all findings are restricted to participants who completed the enhanced digital intervention. Data saturation was assessed retrospectively after coding all 23 interview transcripts and was reached at the 18th interview, with no new themes identified in the subsequent 5 interviews. The analysis revealed a dynamic patient journey, summarized in Table 3. A patient journey map typically consists of a horizontal axis (timeline) and a vertical axis (task axis). Guided by Empowerment Theory, this study divides the typical journey of patients with CKD in a digital sodium restriction intervention along the horizontal axis into three phases: (1) initial contact and expectation phase (week 1); (2) sustained interaction and empowerment phase (weeks 2 - 4); and (3) platform-facilitated relationship evolution phase (week 5 and beyond). The vertical axis encompasses 4 dimensions—tasks, emotions, pain points, and needs—yielding a total of 28 themes. These themes capture the realities, emotional shifts, barriers, and needs faced by patients at each stage of the digital sodium restriction intervention, comprehensively mapping the storyline across different dimensions throughout each phase. This reveals the dynamic evolution of patients from adherence to self-empowerment during the digital sodium restriction intervention, ultimately forming a typical patient journey map for patients with CKD in such interventions, as shown in Figure 3.

XSL•FO
RenderX

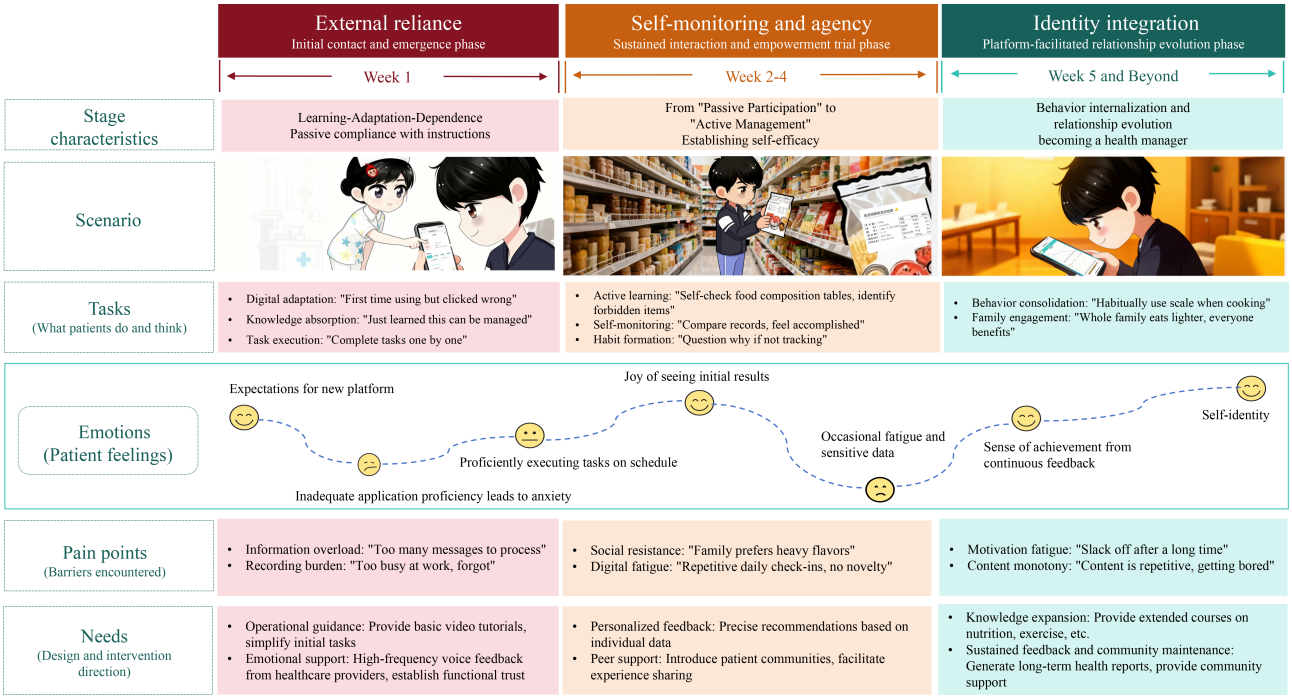
Table . Participant characteristics (qualitative study).

Number	Sex	Age (y)	Educational level	Profession	CKD[a] stage
P1	Female	43	Primary education or below	Unemployed	5
P2	Male	52	Junior middle school	Farmer	2
P3	Female	29	Bachelor’s degree or above	Office worker	1
P4	Female	24	Associate degree	Professionals	1
P5	Female	37	Bachelor’s degree or above	Civil servant	1
P6	Male	61	Bachelor’s degree or above	Professionals	5
P7	Male	38	Junior middle school	Migrant worker	1
P8	Female	55	Senior high school	Retired	1
P9	Female	31	Bachelor’s degree or above	Civil servant	1
P10	Female	56	Primary education or below	Office worker	2
P11	Female	22	Junior middle school	Office worker	1
P12	Male	45	Associate degree	Civil servant	2
P13	Female	45	Junior middle school	Unemployed	3
P14	Female	28	Bachelor’s degree or above	Unemployed	1
P15	Male	47	Senior high school	Migrant worker	2
P16	Female	24	Junior middle school	Office worker	1
P17	Female	47	Junior middle school	Unemployed	1
P18	Male	43	Junior middle school	Migrant worker	3
P19	Male	27	Associate degree	Civil servant	4
P20	Male	50	Senior high school	Migrant worker	4
P21	Male	26	Associate degree	Professionals	1
P22	Female	37	Bachelor’s degree or above	Teacher	2
P23	Male	20	Associate degree	Student	1

[a]CKD: chronic kidney disease.

XSL•FO
RenderX

Figure 3. Patient empowerment journey among enhanced digital intervention completers.



Initial Contact and Expectation Phase (Learning-Adapting-Depending)

Participants described initial excitement and hope (“I wanted to try anything that could help my kidneys” [P5/13]), which was quickly tempered by information overload and technical anxiety (“So many messages… I didn’t know what to do first” [P7]). Key barriers included difficulty navigating the app interface and uncertainty about correct data entry. Emotional dependency on nurse feedback emerged as a critical motivator for early adherence.

Sustained Interaction and Empowerment Phase (From Passive to Active Management)

As familiarity grew, participants transitioned from passive compliance to active self-monitoring and habit formation. Many reported checking food labels spontaneously (“Now I always look at the nutrition facts” [P5/7/21]) and using salt substitutes. Personalized feedback and peer comparison features enhanced motivation. However, digital fatigue (“It felt repetitive after a while” [P18]) and social resistance (“My family said the food was too bland” [P1/9]) emerged as significant challenges.

Platform-Facilitated Relationship Evolution Phase (Identity Integration)

In later stages, participants began identifying as health managers rather than passive patients. They shared knowledge with family members (“My kids now remind me not to eat pickles” [P13/16/20]) and advocated for healthier meals. Despite these positive shifts, motivational decay and content monotony threatened long-term engagement. Participants expressed strong desires for expanded educational content (eg, protein and potassium management) and ongoing community support.

Mixed Methods Integration

Integrated analyses were performed using a joint display matrix (Table 4) following GRAMMS guidelines for mixed methods reporting. Three patterns of integration were identified. First, convergent findings demonstrated significant improvements in sodium-related knowledge in the enhanced digital intervention group (z=2.84, P=.005). These quantitative improvements were supported by qualitative accounts of clear, structured learning and enhanced confidence in sodium management, providing consistent support for the improvement in knowledge gains observed in the quantitative analysis. Second, supplementary findings highlighted a critical measurement paradox: completion of 24-hour urinary sodium collection was only 24.2% (24/99). Qualitative data further identified substantial procedural burden and poor compatibility with daily routines as key barriers, indicating that the gold-standard assessment itself impeded adherence. Third, explanatory findings clarified the absence of significant between-group differences in blood pressure or QOL (P>.05). Identified barriers, including family resistance, digital fatigue, and limited long-term support, explained why improvements in knowledge did not translate into measurable changes in clinical or patient-reported outcomes. No divergent findings were observed. The integrated synthesis indicates that although the intervention improved sodium-related knowledge, translation into behavioral and clinical changes was limited by implementation barriers and measurement burden.

XSL•FO
RenderX

Table . Joint display matrix for mixed methods integration.

Quantitative outcome	Qualitative finding	Integration type	Integrated interpretation
Sodium control knowledge: intervention 8 (7-8), control 6 (6-8); z=2.84, P=.005	Participants reported clear learning, easy-to-understand modules, and increased confidence in sodium management.	Convergence	The intervention effectively improved sodium-related knowledge, consistent with participant experience.
24 h urine completion: 24.2% (24/99); 75.7% (75/99) missing	Participants reported high burden, time intensity, and poor fit with daily life; urine collection was the top barrier.	Supplementation	Low completion is explained by substantial patient burden, revealing a "measurement paradox" in gold-standard outcome assessment.
BP[a] or QoL[b]: no significant between-group differences (P>.05)	Barriers included family resistance, digital fatigue, repetitive tasks, and insufficient long-term support.	Explanation	Neutral clinical outcomes are explained by real-world implementation barriers that prevented knowledge from translating to behavior change.

[a]BP: blood pressure.

[b]QOL: quality of life.

Discussion

Principal Findings

This study found that the nurse-led digital intervention improved participants' sodium-related knowledge as a secondary exploratory outcome. However, the completion rate of the primary outcome (24-h urinary sodium collection) was only 24.2% (24/99), which precluded evaluation of the intervention's true efficacy. These findings are consistent with Orem Self-Care Deficit Nursing Theory. Qualitative interviews identified a 3-phase patient engagement journey: initial enthusiasm, active participation, and subsequent disengagement driven by digital fatigue and family barriers. Overall, while digital interventions can enhance health knowledge, high-burden assessment tools and poor adaptation to daily life significantly hinder sustained behavioral changes.

From Information Delivery to Multidimensional Empowerment, Bridging the Knowledge-Action Gap

The improvement in knowledge without obvious behavioral changes supports Orem theory, which holds that self-care agency requires adequate knowledge, skills, motivation, and environmental support [36-39]. This reflects a typical knowledge-action gap. Digital tools can efficiently disseminate standardized health information on a large scale, but knowledge alone cannot drive long-term behavioral changes [40-43]. Sodium restriction is affected by family dynamics, cooking habits, and taste preferences—factors far more complex than simple information delivery [44,45]. Therefore, future interventions should shift from one-way knowledge popularization to multidimensional empowerment, comprehensively addressing both structural and interpersonal barriers [46-48].

Selecting Scientifically Sound and Feasible Outcome Measures

This study revealed a measurement paradox. Although 24-hour urine collection is the gold standard for sodium intake assessment [49], its complicated operation and heavy burden lead to low patient adherence and high measurement error [50,51]. This phenomenon aligns with Orem's view that self-care interventions must match patients' actual capabilities, which limits its application in clinical practice and research. By contrast, the spot urine sodium-to-creatinine ratio is simple to operate and has been validated to predict disease progression in multiple studies [52,53]. AI-based real-time urine testing may balance accuracy and convenience; however, its reliability across different patient groups requires further verification before large-scale application [54-56]. We therefore suggest researchers prioritize outcome measures that offer both scientific rigor and real-world feasibility, rather than pursuing methodological purity at the expense of practicality, to improve the clinical translation of research findings.

Constructing a Dynamic, Stage-Matched Intervention Framework

Family resistance, lifestyle constraints, and digital fatigue are key barriers to long-term intervention effects [57-59]. Digital health success relies on continuous patient engagement rather than short-term adoption [60-62]. Based on Orem's self-care theory, we propose a stage-matched intervention framework that includes 3 core components: deficit identification, compensatory support, and self-care capacity improvement [63]. In the initial stage, medical staff identify patients' knowledge gaps and practical obstacles to build trust. In the middle stage, personalized digital guidance is used to improve patients' self-efficacy and bridge capability gaps. The later stage helps patients maintain healthy behaviors through community interaction and independent self-care management. This framework transforms static, one-size-fits-all intervention modes into adaptive, dynamic support that addresses the multidimensional self-care deficits in CKD sodium management. Future mixed methods studies should further explore the mechanisms of patient engagement to inform the development of truly patient-centered digital intervention solutions.

Strengths and Limitations

This study adopted a parallel mixed methods design. Combining quantitative data and qualitative narratives clarifies intervention effects and their underlying mechanisms—an integrated approach that remains underused in digital therapeutics research.

XSL•FO
RenderX

Conducted in routine outpatient settings without additional resource input, the study demonstrates good ecological validity and practicality. The proposed patient empowerment pathway also provides a new perspective for exploring digital health engagement among patients with CKD.

Several limitations of this study should be acknowledged. First, the 24.2% (24/99) completion rate of 24-hour urine collection resulted in massive missing data, rendering ITT analysis and primary outcome testing invalid. Consequently, the true intervention effect could not be confirmed. Second, qualitative participants were limited to intervention completers, which may introduce selection bias, potentially overestimating engagement and underestimating implementation barriers, even after sensitivity analysis. Third, the single-center setting and small sample size limit the generalizability of findings across geographic, demographic, and statistical dimensions. Fourth, we did not analyze outcome differences across CKD stages 1 to 5, further limiting the applicability of the results.

Conclusions

This exploratory mixed methods study indicated that the nurse-led digital sodium restriction intervention was associated with modest gains in participants' sodium-related knowledge as a secondary outcome. However, this study encountered prominent practical difficulties in implementing high-load outcome monitoring, with the completion rate of 24-hour urine collection reaching only 24.2% (24/99). The misalignment between idealized trial protocols and patients' daily realities underscores the need for more feasible, patient-centered outcome measures and adaptive intervention designs. Rather than drawing conclusions about efficacy, this study provides critical insights into implementation barriers and informs the development of future person-centered digital tools. Digital health interventions may be improved via technical enhancements and better adaptation to patients' daily lives. Future research should use implementation science and mixed methods to explore intervention mechanisms, thereby narrowing the gap between health knowledge and self-management in chronic disease management.

Acknowledgments

We would like to express our sincere thanks to all the participants who participated in this project and helped us complete the entire study. We also thank Senmei Company (Chongqing) for providing us with technical support and the Department of Nephrology at the Second Affiliated Hospital of the Army Military Medical University for providing us with the research platform, which ensured that our research was successfully completed. Generative AI was used for post-text verification.

Funding

This study was supported by the Chongqing Technology Innovation and Application Development (Key Research and Development Project; grant number CSTB2023TIAD-KPX0060) and Chongqing Municipal Science-Technology Bureau and Health Commission jointly funded the project (grant number 2022MSXM026).

Authors' Contributions

Conceptualization: J Zhao, YS
Data curation: HP, SP, XT, YL, YZ
Formal analysis: J Zhang, YS
Funding acquisition: J Zhao, YS
Investigation: HP, J Zhang, SP, XT, YL, YS, YZ
Methodology: HP, J Zhao, YS,
Project administration: J Zhang, J Zhao, YS
Resources: HP
Supervision: J Zhao
Validation: YS
Visualization: HP
Writing – original draft: YS
Writing – review & editing: J Zhao, YS

The first author is the principal investigator of this funded research project and takes full responsibility for study design, data collection, statistical analysis, manuscript drafting and final submission. All authors confirm that there are no commercial benefits, industry sponsorships, paid consulting relationships or other potential financial conflicts of interest that may bias the research findings.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Sodium salt knowledge questionnaire.
[PDF File, 109 KB - nursing_v9i1e94330_app1.pdf]

Multimedia Appendix 2
Patient interview guide and coding codebook.
[DOCX File, 37 KB - nursing_v9i1e94330_app2.docx]

Checklist 1
CONSORT-eHEALTH checklist (V 1.6).
[PDF File, 193 KB - nursing_v9i1e94330_app3.pdf]

Checklist 2
COREQ checklist.
[PDF File, 98 KB - nursing_v9i1e94330_app4.pdf]

Checklist 3
GRAMMS checklist.
[PDF File, 97 KB - nursing_v9i1e94330_app5.pdf]

References

1. Rafferty Q, Stafford LK, Vos T, GBD 2023 Kidney Failure with Replacement Therapy Collaborators, et al. Global, regional, and national prevalence of kidney failure with replacement therapy and associated aetiologies, 1990–2023: a systematic analysis for the global burden of disease study 2023. Lancet Glob Health 2025 Aug;13:e1378-e1395. [doi: 10.2139/ssrn.5187175]
2. Chan W. Chronic kidney disease and nutrition support. Nutr Clin Pract 2021 Apr;36(2):312-330. [doi: 10.1002/ncp.10658] [Medline: 33734473]
3. Romagnani P, Agarwal R, Chan JCN, et al. Chronic kidney disease. Nat Rev Dis Primers 2025 Jan 30;11(1):8. [doi: 10.1038/s41572-024-00589-9] [Medline: 39885176]
4. Borrelli S, Provenzano M, Gagliardi I, et al. Sodium intake and chronic kidney disease. Int J Mol Sci 2020 Jul 3;21(13):4744. [doi: 10.3390/ijms21134744] [Medline: 32635265]
5. Kang M, Kang E, Ryu H, et al. Measured sodium excretion is associated with CKD progression: results from the KNOW-CKD study. Nephrol Dial Transplant 2021 Feb 20;36(3):512-519. [doi: 10.1093/ndt/gfaa107] [Medline: 32582942]
6. Nerbass FB, Calice-Silva V, Pecoits-Filho R. Sodium intake and blood pressure in patients with chronic kidney disease: a salty relationship. Blood Purif 2018;45(1-3):166-172. [doi: 10.1159/000485154] [Medline: 29478050]
7. Schrauben SJ, Inamdar A, Yule C, et al. Effects of dietary app-supported tele-counseling on sodium intake, diet quality, and blood pressure in patients with diabetes and kidney disease. J Ren Nutr 2022 Jan;32(1):39-50. [doi: 10.1053/j.jrn.2021.08.006] [Medline: 34649784]
8. Sun L, Duan S, Zuo C, et al. Characteristics of urinary sodium excretion in patients with chronic kidney disease in Jiangsu, China. J Clin Hypertens (Greenwich) 2021 Jul;23(7):1372-1379. [doi: 10.1111/jch.14307] [Medline: 34132028]
9. Xing XY, Zhao Y, Sam NB, et al. Salt reduction behavior of adults in Anhui province in 2019: a cross-sectional survey of 3,378 participants. Front Public Health 2023;11:1242969. [doi: 10.3389/fpubh.2023.1242969] [Medline: 37908687]
10. Digital therapeutics definition and core principles. : Digital Therapeutics Alliance (DTA); 2019 URL: https://dtxalliance.org/wp-content/uploads/2021/01/DTA_DTx-Definition-and-Core-Principles.pdf [accessed 2026-06-26]
11. Hong JS, Wasden C, Han DH. Introduction of digital therapeutics. Comput Methods Programs Biomed 2021 Sep;209:106319. [doi: 10.1016/j.cmpb.2021.106319] [Medline: 34364181]
12. Pereira AP, Janela D, Areias AC, et al. Innovating care for postmenopausal women using a digital approach for pelvic floor dysfunctions: prospective longitudinal cohort study. JMIR Mhealth Uhealth 2025 Apr 2;13:e68242. [doi: 10.2196/68242] [Medline: 40173388]
13. Hartweg DL, Metcalfe SA. Orem's Self-Care Deficit Nursing Theory: relevance and need for refinement. Nurs Sci Q 2022 Jan;35(1):70-76. [doi: 10.1177/08943184211051369] [Medline: 34939484]
14. Pongpoottipatchara R, Anonjarn K, Prawatwong W, Singha-Dong N. Implementation of digital device-assisted and nurse-led case management to promote self-management in adults with noncommunicable diseases: protocol for a single-arm intervention study. JMIR Res Protoc 2026 Jan 22;15:e80083. [doi: 10.2196/80083] [Medline: 41569800]
15. Hua J, Sun D, Wang H, et al. Development of a nurse-led mHealth intervention framework for patients with chronic diseases: a systematic review and Delphi study. Digit Health 2025;11:20552076251387050. [doi: 10.1177/20552076251387050] [Medline: 41132439]

XSL•FO
RenderX

16. Forma F, Pratiwadi R, El-Moustaid F, Smith N, Thorndike F, Velez F. Network meta-analysis comparing the effectiveness of a prescription digital therapeutic for chronic insomnia to medications and face-to-face cognitive behavioral therapy in adults. Curr Med Res Opin 2022 Oct;38(10):1727-1738. [doi: 10.1080/03007995.2022.2108616] [Medline: 35938209]
17. Kario K, Harada N, Okura A. Digital therapeutics in hypertension: evidence and perspectives. Hypertension 2022 Oct;79(10):2148-2158. [doi: 10.1161/HYPERTENSIONAHA.122.19414] [Medline: 35726619]
18. Kario K, Nomura A, Harada N, et al. Efficacy of a digital therapeutics system in the management of essential hypertension: the HERB-DH1 pivotal trial. Eur Heart J 2021 Oct 21;42(40):4111-4122. [doi: 10.1093/eurheartj/ehab559] [Medline: 34455443]
19. Lutz J, Offidani E, Taraboanta L, Lakhan SE, Campellone TR. Appropriate controls for digital therapeutic clinical trials: a narrative review of control conditions in clinical trials of digital therapeutics (DTx) deploying psychosocial, cognitive, or behavioral content. Front Digit Health 2022;4:823977. [doi: 10.3389/fdgth.2022.823977] [Medline: 36060538]
20. Greenwood SA, Young HML, Briggs J, et al. Evaluating the effect of a digital health intervention to enhance physical activity in people with chronic kidney disease (Kidney BEAM): a multicentre, randomised controlled trial in the UK. Lancet Digit Health 2024 Jan;6(1):e23-e32. [doi: 10.1016/S2589-7500(23)00204-2] [Medline: 37968170]
21. McLean RM. Measuring population sodium intake: a review of methods. Nutrients 2014 Oct 28;6(11):4651-4662. [doi: 10.3390/nu6114651] [Medline: 25353661]
22. Liu W, Yu X, Wang J, et al. Improving kidney outcomes in patients with nondiabetic chronic kidney disease through an artificial intelligence-based health coaching mobile app: retrospective cohort study. JMIR Mhealth Uhealth 2023 Jun 1;11:e45531. [doi: 10.2196/45531] [Medline: 37261895]
23. Ma J, Wang J, Ying J, et al. Long-term efficacy of an AI-based health coaching mobile app in slowing the progression of nondialysis-dependent chronic kidney disease: retrospective cohort study. J Med Internet Res 2024 Nov 25;26:e54206. [doi: 10.2196/54206] [Medline: 39402012]
24. Lopez-Alcalde J, Wieland LS, Yan Y, et al. Methodological challenges in randomized controlled trials of mhealth interventions: cross-sectional survey study and consensus-based recommendations. J Med Internet Res 2024 Dec 19;26:e53187. [doi: 10.2196/53187] [Medline: 39700488]
25. Steinman L, Chadwick K, Chavez Santos E, et al. Remote evidence-based programs for health promotion to support older adults during the COVID-19 pandemic and beyond: mixed methods outcome evaluation. JMIR Aging 2024 Jun 13;7:e52069. [doi: 10.2196/52069] [Medline: 38869932]
26. Younas A, Kvist T, Stolt M. Framework and key considerations for designing and conducting critical mixed-methods research (cMMR). J Epidemiol Community Health 2025 Dec 10;80(1):55-61. [doi: 10.1136/jech-2025-224437] [Medline: 40983512]
27. Östlund U, Kidd L, Wengström Y, Rowa-Dewar N. Combining qualitative and quantitative research within mixed method research designs: a methodological review. Int J Nurs Stud 2011 Mar;48(3):369-383. [doi: 10.1016/j.ijnurstu.2010.10.005] [Medline: 21084086]
28. Inker LA, Eneanya ND, Coresh J, et al. New creatinine- and cystatin C-based equations to estimate GFR without race. N Engl J Med 2021 Nov 4;385(19):1737-1749. [doi: 10.1056/NEJMoa2102953] [Medline: 34554658]
29. Shi Y, Peng H, Chen Y, et al. Implementation of cluster-based management strategies for patients with chronic kidney disease. J Nurse Pract 2020 Oct;16(9):689-694. [doi: 10.1016/j.nurpra.2020.07.003]
30. Shi Y, Pu S, Peng H, et al. Impact of mobile application and outpatient follow-up on renal endpoints and physiological indices in patients with chronic kidney disease: a retrospective cohort study in Southwest China. BMC Med Inform Decis Mak 2024 Jun 12;24(1):163. [doi: 10.1186/s12911-024-02567-3] [Medline: 38867251]
31. Humalda JK, Klaassen G, de Vries H, et al. A self-management approach for dietary sodium restriction in patients with CKD: a randomized controlled trial. Am J Kidney Dis 2020 Jun;75(6):847-856. [doi: 10.1053/j.ajkd.2019.10.012] [Medline: 31955921]
32. Cogswell ME, Wang CY, Chen TC, et al. Validity of predictive equations for 24-h urinary sodium excretion in adults aged 18-39 y. Am J Clin Nutr 2013 Dec;98(6):1502-1513. [doi: 10.3945/ajcn.113.059436] [Medline: 24047921]
33. Muntner P, Shimbo D, Carey RM, et al. Measurement of blood pressure in humans: a scientific statement from the American Heart Association. Hypertension 2019 May;73(5):e35-e66. [doi: 10.1161/HYP.0000000000000087] [Medline: 30827125]
34. Korevaar JC, Merkus MP, Jansen MAM, et al. Validation of the KDQOL-SF: a dialysis-targeted health measure. Qual Life Res 2002 Aug;11(5):437-447. [doi: 10.1023/a:1015631411960] [Medline: 12113391]
35. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
36. Conner M, Wilding S, Wright CE, Sheeran P. How does self-control promote health behaviors? A multi-behavior test of five potential pathways. Ann Behav Med 2023 Apr 12;57(4):313-322. [doi: 10.1093/abm/kaac053] [Medline: 36206193]
37. Kelly MP, Barker M. Why is changing health-related behaviour so difficult? Public Health 2016 Jul;136:109-116. [doi: 10.1016/j.puhe.2016.03.030] [Medline: 27184821]
38. Phipps DJ, Hagger MS, Hamilton K. Evidence inhibitory self-control moderates effects of habit on complex but not simple health behaviors. Appl Psychol Health Well Being 2025 Feb;17(1):e12642. [doi: 10.1111/aphw.12642] [Medline: 39710871]

XSL•FO
RenderX

39. Yang MZ, Conner M, Sheeran P. The role of affect regulation in health behavior change. Health Psychol 2025 Sep;44(9):876-885. [doi: 10.1037/hea0001507] [Medline: 40244985]
40. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 4;3(4):e101. [doi: 10.2196/mhealth.4924] [Medline: 26537656]
41. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. J Med Internet Res 2019 Jul 22;21(7):e12869. [doi: 10.2196/12869] [Medline: 31333198]
42. Wang Y, Min J, Khuri J, et al. Effectiveness of mobile health interventions on diabetes and obesity treatment and management: systematic review of systematic reviews. JMIR Mhealth Uhealth 2020 Apr 28;8(4):e15400. [doi: 10.2196/15400] [Medline: 32343253]
43. Zhong C, Luo X, Tan M, et al. Digital health interventions to improve mental health in patients with cancer: umbrella review. J Med Internet Res 2025 Feb 21;27:e69621. [doi: 10.2196/69621] [Medline: 39984165]
44. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
45. Palmer SM, Knoblauch ST, Winham DM, Hiller MB, Shelley MC. Putting knowledge into practice: low-income women talk about food choice decisions. Int J Environ Res Public Health 2020 Jul 15;17(14):5092. [doi: 10.3390/ijerph17145092] [Medline: 32679700]
46. Gao J, Pan Q, Li SL, et al. Influencing factors of health promotion behaviour in patients with aortic dissection: a qualitative study using the COM-B model. BMJ Open 2025 Mar 5;15(3):e076181. [doi: 10.1136/bmjopen-2023-076181] [Medline: 40044200]
47. Liu X, Zhang J, Guan X, Cao Y. Barriers and facilitators to cognitive function interventions in rural diabetic older adults: using the COM-B model and theoretical domains framework. J Adv Nurs 2026 Mar;82(3):2320-2331. [doi: 10.1111/jan.70030] [Medline: 40538250]
48. Reedman SE, Jayan L, Boyd RN, Ziviani J, Elliott C, Sakzewski L. Descriptive contents analysis of ParticiPAte CP: a participation-focused intervention to promote physical activity participation in children with cerebral palsy. Disabil Rehabil 2022 Nov;44(23):7167-7177. [doi: 10.1080/09638288.2021.1985636] [Medline: 34624202]
49. Ogata S, Akashi Y, Sakusabe T, et al. A multiple 24-hour urine collection study indicates that kidney function decline is related to urinary sodium and potassium excretion in patients with chronic kidney disease. Kidney Int 2022 Jan;101(1):164-173. [doi: 10.1016/j.kint.2021.10.030] [Medline: 34774553]
50. Carnes K, Howe A, Feustel PJ, Listman JA, White M, Kogan BA. 24-hour urine collection for first time pediatric stone formers: is it worth it? J Pediatr Urol 2021 Jun;17(3):387. [doi: 10.1016/j.jpurol.2020.12.001] [Medline: 33762156]
51. Santos JA, Bolton KA, Rosewarne E, et al. Agreement between 24-hour urine and 24-hour food recall in measuring salt intake in primary school children in Australia. Nutr J 2022 Nov 15;21(1):68. [doi: 10.1186/s12937-022-00823-8] [Medline: 36380391]
52. Wan ER, Cross J, Sofat R, Walsh SB. 24-hour vs. spot urinary sodium and potassium measurements in adult hypertensive patients: a cohort validation study. Am J Hypertens 2019 Sep 24;32(10):983-991. [doi: 10.1093/ajh/hpz104] [Medline: 31247068]
53. Tagawa K, Tsuru Y, Yokoi K, Aonuma T, Hashimoto J. Being overweight worsens the relationship between urinary sodium excretion and albuminuria: the Wakuya study. Eur J Clin Nutr 2023 Nov;77(11):1044-1050. [doi: 10.1038/s41430-023-01327-2] [Medline: 37587243]
54. Bentzen SM. Artificial intelligence in health care: a rallying cry for critical clinical research and ethical thinking. Clin Oncol 2025 May;41:103798. [doi: 10.1016/j.clon.2025.103798]
55. Inagaki K, Tsuriya D, Hashimoto T, Nakamura K. Verification of the reliability of an automated urine test strip colorimetric program using colorimetric analysis: survey study. JMIR Form Res 2025 Jan 14;9:e62772. [doi: 10.2196/62772] [Medline: 39810396]
56. Xu H, Hua Y, Li H, Wang J, Liu G, Li X. Smartphone-based colorimetric correction analysis system for long-term urine monitoring of chronic kidney disease. Anal Biochem 2025 Jul;702:115824. [doi: 10.1016/j.ab.2025.115824] [Medline: 40024322]
57. Duffy A, Boroumandzad N, Sherman AL, Christie G, Riadi I, Moreno S. Examining challenges to co-design digital health interventions with end users: systematic review. J Med Internet Res 2025 Mar 14;27:e50178. [doi: 10.2196/50178] [Medline: 40085834]
58. Main A, McCartney H, Ibrar M, et al. Patients' experiences of digital health interventions for the self-management of chronic pain: systematic review and thematic synthesis. J Med Internet Res 2025 Mar 18;27:e69100. [doi: 10.2196/69100] [Medline: 40101209]
59. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]

XSL•FO
RenderX

60. Nahum-Shani I, Hekler EB, Spruijt-Metz D. Building health behavior models to guide the development of just-in-time adaptive interventions: a pragmatic framework. Health Psychol 2015 Dec;34S:1209-1219. [doi: 10.1037/hea0000306] [Medline: 26651462]
61. van Genugten CR, Thong MSY, van Ballegooijen W, et al. Beyond the current state of just-in-time adaptive interventions in mental health: a qualitative systematic review. Front Digit Health 2025;7:1460167. [doi: 10.3389/fdgth.2025.1460167] [Medline: 39935463]
62. Willms A, Liu S. Exploring the feasibility of using ChatGPT to create just-in-time adaptive physical activity mHealth intervention content: case study. JMIR Med Educ 2024 Feb 29;10:e51426. [doi: 10.2196/51426] [Medline: 38421689]
63. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 May 10;19(5):e153. [doi: 10.2196/jmir.7725] [Medline: 28490417]

Abbreviations

CKD: chronic kidney disease
GRAMMS: Good Reporting of A Mixed Methods Study
ITT: intention-to-treat
KDQOL-SF: Kidney Disease Quality of Life-Short Form
MMSE: Mini-Mental State Examination
QOL: quality of life
RCT: randomized controlled trial
SF-36: 36-Item Short Form Survey

Edited by E Borycki; submitted 27.Feb.2026; peer-reviewed by FJ Yang, H Arooj, Q Jiang-Siebert; revised version received 13.Jun.2026; accepted 17.Jun.2026; published 20.Jul.2026.

Please cite as:
Shi Y, Peng H, Zhang Y, Zhang J, Li Y, Pu S, Tang X, Zhao J
Nurse-Led Digital Intervention for Sodium Restriction in Chronic Kidney Disease: Mixed Methods Implementation Study
JMIR Nursing 2026;9:e94330
URL: https://nursing.jmir.org/2026/1/e94330
doi:10.2196/94330

© Yu Shi, Hongmei Peng, Youying Zhang, Jingyu Zhang, Yang Li, Shi Pu, Xiangchun Tang, Jinghong Zhao. Originally published in JMIR Nursing (https://nursing.jmir.org), 20.Jul.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Patients' Perceptions of the Role of Nursing in Substance Use Disorder Treatment Programs: Qualitative Study

Clara Lucas Guerra[1,2,3*], MSN, PhD; Marina Gómez de Quero Córdoba[4*], MSN, PhD; Bruno García Tardón[2,3,5*], PhD; Guadalupe González Naranjo[2,3*], MSN; Noemí Mayoral Gonzalo[2,3*], MSN

[1]Faculty of Nursing, Physiotherapy and Podiatry, Complutense University of Madrid, C/Emiliano Barral, 18A, 2º3, Madrid, Spain

[2]Faculty of Health Sciences - HM Hospitals, Camilo Jose Cela University, Villanueva de la Cañada, Madrid, Spain

[3]HM Hospitals Health Research Institute, Madrid, Spain

[4]Research Group on Advanced Nursing (CARING)-161, Rovira i Virgili University, Tarragona, Spain

[5]Faculty of Education, Camilo Jose Cela University, Madrid, Spain

[*]all authors contributed equally

Corresponding Author:
Clara Lucas Guerra, MSN, PhD
Faculty of Nursing, Physiotherapy and Podiatry, Complutense University of Madrid, C/Emiliano Barral, 18A, 2º3, Madrid, Spain

Abstract

Background: Substance use disorders are chronic conditions with significant personal and social consequences. Nursing care plays a key role in outpatient and community-based rehabilitation programs, yet patients' perspectives on this role remain underexplored.

Objective: This study aimed to explore how patients with substance use disorders perceive and interpret nursing care in community addiction treatment centers operated by the Spanish Red Cross in Madrid, Spain. Specifically, it sought to describe the organizational and care-related roles attributed to community addiction treatment centers and analyze patients' perceptions of nurses' technical and relational functions.

Methods: A phenomenological qualitative design was used. Fourteen in-depth, semistructured interviews were conducted with patients undergoing treatment at Red Cross centers. Participants were selected through purposive sampling to ensure diversity in age, gender, substance use, and treatment experiences. The data were analyzed using systematic text condensation and supported by the ATLAS.ti software.

Results: The center is perceived as a significant space not only for offering healthy leisure activities outside the context of substance use but also as a supportive environment that fosters a sense of belonging to a community. Patients valued the emotional support, empathy, and relational care provided by nurses, often highlighting their role in building trust and offering personalized attention. However, there was limited awareness of nurses' technical competencies.

Conclusions: These findings underscore the importance of holistic, patient-centered care and the need to enhance the visibility and recognition of nursing roles in addiction treatment settings.

(JMIR Nursing 2026;9:e82401) doi:10.2196/82401

KEYWORDS

nurse's role; nursing staff; qualitative research; substance-related disorder; psychiatric nursing

Introduction

Background

Substance use disorders (SUDs) are defined by the *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*) as the "association of cognitive, behavioral, and physiological symptoms that indicate that the person continues to use a substance despite significant problems related to that substance" [1]. Globally, substance use represents a major public health concern and contributes substantially to the burden of disease, being associated with increased mortality, disability, and chronic conditions, such as liver disease, cancer, and infectious diseases [2-4]. SUDs commonly coexist with psychiatric disorders, particularly mood and anxiety disorders, which complicate treatment and worsen prognosis [3,5,6]. Consequently, SUDs can have a profound impact on a person's identity, causing individuals to lose their sense of self, which can manifest itself as a change in values and priorities, a deterioration of self-image, social disconnection, or a break with previous roles [5,6].

This individual impact underscores why SUDs represent a global challenge with serious consequences at the individual level. In 2025, there were 51,255 admissions to treatment for

XSL•FO
RenderX

psychoactive substance dependence in Spain, representing an increase of 9.7% over the previous year [7]. Moreover, adherence to rehabilitation programs is often limited, with high dropout rates reported internationally and frequently linked to psychosocial vulnerability and difficulties in establishing therapeutic relationships with health care professionals [3].

Beyond physical and mental health consequences, SUDs also affect identity and social functioning, contributing to stigma, social disconnection, and loss of social roles [3]. In Europe, treatment demand continues to rise, with 51,255 admissions for psychoactive substance dependence recorded in 2025, representing a 9.7% increase compared with the previous year [7,8].

Although SUDs affect people of all genders, treatment-seeking populations in community addiction services are often predominantly male, which may lead to the underrepresentation of women's experiences in qualitative studies [6-8]. This is of relevance given the established role of gender in shaping pathways into substance use, the associated stigma, exposure to violence or trauma, caregiving responsibilities, and help-seeking behaviors. These factors might shape how nursing care is perceived and what forms of support are most beneficial. Therefore, incorporating a gender perspective might be essential when interpreting patients' accounts and designing nursing interventions in community addiction treatment centers [9].

Prior Work

Evidence-based rehabilitation programs have been shown to reduce the use of illegal opioids and improve patients' physical and mental well-being, as well as their quality of life [8]. These programs adopt a holistic approach, with a multidisciplinary team comprising nurses, psychologists, physicians, social workers, and occupational therapists. Furthermore, the centers where these programs are implemented administer substitution treatment and provide social support and nursing care [8,10]. The ideal outcome of substance use rehabilitation is a state of balance and control, where the person feels empowered to manage their life without relying on substances that alter their mental state [7,8]. This process involves a reconstruction of the individual's identity through a series of behavioral modifications, facilitated by health care professionals [7,9,10]. However, it is estimated that up to 40% of patients withdraw from these programs. In this context, the support and participation of a multidisciplinary team, including nurses, is an empowering element in addressing this condition [11].

The concept of care is understood as a natural phenomenon. The patient's world, vulnerability, health, and suffering form the core of this concept [12]. The science of care not only facilitates the treatment of disease but also addresses patients' emotional and social needs, thereby promoting their overall well-being. The quality of care is known to improve when health care professionals make a conscious effort to understand the world from the patient's perspective. This, in turn, strengthens the therapeutic relationship and promotes better care outcomes. A profound understanding of the experiences and perspectives of patients with SUDs is imperative for the development of patient-centered care. This knowledge facilitates the design of personalized interventions, tailored to the specific needs of this vulnerable population [13,14]. By exploring the subjective experiences of these patients in relation to their care, areas for improvement in health services can be identified, more empathetic and culturally sensitive practices can be promoted, and barriers to access and adherence to treatment can be addressed. This comprehensive strategy for addressing SUDs has the capacity to markedly enhance patients' quality of life, diminish the social disapproval associated with substance use, and ensure optimal long-term therapeutic outcomes [15].

Despite the presence of nursing in SUD rehabilitation programs, the literature indicates that its role continues to be invisible and poorly understood by both patients and society. In community and outpatient settings, nursing work tends to be diluted within the multidisciplinary team, making it difficult to identify its specific functions [13,14]. However, in addiction treatment centers (CATCs), nurses carry out key interventions that include comprehensive patient assessment, clinical follow-up, health education, administration and supervision of substitution treatments, early detection of relapses, crisis care, and, especially, the establishment of a therapeutic relationship based on trust and continuous support during the rehabilitation process [16]. This apparent contradiction between the relevance of care and the low social recognition of the nursing role justifies the need to explore how users themselves perceive and understand this care [13,17].

Study Objectives

However, there is a significant gap in the literature on the perception of patients with SUDs regarding the role of nurses in rehabilitation programs [16,17].

The primary aim of this study is to explore the experiences and perceptions of patients with SUDs regarding the nursing care they receive in outpatient and community settings in Spain. Specifically, the study seeks to analyze patients' perception of the role of nursing staff, considering both the technical and relational dimensions of nursing care as experienced by users and to describe the organizational and care role that users attribute to Red Cross CATCs in Madrid (Spain), including resources, routines, and processes involved.

Methods

The question of how the role of nursing is perceived in the context of patients with SUDs undergoing rehabilitation programs lends itself to a qualitative approach, specifically a phenomenological study. Qualitative research focuses on understanding the world of the participants [18].

Study Design

The researchers were interested in how patients recover from SUDs and how they perceive the role of nurses in this process. It was therefore determined that qualitative research would best meet the objective. To gain a deeper understanding of this experience, a phenomenological study was conducted, as in-depth interviews allow the emotional, social, and personal nuances of the rehabilitation process to be captured, as well as the importance of support from nurses [19-21]. This research was conducted based on the recommendations of the COREQ guide [22].

XSL•FO
RenderX

Participants

Patients diagnosed with SUDs according to the *DSM-5* [1] were selected.

Eligibility Criteria

The inclusion criteria established at the beginning of the study were as follows: participants had to be enrolled in a rehabilitation program, be over 18 years of age, and be able to speak Spanish. Patients who did not wish to participate in the study or who were physically and/or mentally unable to complete the interview properly were excluded from the study.

Study Sample

A sample of participants was intentionally selected to represent a range of age groups, socioeconomic backgrounds, recovery times, genders, and experiences regarding substance use and the impact of nurses during the rehabilitation process. Furthermore, the data were collated pertaining to the participants' date of birth, gender, marital status, substance of abuse, alcohol consumption, and number of prior admissions to rehabilitation programs (Multimedia Appendix 1).

Of the 15 participants, 14 were male, and 1 was female. The ages of the participants ranged from 32 to 78 years (Table S1 in Multimedia Appendix 1). The participants' demographic profile was as follows: 8 were unmarried, 4 were divorced, 2 were married, and none were widowed. The participants consumed a variety of substances, including alcohol, cannabis, cocaine, heroin, opioids, methamphetamines, ecstasy, ketamine, hallucinogens and magic mushrooms, spice, mephedrone, salvia, gamma hydroxybutyrate, and lysergic acid diethylamide. The substances most frequently used were cannabis, heroin, cocaine, and alcohol. The number of previous admissions to rehabilitation programs ranged from 1 to 10. All participants were enrolled in the Red Cross rehabilitation program in Madrid, Spain.

Access to the Population

Initially, the researchers identified the various support structures for treatment and posttreatment within the Red Cross organization. In-depth interviews were conducted at 2 Red Cross CATCs in Madrid (Spain).

The management team and the multidisciplinary team at the centers were contacted, and meetings were organized to discuss the research objective and obtain the help of the various professionals in identifying participants. Participants were recruited through purposive sampling, following a preliminary presentation of the study to patients at the Red Cross centers together with the center's management team. Factors that would facilitate a more nuanced categorization of the work were considered, particularly regarding the perception of the nursing role in individual interviews.

Data Collection

In-depth, individual, semistructured interviews were conducted in person with 15 participants. One participant decided to leave the interview after 14 minutes of recording, so a total of 14 patients completed the interviews. The average interview duration was 45 (SD 15.3) minutes, and the interviews were conducted in convenient settings, such as individual rooms at the CATCs, to ensure privacy. The interviews were conducted between January and April 2023. All interviews commenced with the following opening statement: "Please provide a brief overview of your professional background and the circumstances that led to your involvement in this field." The flexibility of this qualitative interview methodology enabled participants to share their experiences from their own perspective. Open-ended questions and prompts were strategically used to explore turning points, support, situations, and coping strategies when participants were reluctant to share them. The following questions and prompts were included: "Please find below a list of the key responsibilities that nurses in this institution typically have" (Multimedia Appendix 1). With the prior consent of the participants, the interviews were audio recorded. The data collection protocol for the interviews was consistent for all patients. Interviews were transcribed immediately after completion. Data collection ceased once sample saturation was reached [23].

Trustworthiness

The criteria used to ensure reliability, in accordance with Guba's contributions to naturalistic research, encompass credibility, transferability, dependability, and confirmability [24].

Credibility

All interviews were recorded, and the exact duration of each session was documented, with additional field notes providing further elaboration and complement.

Transference

The processes followed, both for data collection and subsequent analysis, have been described in detail.

Dependence

The data collection process and the instrument used for conducting the interview have been delineated and included (Multimedia Appendix 1). Furthermore, members have been added to the working team at different stages to carry out triangulation processes in the negotiation of meanings.

Neutrality

The saturation of messages and the inclusion of direct quotes in the final report ensure neutrality in the analysis of the data.

Analysis

The transcriptions of these recordings were then produced. Subsequently, a comparison was made between the recordings and transcripts to ensure accuracy. Thereafter, a systematic condensation of the text was carried out in the analytical process, with the following 4 steps [25]: first, the transcripts were read to obtain an overview and identify preliminary themes, with annotations made in the text. Second, a series of codes and code groups were developed. In the event of the emergence of divergent codes, the most appropriate ones were agreed upon to develop a shared understanding. Third, each code group was identified and condensed into categories, which were incorporated into the report, with relevant quotes selected for inclusion. Finally, the condensed subgroups were synthesized to generalize descriptions and concepts of the patients' experiences [26,27]. The data analysis was iterative and began

XSL•FO
RenderX

concurrently with data collection, using the constant comparison method as described by Strauss and Corbin [23] to identify commonalities and variations across cases progressively incorporated.

The ATLAS.ti program (version 25.0.1) was used for the analysis, as were the Office suite and Microsoft Excel for descriptive data analysis, and Microsoft Teams for facilitating digital collaboration during coding.

Ethical Considerations

This study has been approved by the Ethics Committee of the Complutense University of Madrid (Reference: CE_20221215 - 06_SAL). Furthermore, the study was conducted in accordance with the principles set out in the Declaration of Helsinki. In addition, to comply with the code of ethics for good practice in the collection of information for this research project, the resources provided by the Complutense University of Madrid have been accessed. The data have been anonymized, with all names appearing in the text being fictitious (Table S1 in Multimedia Appendix 1). Reasonable and appropriate physical, administrative, and technical measures have been taken to protect personal data from loss, misuse, unauthorized access, disclosure, alteration, or destruction. Participants were at liberty to accept or decline participation, and those who agreed to participate signed consent forms. It is important to note that no form of compensation was provided or offered to any of the participants.

Results

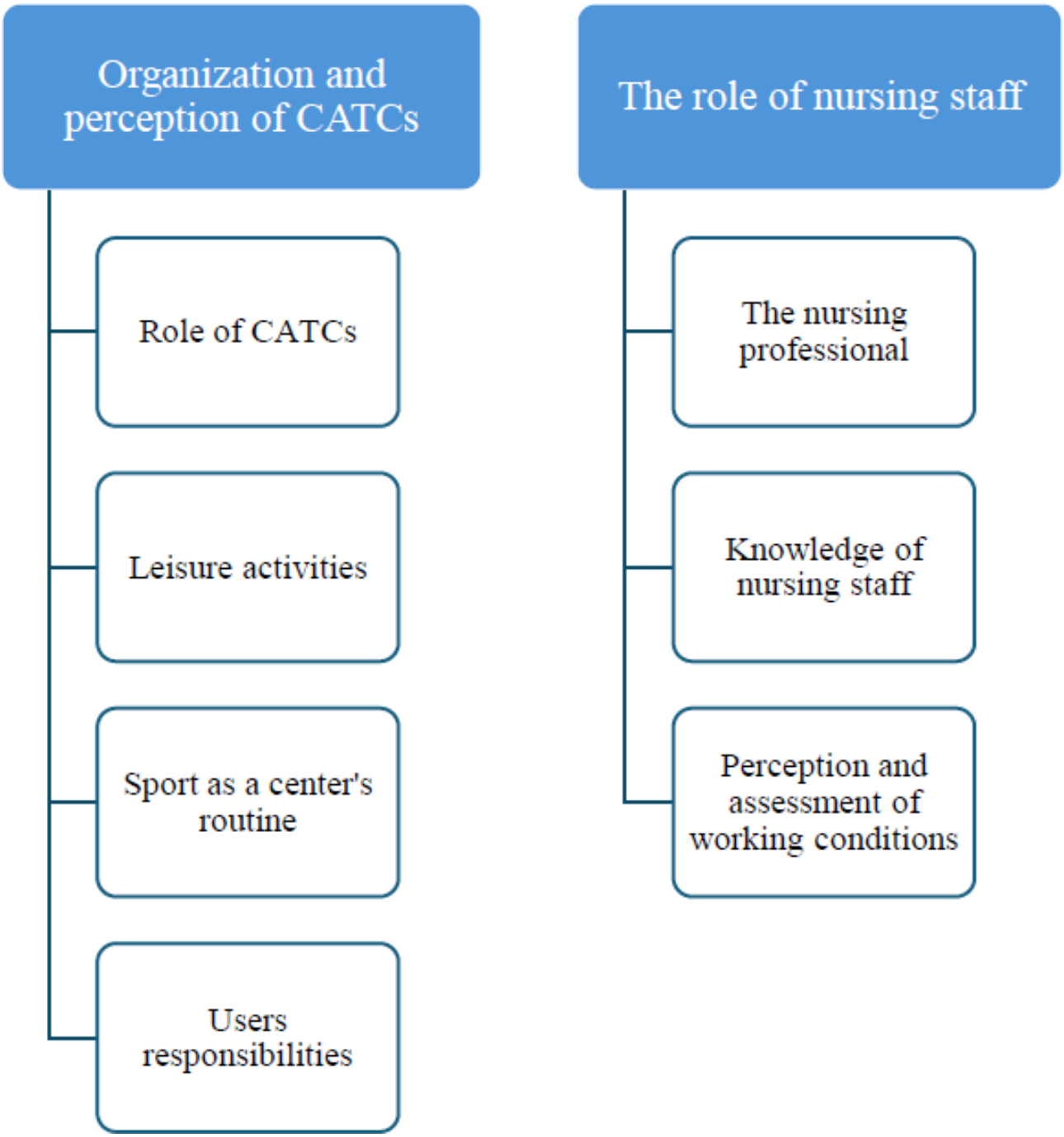
Participant Characteristics

All participants had received a confirmed diagnosis of SUDs prior to the interview. Of the 15 participants, 14 were male, 8 (54%) were single, 5 (33%) were divorced, and 2 (13%) were married. The mean age of the subjects was 48.2 years (SD 9.28). The substances most frequently consumed were cocaine—used by 13 (86%) of the participants—cannabis and alcohol (Multimedia Appendix 2). The mean number of previous admissions to rehabilitation centers within the sample was 2.5 (SD 2.4).

In the following section, the results of the empirical investigation will be presented. The data obtained from the interviews have been organized into 2 overarching categories, which are related to the research objectives and questions (Figure 1). All these categories have been framed from the perspective of the users.

XSL•FO
RenderX

Figure 1. Emerging categories from patient narratives: organization of community addiction treatment centers (CATCs) and the nursing role.



The Role of CATCs

Key Functions of CATCs

The role of Red Cross CATCs is pivotal in helping participants to stop using substances. The interviews revealed several issues. First, these centers play a significant role in helping individuals stop using substances, particularly through measures such as prescribing methadone as a substitute for heroin: “This is how I became acquainted with the CATCs and the process of heroin substitution” (Óscar). It was also noted that the center offers users the possibility to have meals on-site and to receive support with various arrangements, including transport and mobility-related matters. “Additionally, my transport card is being loaded, which is a modest amount of four euros and a few cents” (Óscar).

The same user commented on the existence of these rehabilitation centers. He said that he had gone “to a hospital because I didn’t know that the CATCs existed” (Óscar).

Leisure Activities

The importance of leisure activities in the daily lives of participants emerged as a recurring theme in the interviews, with all interviewees highlighting it as a central focus. Javier succinctly summarizes the routine of a typical day at the center, characterizing it as “attending therapy sessions.” This assertion is corroborated by the observations of other participants, such as Oscar, who states: “At the Red Cross therapy centre (...) it

XSL•FO
RenderX

helps me a lot because I was busy all day and I didn't have any nonsense going through my head." These centers, as a result, organize a variety of activities and "workshops on drugs, gender, reading ..." (Chema), which are emphasized in the interviews. Daniel elucidates the manner these centers provide support:

> *It's like, like, like a house where you can't just sleep. They also make us feel like this is our home. The Red Cross treats us very well and they really love us; they truly love us. I mean, it's not just a pat on the back and that's it. They are ... it's a complete embrace, from heart to heart.*

As well as the points raised by Chema, there have been repeated mentions of board games, music, and dance workshops. The latter has been of particular interest, given the work carried out by the volunteer running them:

> *She herself, who is not like all the other workshops that are all about that, about drugs, about you, about these, about others ... No, she plays music and says, "Now you do whatever you want" and "Now you dance, and now you ... and now everyone together here and now everyone there ... now everyone ..." I mean ... she makes you forget everything at that moment. No, you don't think about your problems, you don't have time to think about anything bad, or anything that could hurt you, that could ... She doesn't give you time, she doesn't let you. She's full of life.* [María]

Sport in the Center's Routines

Sport has also been mentioned in numerous interviews, often referred to as "sports activities in the morning" (Paco). To that end, the following activities have been outlined:

- "Playing table tennis or cycling or exercising." (Rodolfo)
- "Some days it's Tai Chi, other days they go walking." (Maria)
- "I played volleyball, which I hadn't played in years." (Roberto)
- "Now we are cycling." (Antonio)

In addition, sport has been referenced on a single occasion as a concomitant treatment for other conditions: "I come walking to get some exercise because of my diabetes" (Chema).

Users' Responsibilities in the Centers

Leisure time is also spent carrying out different responsibilities assigned by the center:

> *Oh, and if it's your turn to do the office work, you have to set the table. If it's your turn to wash up, you have to wash up too, and so on.* [Fulgencio]

> *"As for routines ... we have a plan too ... there's a ... a notebook where they write down when we have to prepare meals, wash the dishes ... That's how we keep busy. .. the manager of the centre writes it down for everyone. And on days when it's my turn, it's someone else's turn ... When it's my turn, I have to serve the tables, then clear the dishes, take out the rubbish bins ... and so on, routine stuff ... To clean the dishes, then dry them ... That's how it works, I don't know, to keep us busier.*" [Óscar]

> *Then one day it's your turn to wash the dishes and another week you're in charge. You're responsible for distributing the food and so on, but hey, they're supervising you.* [Rodolfo]

Functions of the Nursing Professional

Role of Nurses in CATCs

A variety of professionals are employed in these centers, including nurses, doctors, psychiatrists, psychologists, and occupational therapists. Nurses have been the central focus of the work, and they have taken up a large part of the interview time. The functions of nursing staff are those that have been most frequently mentioned in the coding process, in relation to the perception of this group. It is generally understood that the profession is held in high regard by its users. There are some very compelling testimonials in this regard:

> *At the Jiménez-Díaz Foundation, nurses treat you excellently. I can't say anything bad about healthcare here in Spain or about the nurses, they are all lovely and very professional.* [Roberto]

These comments are not isolated, as they are reinforced by other comments that echo the same sentiment.

For example, Roberto stated:

> *The nurse who works with the doctor [at the CATC] (...) is a lovely girl, she is super friendly, she is a sweetheart. Professionally, she is excellent, as they say. And also, she was very concerned and asked me: are you okay?*

Daniel also commented that "the treatment by the nursing staff [at the CATC] was very good, very good, very good." He particularly highlighted the treatment by the nurse above all others. When asked about the care that had meant the most to him, he pointed to the emotional support provided by a nurse:

> *I'm very clear about this. It's (...) a nurse I had at the shelter [at the CATC] (...), accompanying me to an ambulance because I have tinnitus. Then they were saying bad things to me, and she was the first person I talked about it. And that, and how she accompanied me and went down the stairs with me ... I was shaking because I didn't want to go to the ambulance, I didn't want to go there. I was terrified of being alone for even a second with the voice. And that's the most the nurses have done for me.* [Daniel]

On occasion, this positive perception was intertwined with the work of the entire team, as Antonio noted:

> *Here [at the CATC] we have a psychologist, a social worker, a nurse, the manager, very good people, the workers too. They get angry, which is normal sometimes, but they are people.*

The testimonies are intricately interwoven throughout the interview, thereby underscoring the human dimensions of the nursing professional. Various statements highlight the staff's

XSL•FO
RenderX

attentiveness and their provision of empathetic listening and support, in response to the most useful care they have received:

> *Therapies where you can talk, like two people. You know what I mean? Being able to sit down, two people. You, with a therapist, being able to sit down and have .. him being able to tell you about his problems and you being able to tell him your problems, having someone there to talk to unburden yourself. Telling him, "Look, I got out of bed and cried." I cried because of this, because of this and because of this, and being able to tell you ... your things and him being able to tell you his things.* [Jorge]

> *This girl … the girl on the other side, at the Red Cross [at the CATC]. She ... I saw that she was concerned about my situation the first day I went there. I ... I thought she was a psychologist rather than a nurse, from the time I spent talking to her and expressing my feelings, my experiences, where I came from, where I was going, and why I found myself where I was. I was there for, I don't know, two hours, three hours, talking to her, and she was a nurse, not a psychologist. And she was concerned about ... about how I was doing.* [Jero]

Even when pointing out other functions, Paco considered that

> *the way the staff [at the CATC] treat you is what matters most ... because that's what affects you the most. But then, of course, each person ... the pills and so on ... also have an influence. But ... the way they treat you is what attracts you the most.*

Rehabilitation programs were also identified as particularly valuable forms of care, with participants citing the significance of being listened to in difficult situations:

> *My nurse [at the primary care centre], who has helped me to trust her a little, weigh myself, tell her when I have used drugs and when I haven't, give thanks, I am grateful to the people who have been helping me since 2015. And even though I messed up and stuff, they did everything they could to help me (...) see what we needed.* [Óscar]

In other words, the same patient insisted:

> *She was teaching nursing classes here, for everyone. That was very positive because there were classes on all kinds of illnesses, how to use pills ... There were also classes on quitting smoking, how to quit smoking ... There are workshops for people who want to quit smoking too. These workshops are run by nurses. That's all I can say for now because I can't remember anything else.* [Óscar]

María was very clear on this point: "They [at the primary care center] ... do a bit of psychology and nursing, kind of combining the two." Another user, Jero, who had already pointed out the importance of nursing staff's ability to listen, emphasized the importance this can have in other therapeutic aspects, such as methadone use:

> *Well, if this is the first day—I said—"it shouldn't take too long." He [at the CATC] offered me something ..., which was methadone, and I said, "Look, I don't want that, I've never taken that in my life, but I'm afraid, because I've heard that there are people who spend their whole lives taking that crap." I said, "Well, let's start with a very, very small amount. Let's see if that's enough for you and then we'll reduce it." "Okay, I'll accept it, because of the way you've treated me, because I can see that you're someone who cares about the situation I'm in without knowing me at all, and because I've ... I've opened up to you and told you about ... my plans." It was really good, to be honest, that's the part I'm most grateful for.* [Jero]

Knowledge of Nursing Staff

During the interviews, references to the work of nurses were particularly prevalent, especially in discussions of more general topics and when the work functions of the entire team were intertwined. There are numerous examples of this phenomenon, particularly in relation to the role of the doctor. For instance, when asked about the duties of a nurse, Paco replied:

> *Well, we ... I, for example, go up to cognitive workshop, which was today ... to the girl in cognitive workshop [at the CATC]. And then, in case you have ... there's the doctor, the doctor in case something happens to you. I have a toothache and I ask him for ... a pill for toothache because it hurts and I still haven't ... I go to the medical centre by bus and I always forget to stop at medicine.*

Another patient intermingled the roles of different professionals in his reflections:

> *I haven't had much … much … much experience with nurses, because as I'm new, I haven't had much time. What I do have is a good day centre, which I've had for two days now, OK? And I only have experience with Susana, who is my psychologist, and with Dr. Felipe, I don't know if you know him. They are the only ones I have experience with.* [Roberto]

Comparisons were even drawn between nurses and doctors:

> *You know more. You are not doctors, but we can trust you more than doctors. I am quite bold, having been a salesperson all my life, so I don't mind, but ... I think a doctor can be more reserved than a nurse.* [Eduardo]

These relationships and role confusions are not exclusive to CATCs, as they also mentioned past experiences in which they highlighted the importance of nurses in primary care centers:

> *At the mutual insurance company, for example, this leg that I told you I hurt while cleaning ambulances, they treated me like a queen. Both in physiotherapy and in ... when it came to medication, creams, everything, and all for free.* [María]

> *Not everyone knows about nursing. There are things that if you have, if you have a wound here, you put*

XSL•FO
RenderX

> *Betadine on it or whatever and that's it. But there are things that they have to give you. They have to tell you, "Look, the stretch consists of squats, this, that and the other, OK?" (...) Well then, I think that time should also be used to ... I see that they devote a lot of time to older people [at the primary care centre].* [Maria]

Perception and Assessment of Working Conditions

Another noteworthy aspect that emerged from the interviews was related to working conditions, including issues of overburden:

> *Well, I can see that he's a bit overworked, but even so, he's always available [at the CATC]. Yes, if I could add something ... a little help, give him some assistance. Because he does tests, dispenses medication, sees patients ... You don't have to make an appointment with him. Because you talk to him beforehand and Julián is immediately available.* [Chema]

> *Appointments take a long time. For example, those at the public centre and at the outpatient clinics too. Sometimes they can't see you, even if it's an emergency, because they have someone else who maybe ... well, you say it yourself ... I mean, I don't pay for my medication, do I? Because I earn very little. So there are times when I've had to buy the medication because I needed it at that moment because they don't have enough nurses [referring to nurses in outpatient/primary care consultations]. They keep cutting back instead of ... providing more for the population.* [María]

This same patient, María, concluded very clearly: "I think people should be given more permanent jobs."

Discussion

Principal Results

Patients with SUDs perceive nursing care positively, although they predominantly highlight relational aspects, such as accompaniment, active listening, and trust. In some cases, there appears to be a lack of awareness regarding nursing functions, particularly those within the scientific and technical domains. The results of this study provide valuable insight into how patients with SUDs value and interpret the nursing care they receive in outpatient and community-based rehabilitation programs in Spain.

From a phenomenological perspective, understanding patients' perceptions of the role of nursing requires first delving into their personal history of substance use. This is not merely a chronological account of events; it is a complex network of lived experiences, emotions, social relationships, and meanings attributed to the substance and its effects [28]. The mean number of previous admissions to rehabilitation centers in the sample was 2.5 (SD 2.4). This aspect also frequently features in patients' accounts, highlighting the chronic nature of SUDs and the difficulties in achieving sustained rehabilitation over time [4,7,28].

Participants offered a largely favorable evaluation of the center, regarding it as a vital asset in their rehabilitation process. This positive perception appears to be associated with the multidisciplinary team, comprising doctors, psychologists, social workers, occupational therapists, and nurses. The team's role is to address both the physical and psychological aspects of treatment and provide comprehensive support that extends beyond a purely clinical approach [12,13,15]. In their testimonies, participants highlight the professional and emotional support they received, as well as the feeling of belonging to a safe and structured environment that not only distances them from problematic consumption environments but also provides for their basic needs and offers opportunities for improvement in various areas.

On the other hand, participating in a variety of group activities, such as physical exercise, healthy leisure initiatives, and workshops focusing on integrating patients into the workplace and fostering personal relationships, was considered vital for sustaining abstinence. Physical exercise is widely used as a complementary therapy in SUD treatment. It is considered a healthy practice, as well as a meaningful and socially integrative activity [29,30]. Similarly, the other leisure workshops offered at the center are seen as potentially transformative by providing entertainment in environments unrelated to substance use and fostering a sense of belonging to a social group [28,29]. One possible explanation for these findings is that the center is perceived not only as a protective factor against relapses, but also as a safe resource for well-being and ongoing support. This finding is consistent with other studies addressing this issue but contrasts with previous studies documenting experiences of stigmatization and discriminatory treatment in hospital settings. The role of the multidisciplinary team could be pivotal in generating a positive overall experience that encourages adherence to the rehabilitation program and social integration [30,31]. This approach is consistent with the Schlossberg transition process model, in which the authors recommend that, to effectively cope with a transition—namely, rehabilitation from SUDs—people need access to diverse types of support [32]. These support categories can range from interpersonal relationships (such as sponsors and spouses) to more distant support systems (such as institutions and communities), including rehabilitation centers [33]. Consequently, rehabilitation programs that adopt a holistic perspective—combining respect for patient autonomy with a flexible and adaptable support structure—align not only with international standards but also with a patient-centered approach [7,10].

Regarding nursing care, the participants in our study recognize nurses as key figures in the rehabilitation programs of the Red Cross CATCs. Most of the nursing practices identified as beneficial by the participants in this study pertain to relational care interventions. They emphasize their close working relationships, active listening skills, and emotional support as key differentiating factors compared to other health care professionals in the multidisciplinary team. Furthermore, they place a high value on the involvement of nursing staff, particularly in health education, support in times of crisis, and management of withdrawal symptoms. This finding aligns with

XSL•FO
RenderX

the results of other studies, which indicate that patients with SUDs perceive increased support and reduced judgment from nursing professionals [34,35]. As demonstrated in previous research, the provision of compassionate care is of paramount importance, even in relation to other populations [36,37].

Comparison With Prior Work

This evidence indicates that the therapeutic relationship is fundamental for patients, regardless of their age or care setting. Skills, such as listening, communicating, and providing reassurance, are valued and recognized as essential, not as inferior or merely technical skills [38]. In addition, other studies have shown that relational interventions generate a wide variety of positive health outcomes. Research has demonstrated that effective therapeutic relationships can enhance treatment adherence and rehabilitation and reduce perceptions of stigma, as outlined in Parkhideh et al [39]. As demonstrated in the literature and as evidenced by our own findings, there is clear evidence of the importance that patients attach to being treated with respect and as people rather than as problems or diseases [40].

Among the recommendations expressed by users, the one that stands out is the suggestion that health care personnel show greater patience during care. This demand reflects a need for more humane and empathetic treatment and highlights the importance of the therapeutic relationship component. Patience is seen as a sign of understanding, respect, dignity, and a willingness to listen, which are all basic elements in building an effective working relationship. This is particularly relevant in contexts involving vulnerable populations, such as individuals with SUDs. Integrating this perspective into clinical practice has the potential to enhance the quality of care and ensure adherence to treatment [35,38].

However, there is a lack of knowledge regarding the roles performed by each member of the health care team. Although the role of the nurse remains constant within the team, participants confuse their role with those of doctors and social workers during the discussions. This is also evident from the absence of any codes or categories relating to other areas of nursing care. One potential explanation for these findings is that patients may not have had sufficient time to comprehend the nature of nursing work. However, it should be noted that all participants had previously been admitted to at least 1 other rehabilitation program. Therefore, the lack of knowledge about nursing work may not be attributable to a lack of contact with nursing staff, but rather to the invisibility of care. The extant literature supports this assertion. In both this population and the general population, care is often not directly associated with nursing. This may mean that the specific and specialized role of nurses in various contexts is not fully recognized. This lack of knowledge about nursing care reflects an underlying problem, which contributes to its social and professional devaluation [41,42].

The participants' accounts also allow us to infer a structural reality that affects the professional practice of nursing: the low institutional and social value placed on this group. This situation is evident in working conditions that often fail to reflect the complexity of the care required by patients with SUDs. These conditions include a lack of financial recognition, excessive workloads, a shortage of adequate rest areas for staff, and limited participation by nurses in decision-making spaces. This invisibility stands in stark contrast to the pivotal role that patients themselves ascribe to nurses in their recovery processes, particularly regarding human support, active listening, and the establishment of long-term therapeutic bonds [43]. This structural invisibility might help explain why some participants were uncertain about the specific responsibilities of each professional, occasionally mixing up nursing functions with those of other team members. This confusion may also arise from overlapping interdisciplinary tasks, inconsistent role introductions—particularly at admission or during care-plan changes—and contextual factors in SUD care, such as time pressure, staff turnover, and psychosocial vulnerability. Therefore, strengthening role communication and increasing the visibility of nursing responsibilities might enhance patients' understanding and recognition of nursing care within community addiction services [41]. Role ambiguity may also limit patients' ability to actively seek nursing support. Brief standardized introductions at admission, visible role identification, and, when feasible, assigning a reference nurse may improve clarity. Future research should further explore patients' perspectives on strategies to enhance understanding of nursing roles in care [41,44].

Limitations

This study is one of a limited number of research projects conducted in Spain that address the perceptions of a particularly challenging-to-reach demographic, such as patients with SUDs. However, it is vital to understand the experiences of individuals with SUDs and their interactions with nursing care in order to improve and personalize interventions and establish effective improvement strategies within this particularly vulnerable population. Nevertheless, this study presents some limitations. Most of the participants in the sample are male, and most of them self-identify as such. This finding is consistent with the available evidence, which points to a higher prevalence of substance use among men [2]. From a gender perspective, this pattern may be influenced by social constructs associated with masculinity. These constructs include the normalization of risk, pressure to demonstrate strength or autonomy, and perceived lower access to emotional support networks [4]. These factors may facilitate the initiation and maintenance of use and therefore bias the findings. Conversely, the participants were members of merely two centers belonging to the same rehabilitation Red Cross program, which may have exerted an influence on the results due to the absence of variety in the care received. Furthermore, no distinction was made between different types of nursing care (specialized, hospital-based, or based on the years of experience of the professionals). To expand the knowledge base regarding this population's perceptions of the nursing role, future studies involving samples of patients from diverse rehabilitation programs and incorporating a gender perspective are essential. Women's perspectives may be insufficiently captured in this study, and future research should purposively recruit more women (and gender-diverse participants when possible) to better inform gender-responsive nursing care in community addiction services.

XSL•FO
RenderX

Conclusions

This study demonstrated that SUD is a chronic condition, with multiple admissions to rehabilitation programs. Participants have provided a positive evaluation of the rehabilitation center, emphasizing the comprehensive support provided by the multidisciplinary team and the safe environment it offers, beyond clinical treatment. Group activities and ongoing support are key factors in maintaining abstinence and promoting social integration.

Patients with SUDs place a high value on the role of nursing, particularly relational interventions over technical interventions. However, this assessment contrasts with a lack of knowledge about certain functions of the nursing profession, which limits their ability to recognize and take advantage of their capacity for action in community settings.

These findings emphasize the necessity to enhance care models through a holistic approach, considering patients' life contexts and recognizing both individual care and structural conditions. Nurses can play an essential role in designing patient-centered interventions to achieve sustainable recovery and a positive therapeutic experience.

Acknowledgments

We would like to express our gratitude to the staff of Red Cross centers in Madrid for their crucial assistance in reaching out to patients.

Artificial intelligence (Microsoft Copilot) was used only to help draft the abstract and did not replace the authors' critical analysis or interpretation.

Funding

The authors declared no financial support was received for this work.

Data Availability

No additional data are available due to data protection requirements.

Authors' Contributions

CLG, MGdQC, and BGT led the design of the study, its analysis, and the preparation of the study. BGT contributed to the data analysis, identification of codes, and categories and selection of relevant quotes. CLG, MGdQC, BGT, NMG, and GGN have contributed to preparing the study. All authors have read and approved the final study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sociodemographic data, participant information, and interview guide.
[DOCX File, 16 KB - nursing_v9i1e82401_app1.docx]

Multimedia Appendix 2
Most used substances among interviewed participants enrolled in the rehabilitation program.
[PNG File, 75 KB - nursing_v9i1e82401_app2.png]

References

1. American Psychiatric Association. Trastornos relacionados con sustancias y trastornos adictivos [Book in Spanish]. In: Manual Diagnóstico y Estadístico de Los Trastornos Mentales (DSM-5), 5th edition: Editorial Médica Panamericana; 2014:481-590 URL: https://www.federaciocatalanatdah.org/wp-content/uploads/2018/12/dsm5-manualdiagnsticoyestadisticodelostrastornosmentales-161006005112.pdf [accessed 2026-03-13]
2. Rehm J, Gmel Sr GE Sr, Gmel G, et al. The relationship between different dimensions of alcohol use and the burden of disease-an update. Addiction 2017 Jun;112(6):968-1001. [doi: 10.1111/add.13757] [Medline: 28220587]
3. Forouzanfar MH, Afshin A, Alexander LT, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet 2016 Oct;388(10053):1659-1724. [doi: 10.1016/S0140-6736(16)31679-8] [Medline: 27733284]
4. Ventura CAA, Carrara BS, Fernandes RHH, et al. General beliefs about illicit drug use and stigma: perspective of people who use illicit drugs. Community Ment Health J 2022 Oct;58(7):1346-1353. [doi: 10.1007/s10597-022-00944-8] [Medline: 35182249]

XSL•FO
RenderX

5. Nelson EUE, Essien NF. Redefining recovery: accounts of treatment experiences of dependent cannabis users in Nigeria. J Subst Abuse Treat 2021 Jun;125:108321. [doi: 10.1016/j.jsat.2021.108321] [Medline: 34016304]
6. Simpson KA, Bolshakova M, Kirkpatrick MG, et al. Characterizing opioid withdrawal experiences and consequences among a community sample of people who use opioids. Subst Use Misuse 2024;59(6):886-894. [doi: 10.1080/10826084.2024.2306221] [Medline: 38287506]
7. European Drug Report 2025: trends and developments. : European Union Drugs Agency (EUDA); 2025 URL: https://www.ektepn.gr/sites/default/files/2025-06/edr-2025-full-book-5.06.2025-en-v3_1.pdf [accessed 2026-03-13]
8. Ledingham E, Adams RS, Heaphy D, Duarte A, Reif S. Perspectives of adults with disabilities and opioid misuse: qualitative findings illuminating experiences with stigma and substance use treatment. Disabil Health J 2022 Jun;15(2S):101292. [doi: 10.1016/j.dhjo.2022.101292] [Medline: 35341717]
9. Towers EB, Williams IL, Qillawala EI, Rissman EF, Lynch WJ. Sex/gender differences in the time-course for the development of substance use disorder: a focus on the telescoping effect. Pharmacol Rev 2023 Mar;75(2):217-249. [doi: 10.1124/pharmrev.121.000361] [Medline: 36781217]
10. Cabrera A G. El modelo transteórico del comportamiento en salud [Article in Spanish]. Rev Fac Nac Salud Pública 2000;18(2):129-138 [FREE Full text] [doi: 10.17533/udea.rfnsp.997]
11. Minozzi S, Amato L, Jahanfar S, Bellisario C, Ferri M, Davoli M. Maintenance agonist treatments for opiate-dependent pregnant women. Cochrane Database Syst Rev 2020 Nov 9;11(11):CD006318. [doi: 10.1002/14651858.CD006318.pub4] [Medline: 33165953]
12. Arman M, Ranheim A, Rydenlund K, Rytterström P, Rehnsfeldt A. The Nordic tradition of caring science: the works of three theorists. Nurs Sci Q 2015 Oct;28(4):288-296. [doi: 10.1177/0894318415599220] [Medline: 26396212]
13. Feo R, Kitson A, Conroy T. How fundamental aspects of nursing care are defined in the literature: a scoping review. J Clin Nurs 2018 Jun;27(11-12):2189-2229. [doi: 10.1111/jocn.14313] [Medline: 29514402]
14. Solheim K, Hegg Reime M, Eide LSP. How do persons who inject drugs experience care from nurses in hospital settings? A qualitative study. Glob Qual Nurs Res 2024;11:23333936241240795. [doi: 10.1177/23333936241240795] [Medline: 38577493]
15. Moyo P, Nishar S, Merrick C, et al. Perspectives on admissions and care for residents with opioid use disorder in skilled nursing facilities. JAMA Netw Open 2024 Feb 5;7(2):e2354746. [doi: 10.1001/jamanetworkopen.2023.54746] [Medline: 38315484]
16. Phoenix BJ, Tierney M, Chapman SA, Spetz J. Influences on advance practice nursing education to prescribe medications for opioid use disorder. Nurs Outlook 2023;71(3):101963. [doi: 10.1016/j.outlook.2023.101963] [Medline: 37003088]
17. Ellefsen R, De Pirro S, Haukland V, Wüsthoff LEC, Arnevik EA. “It’s a delicate balance”: clinicians’ experiences of providing heroin-assisted treatment. Harm Reduct J 2024 Dec 31;21(1):230. [doi: 10.1186/s12954-024-01135-2] [Medline: 39741328]
18. Merriam SB, Tisdell EJ. Qualitative Research: A Guide to Design and Implementation: Jossey-Bass Inc Pub; 2016. URL: https://www.amazon.in/Qualitative-Research-Guide-Design-Implementation/dp/111900361X [accessed 2026-03-13]
19. Creswell JW. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches: Sage Publications; 2014. URL: https://collegepublishing.sagepub.com/products/research-design-5-255675 [accessed 2026-03-20]
20. Lichtman M. Understanding and Evaluating Qualitative Educational Research: SAGE Publications Inc; 2011. URL: https://sk.sagepub.com/book/edvol/understanding-and-evaluating-qualitative-educational-research/toc#_ [accessed 2026-03-13]
21. Marshall C, Rossman GB. Designing Qualitative Research, 5th edition: Sage Publications; 2011. URL: https://books.google.co.in/books?id=RbqXGjKHALoC&printsec=frontcover&source=gbs_vpt_reviews#v=onepage&q&f=false [accessed 2026-03-13]
22. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
23. Strauss AL, Corbin J. Bases de La Investigación Cualitativa: Técnicas y Procedimientos Para Desarrollar La Teoría Fundamentada [Book in Spanish]: Editorial Universidad de Antioquia; 2002. URL: https://openlibrary.org/books/OL26228230M/Bases_de_la_investigaci%C3%B3n_cualitativa_t%C3%A9cnicas_y_procedimientos_para_desarrollar_la_teor%C3%ADa_fundament [accessed 2026-03-13]
24. Guba EG. Criterios de credibilidad en la investigación naturalista [Book in Spanish]. In: La Enseñanza: Su Teoría y Su Práctica: Ediciones Akal; 1989:148-165 URL: https://dialnet.unirioja.es/servlet/articulo?codigo=6220913 [accessed 2026-03-13]
25. Malterud K. Systematic text condensation: a strategy for qualitative analysis. Scand J Public Health 2012 Dec;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
26. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16. [doi: 10.1177/1609406917733847]
27. Fossey E, Harvey C, McDermott F, Davidson L. Understanding and evaluating qualitative research. Aust N Z J Psychiatry 2002 Dec;36(6):717-732. [doi: 10.1046/j.1440-1614.2002.01100.x] [Medline: 12406114]

XSL•FO
RenderX

28. Engelgardt P, Krzyżanowski M, Borkowska-Sztachańska M, Wasilewska A, Ciucias M. The impact of lifetime substance use on psychiatric comorbidities and treatment seeking in patients with alcohol use disorders. Sci Rep 2024 Jun 20;14(1):14257. [doi: 10.1038/s41598-024-65028-x] [Medline: 38902395]
29. Schoenberger SF, Park TW, dellaBitta V, Hadland SE, Bagley SM. “My life isn’t defined by substance use”: recovery perspectives among young adults with substance use disorder. J Gen Intern Med 2022 Mar;37(4):816-822. [doi: 10.1007/s11606-021-06934-y] [Medline: 34100229]
30. Furulund E, Carlsen SEL, Druckrey-Fiskaaen KT, Madebo T, Fadnes LT, Lid TG. A qualitative study of experiences with physical activity among people receiving opioid agonist therapy. Subst Abuse Treat Prev Policy 2024 May 6;19(1):26. [doi: 10.1186/s13011-024-00607-9] [Medline: 38711108]
31. Simonton AJ, Young CC, García AA, Bartholomew JB, Brown RA. A cross-sectional study of physical activity attitudes and preferences of individuals with opioid use disorder. Ment Health Phys Act 2022 Mar;22:100444. [doi: 10.1016/j.mhpa.2022.100444]
32. Schlossberg NK. A model for analyzing human adaptation to transition. Couns Psychol 1981 Jun;9(2):2-18. [doi: 10.1177/001100008100900202]
33. Anderson ML, Goodman J, Schlossberg NK. Counseling Adults in Transition: Linking Schlossberg’s Theory With Practice in a Diverse World: Springer; 2012. URL: https://scholarworks.wmich.edu/books/15/ [accessed 2026-03-13]
34. Elswick A, Fallin-Bennett A. Voices of hope: a feasibility study of telephone recovery support. Addict Behav 2020 Mar;102:106182. [doi: 10.1016/j.addbeh.2019.106182] [Medline: 31778847]
35. Ould Brahim L, Hanganu C, Gros CP. Understanding helpful nursing care from the perspective of mental health inpatients with a dual diagnosis: a qualitative descriptive study. J Am Psychiatr Nurses Assoc 2020;26(3):250-261. [doi: 10.1177/1078390319878773] [Medline: 31587601]
36. Torres VN, Williams EC, Ceballos RM, Donovan DM, Ornelas IJ. Discrimination, acculturative stress, alcohol use and their associations with alcohol-related consequences among Latino immigrant men. J Ethn Subst Abuse 2024;23(1):150-165. [doi: 10.1080/15332640.2022.2077273] [Medline: 35634786]
37. Leza L, Haro B, López-Goñi JJ, Fernández-Montalvo J. Substance use disorder and lifetime suicidal behaviour: a scoping review. Psychiatry Res 2024 Apr;334:115830. [doi: 10.1016/j.psychres.2024.115830] [Medline: 38432115]
38. DeFrino DT. A theory of the relational work of nurses. Res Theory Nurs Pract 2009;23(4):294-311. [doi: 10.1891/1541-6577.23.4.294] [Medline: 19999747]
39. Parkhideh A, Beiting KJ, Yang M, Landi AJ, Levine S. Substance use disorder care in skilled nursing facilities: characterizing resident experiences. J Addict Med 2024;18(5):533-539. [doi: 10.1097/ADM.0000000000001318] [Medline: 38842178]
40. Moreno-Poyato AR, Montesó-Curto P, Delgado-Hito P, et al. The therapeutic relationship in inpatient psychiatric care: a narrative review of the perspective of nurses and patients. Arch Psychiatr Nurs 2016 Dec;30(6):782-787. [doi: 10.1016/j.apnu.2016.03.001] [Medline: 27888975]
41. Mourão Netto JJ. Invisibility and devaluation of nursing work: related factors and coping strategies. Invest Educ Enferm 2024 Mar;42(1):e01. [doi: 10.17533/udea.iee.v42n1e01] [Medline: 39083813]
42. Lund SB, Skolbekken JA, Mosqueda L, Malmedal W. Making neglect invisible: a qualitative study among nursing home staff in Norway. Healthcare (Basel) 2023 May 12;11(10):1415. [doi: 10.3390/healthcare11101415] [Medline: 37239698]
43. Nunes JC, Costa GPA, Weleff J, Rogan M, Compton P, De Aquino JP. Assessing pain in persons with opioid use disorder: approaches, techniques and special considerations. Br J Clin Pharmacol 2024 Dec;90(12):2985-3002. [doi: 10.1111/bcp.16055] [Medline: 38556851]
44. Roodbeen RTJ, Bruinsma J, Rozema AD, Crutzen R, Stutterheim SE. Accessibility, clarity, and organizational opportunities to enhance interprofessional collaboration in alcohol interventions: a qualitative study. BMC Health Serv Res 2025 Oct 21;25(1):1384. [doi: 10.1186/s12913-025-13552-5] [Medline: 41121102]

Abbreviations

CATC: community addiction treatment center
***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
SUD: substance use disorder

Edited by E Borycki; submitted 14.Aug.2025; peer-reviewed by MR López, RG Hervías; accepted 24.Feb.2026; published 31.Mar.2026.

Please cite as:
Lucas Guerra C, Gómez de Quero Córdoba M, García Tardón B, González Naranjo G, Mayoral Gonzalo N
Patients’ Perceptions of the Role of Nursing in Substance Use Disorder Treatment Programs: Qualitative Study
JMIR Nursing 2026;9:e82401
URL: https://nursing.jmir.org/2026/1/e82401
doi:10.2196/82401

XSL•FO
RenderX

© Clara Lucas Guerra, Marina Gómez de Quero Córdoba, Bruno García Tardón, Guadalupe González Naranjo, Noemí Mayoral Gonzalo. Originally published in JMIR Nursing (https://nursing.jmir.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multicomponent Supportive-Educative mHealth Intervention for Self-Care Among Individuals With Osteoporosis in Thailand: Exploratory Pretest-Posttest Pilot Study

Natthida Pangutha[1], RN, MNS; Sirinapha Jittimanee[2], RN, MSN, PhD; Porramat Saksaen[3], RN, MNS

[1]Faculty of Nursing, Chulalongkorn University, Pathumwan, Bangkok, Thailand

[2]Center of Excellence for Enhancing Well-being in Vulnerable and Chronic Illness Populations, Faculty of Nursing, Chulalongkorn University, 254 Rama 1 Road, Wangmai Subdistrict, Pathumwan District, Bangkok, Thailand

[3]Chaiprakan Hospital, Ministry of Public Health, Chaiprakan district, Chiang Mai province, Thailand

Corresponding Author:
Sirinapha Jittimanee, RN, MSN, PhD
Center of Excellence for Enhancing Well-being in Vulnerable and Chronic Illness Populations, Faculty of Nursing, Chulalongkorn University, 254 Rama 1 Road, Wangmai Subdistrict, Pathumwan District, Bangkok, Thailand

Abstract

Background: Poor self-care impairs osteoporosis management. Disease-related knowledge and family support have influenced self-care; however, the combined association of these factors with self-care remains insufficiently understood.

Objective: This study aimed to compare self-care among individuals with osteoporosis who received a multicomponent supportive-educative mobile health intervention with that of individuals who received usual care.

Methods: This exploratory pilot study using a pretest-posttest design with a nonequivalent control group included 46 individuals aged 50 years and older who had been diagnosed with osteoporosis and were receiving care at the orthopedic outpatient department of a university hospital in Bangkok. Participants were matched by sex, age, and back pain scores and then sequentially allocated to an intervention group (n=23, 50%) and a comparison group (n=23, 50%). The intervention group received a 4-week multicomponent supportive-educative intervention based on Orem's self-care theory, delivered partly through a closed Facebook group. The intervention included educational guidance, video-based teaching, psychological support, family involvement, and remote nurse interaction. The comparison group received usual care. Self-care was measured using the Self-Care of Osteoporosis Scale. Chi-square tests, Fisher exact tests, independent-sample 2-tailed *t* tests, analysis of covariance, and multivariate analysis of covariance were used to examine differences between groups.

Results: After controlling for baseline self-care scores, participants in the intervention group had significantly higher postintervention self-care scores than those in the comparison group ($F_{1,43}$=459.29; P<.001). Significant group differences were also observed across all 3 domains of self-care after controlling for baseline scores. Compared with the control group, the intervention group demonstrated significantly higher postintervention scores in diet and physical activity ($F_{1,41}$=245.04; P<.001), safety performance ($F_{1,41}$=171.20; P<.001), and medication adherence ($F_{1,41}$=144.54; P<.001).

Conclusions: The multicomponent supportive-educative intervention delivered partly through Facebook was associated with improvements in self-care among individuals with osteoporosis. Future rigorous trials are needed to confirm its efficacy.

(JMIR Nursing 2026;9:e101025) doi:10.2196/101025

KEYWORDS

osteoporosis; self-care; telemedicine; family support; exercise; accident prevention

Introduction

Osteoporosis affects 23.1% of older women and 11.7% of older men worldwide [1]. It is associated with increased risk of osteoporotic fractures, particularly at the femoral neck (31.9%), spine (28.8%), and hip (9.9%) [2]. The causes of osteoporosis include reduced estrogen levels after menopause, inadequate intake of vitamin D and calcium, the use of certain medications (eg, anticoagulants), and certain diseases (metabolic and hormonal disorders) [3]. Effective management of osteoporosis relies on both medical treatment and individuals' engagement in self-care, including medication adherence, regular exercise, adequate dietary intake, and fall prevention, to reduce the risk of complications.

Self-care among individuals with osteoporosis remains a significant challenge. One study reported that 100% of women with osteoporosis did not engage in exercise, whereas mean daily calcium and vitamin D intake was 546.7 (SD 60.28) mg

XSL•FO
RenderX

and 2.0 (SD 0.34) µg, respectively [4]. Both values were below the recommended daily intake of 1000 to 1200 mg for calcium and 15 to 20 µg for vitamin D [5]. These behaviors include discontinuing medication without medical consultation [6]. Self-care for osteoporosis refers to behaviors undertaken to manage the condition and prevent complications. Nurses play a vital role in supporting and encouraging individuals with osteoporosis to adopt these self-care practices [7].

Past research has identified several factors influencing self-care in individuals with osteoporosis, including disease-related knowledge, social support, patient–health care provider relationship, medication-related side effects, and patient participation in treatment. In one study, poor disease-related knowledge was reported in 52.8% of patients and was associated with inadequate intake of vitamin D and calcium [8]. Social support and a strong patient-physician relationship have been positively associated with self-care behaviors [9]. Medication-related side effects, including those associated with self-administered Forteo, also influence medication self-care [6]. In addition, patient participation in treatment decisions has been linked to improved treatment outcomes [10].

Interventions to promote self-care among individuals with osteoporosis have been widely studied. These include educational video-based programs, group training sessions, telephone follow-ups, and support from health care teams. There is evidence suggesting that such interventions significantly improve medication adherence and reduce the risk of fractures [11,12]. However, most of the previous interventions focus primarily on individuals (participants engaged in the interventions themselves without the involvement of family members) even though many are older adults who may require assistance from family members and may find it difficult to follow instructions independently. Integrating family support with the supportive-educative nursing system by Orem et al [13], which emphasizes teaching, guidance, emotional support, and the creation of an environment conducive to self-care, may better address these needs and enhance self-care among individuals with osteoporosis.

In Thailand, 1 in 5 women aged 40 to 80 years has osteoporosis, and the prevalence exceeds 50% among those aged 70 years and above [14]. However, challenges similar to those in other countries persist; only 67.2% of individuals have adequate calcium intake, and overall levels of physical activity remain low [15]. These problems highlight the need for interventions to promote self-care among individuals with osteoporosis. With the rapid advancement of mobile health (mHealth) technologies and the widespread use of social media, these platforms offer promising opportunities for delivering supportive-educative interventions and promoting healthy behaviors, as reported in previous studies [16]. Family support also plays a significant role in enhancing self-care, as shown in previous studies [17]. Addressing this gap is crucial for enhancing self-care among individuals with osteoporosis. Therefore, this study aimed to compare self-care between individuals with osteoporosis who received a supportive-educative mHealth intervention and those who received usual care. We hypothesized that participants receiving the multicomponent supportive-educative mHealth intervention would demonstrate significantly higher overall self-care mean scores and higher individual mean scores in the 3 domains of self-care (medication adherence, diet and physical activity, and safety performance) than those receiving usual care.

Methods

Study Design

This exploratory pilot study used a pretest-posttest design with a nonequivalent control group. The study followed the TREND (Transparent Reporting of Evaluations With Nonrandomized Designs) checklist [18].

Study Setting

This study was conducted at an outpatient clinic in a 1200-bed university hospital in Bangkok. In 2025, a total of 256 individuals with osteoporosis received care at this hospital, where osteoporosis services are provided at the orthopedic outpatient department every Wednesday. Medical appointments are scheduled every 4 to 6 weeks.

Participants

Eligibility Criteria

Participants were individuals diagnosed with osteoporosis. Eligibility criteria were (1) age of 50 years or older, (2) receipt of treatment for 6 months or more, (3) residence with family members or relatives, and (4) either use of Facebook or having a family member who used Facebook. Individuals were excluded if they were severely ill; were scheduled for surgery during the intervention period; had speech or hearing impairments; or had cognitive impairment, defined as a score of 22 or lower on the Thai version of the Mini-Mental State Examination.

The sample size was calculated using an effect size of 0.8, a significance level of .05, a power of 0.80, and a *t* test, resulting in a required sample size of 21 individuals per group. Allowing for a 10% dropout rate, the final sample size was 23 participants per group.

Method of Recruitment

Eligible participants were consecutively identified by nurses at the time of clinic check-in. Nurses provided a flyer to invite them to participate in the study. Interested individuals were referred to the researcher, who explained the study procedures. During the enrollment period between July 2025 and August 2025, a total of 120 individuals with osteoporosis attended the clinic, of whom 88 (73.3%) were eligible and invited, and 46 (52.3%) agreed to participate. All individuals who declined participation were female. Participants were matched 1:1 based on sex, age, and back pain score because these variables influenced self-care [6,9].

Assignment Method

Participants were assigned alternately to the intervention and comparison groups in a 1:1 ratio. This assignment method was used to maintain comparable group sizes and facilitate group matching based on key participant characteristics.

XSL•FO
RenderX

Blinding

Due to the nature of the multicomponent supportive-educative mHealth intervention, blinding of participants and researchers was not feasible. Participants were aware of their group assignment, and the researchers delivering the intervention were not blinded.

Usual Care

During the study period, the comparison group continued to receive usual care. Nurses provided health education focusing on medication, diet, and exercise. This education was delivered orally, without the use of written materials or video clips for exercise demonstration. Individuals could confirm or reschedule their appointments through the hospital application.

Intervention Program

Intervention Description

The multicomponent intervention was delivered through a private Facebook group over a 4-week period (28 days). It consisted of 4 weekly 15-minute educational sessions conducted in small groups of 3 to 4 participants. Live Facebook sessions were held daily between 3 PM and 4 PM to accommodate participants' schedules. Each participant and 1 family member attended 1 session per week using a shared study-specific Facebook account. Participants who missed their scheduled session received a reminder from the researcher and attended a subsequent session. Outcome assessments of self-care were conducted at 2 time points: baseline and 1 week after completion of the intervention during face-to-face visits at the outpatient department.

The multicomponent intervention focused on three domains of self-care: (1) medication adherence, (2) diet and physical activity, and (3) safety performance. It was developed based on Orem's self-care theory [13], specifically the supportive-educative nursing system, and incorporated the helping methods by Orem et al [13] by (1) guiding and directing participants to strengthen self-care ability and decision-making; (2) teaching through four 10-minute educational video clips covering osteoporosis and medication, nutrition, exercise, and fall prevention; (3) providing psychological support through online family reflection sessions; and (4) creating a supportive learning environment through regular nurse-participant interactions (Multimedia Appendix 1). A printed handbook containing the same educational content as the 4 video clips was provided to reinforce learning throughout the intervention. An overview of the intervention is shown in Table 1.

XSL•FO
RenderX

Table . Intervention overview.

	Objectives	Key activities	Delivery method	Helping methods by Orem et al [13]	Monitoring and evaluation
Session 1 (baseline)	Establish relationships with individuals and families	• Introduce the intervention • Help participants and families share self-care experiences • Assess baseline self-care and back pain scores using an interviewer-administered questionnaire • Provide self-care handbook	Face-to-face	Guiding and directing	Baseline self-care
Session 2 (week 1)	Explain disease management	• Learn about osteoporosis, medication, nutrition, and exercise • Upload four 10-min video clips, which are available for repeated viewing • Participate in questions and answers	Online group session via closed Facebook group	Teaching	Ability to identify content from the video and handbook
Session 3 (week 2)	Reflect on self-care experience	• Facilitate the sharing of experience reflection sessions by participants and families applying self-care • Facilitate group discussion and questions and answers	Online group session via closed Facebook group	Supporting	Self-report on activities
Session 4 (week 3)	Apply strategies to improve home safety	• Shared experiences • Review environmental safety and accident prevention through video • Facilitate group discussion	Online group session via closed Facebook group	Supporting	Self-report of home changes
Session 5 (week 4)	Use available resources to sustain self-care	• Discuss barriers and apply problem-solving strategies • Engage family support • Reflect on resource use	Online group session via closed Facebook group	Providing an environment that promotes personal development	Family attendance

Family Involvement

Family members participated in all intervention sessions delivered through the Facebook group. They received the same educational content as participants and were encouraged to support self-care behaviors, including medication adherence, dietary management, physical activity, and home safety practices. During the sessions, family members were invited to share their experiences and discuss strategies they used to support participants in managing osteoporosis.

XSL•FO
RenderX

Facebook Group Privacy Protection

To protect their privacy, participants were asked to create new Facebook accounts exclusively for this study. Participants were instructed to use pseudonyms rather than their real names. Each participant and 1 family member shared the same study-specific account, resulting in 23 Facebook accounts. All live Facebook group sessions were conducted with participants' cameras turned off, and only pseudonyms were displayed during the sessions to maintain confidentiality.

Intervention Fidelity

Intervention fidelity was monitored throughout the study to assess intervention delivery and participant engagement. All intervention sessions were delivered by the principal investigator according to a standardized intervention protocol. Four educational videos covering osteoporosis and medication, nutrition, exercise, and fall prevention were uploaded to the private Facebook group. Facebook analytics recorded 71 views for the osteoporosis and medication video, 165 views for the exercise video, 73 views for the nutrition video, and 90 views for the accident prevention video. Because Facebook did not permit tracking of individual viewing histories or downloads, participant engagement was also assessed through attendance to the weekly live Facebook sessions. All participants attended 1 live session per week, resulting in a 100% (23/23) attendance rate across the 4 scheduled sessions for each participant.

Participants and their family members were asked open-ended questions about the educational content, video viewing experience, attendance to the online sessions, and implementation of self-care activities. Participants were able to accurately describe the educational content and discuss its application to self-care, indicating that all participants had viewed the videos. Any misunderstandings were clarified immediately by the principal investigator during the discussions. All 23 Facebook study accounts completed the intervention, and 3 (13%) participants who missed their originally scheduled sessions attended an alternative session after receiving a reminder from the researcher.

Outcome

The primary outcome was self-care, which was measured using the Self-Care of Osteoporosis Scale (SCOS) [7], a self-reported instrument.

Instruments

Overview

The SCOS, developed by Cittadini et al [7], consists of 15 items across 3 domains: medication adherence, diet and exercise, and safety to prevent accidents. The scale uses a 5-point Likert scale ranging from "never" (1 point) to "always" (5 points). The total SCOS score was calculated as the mean of the 15 item scores, yielding a possible score ranging from 1 to 5, with higher mean scores indicating better self-care behaviors. The original instrument demonstrated excellent internal consistency, with Cronbach α values of 0.94 for the overall scale, 0.84 for medication adherence, 0.88 for diet and physical activity, and 0.85 for safety performance. As the SCOS had not previously been used in Thailand, permission was obtained from the original instrument developer, and the instrument was translated using a standard forward-backward translation process followed by expert review. In the present study, the Thai version demonstrated acceptable internal consistency at baseline, with a Cronbach α value of 0.93 for the overall scale, 0.56 for medication adherence, 0.80 for diet and physical activity, and 0.70 for safety performance. At the postintervention time point, Cronbach α values were 0.95 for the overall scale, 0.79 for medication adherence, 0.91 for diet and physical activity, and 0.89 for safety performance.

Back pain was assessed using the visual analog scale, a 10-cm line representing pain intensity from 0 to 10, with each centimeter corresponding to 1 point [19].

A process evaluation form was used to assess participants' engagement in reflective discussions after viewing the videos. Four open-ended questions were used to explore participants' experiences: (1) weekly viewing experience, (2) most preferred content and reasons, (3) least preferred content and reasons, and (4) suggestions for improving content or presentation.

Validity Testing and Tryout

The content validity of the intervention and instruments was evaluated by 5 experts in osteoporosis and physical therapy. The content validity index was 0.86, indicating good content validity. The instruments were also pilot-tested with 30 individuals with osteoporosis who met the inclusion criteria but were not enrolled in the study to assess the clarity and readability of the instruments.

Unit of Analysis

Although family members participated in the intervention sessions, outcome measures were collected only from participants with osteoporosis. Therefore, the individual participant was the unit of analysis.

Statistical Analysis

Data were analyzed using SPSS (version 29.0; IBM Corp). Normality was assessed using the Kolmogorov-Smirnov test, and all variables were normally distributed; therefore, parametric tests were applied. Baseline homogeneity between groups was examined using the chi-square test, Fisher exact test, and independent-sample *t* test as appropriate. Differences in overall self-care following the intervention were analyzed using analysis of covariance (ANCOVA), whereas differences in the 3 domains of self-care were examined using multivariate ANCOVA (MANCOVA). Statistical significance was set at $P<.05$. To facilitate transparency and reproducibility, statistical outputs from SPSS, including baseline comparisons and ANCOVA and MANCOVA results, are provided in Multimedia Appendix 2.

Prior to the main analyses, the assumptions underlying ANCOVA and MANCOVA were evaluated. Normality of residuals was assessed using the Shapiro-Wilk and Kolmogorov-Smirnov tests. The Shapiro-Wilk test indicated a slight deviation from normality ($W=0.946$; $P=.03$), whereas the Kolmogorov-Smirnov test indicated no significant departure from normality ($P=.07$). Although the Shapiro-Wilk test suggested a deviation from normality, ANCOVA and MANCOVA are considered robust to moderate violations of

XSL•FO
RenderX

normality when group sizes are equal (n=23 per group); therefore, parametric analyses were retained. Linearity between the covariate (baseline self-care score) and the postintervention outcome was confirmed ($F_{1,43}$=113.914; P<.001). The assumption of homogeneity of regression slopes was evaluated by including a group × baseline self-care mean score interaction term in the ANCOVA model and was satisfied as the interaction was not statistically significant ($F_{1,42}$=3.626; P=.06). The Levene test indicated that the assumption of homogeneity of variance was satisfied for the overall self-care score ($F_{1,44}$=2.989; P=.09).

For MANCOVA, the Bartlett test of sphericity confirmed significant intercorrelations among the 3 self-care domains, supporting the use of a multivariate approach (χ^2_3=47.3; P<.001). Pearson correlation coefficients among the 3 postintervention domain scores ranged from 0.739 to 0.861 (P<.001 in all cases), indicating no evidence of multicollinearity (r<0.90). Homogeneity of variance-covariance matrices, assessed using the Box M test, was violated (M=21.231; F_6=3.276; P=.003). Therefore, the Pillai trace was used to evaluate multivariate group differences because it is robust to violations of the assumption of homogeneity of covariance matrices. However, the baseline mean score for the diet and physical activity domain differed significantly between the groups (P=.04), whereas the baseline mean scores for the medication adherence and safety performance domains did not. To account for this baseline difference, the corresponding baseline SCOS domain mean scores were included as covariates in the MANCOVA model. Therefore, the postintervention group comparisons were based on adjusted mean scores after controlling for the corresponding baseline SCOS domain mean scores, minimizing the potential influence of baseline imbalance on the interpretation of the between-group differences.

Ethical Considerations

Ethics approval was obtained from Chulalongkorn University (022/68) and the Faculty of Medicine of Ramathibodi Hospital (MURA2025/477). All participants provided written informed consent prior to participation. To ensure confidentiality, all data were stored in password-protected files accessible only to the research team. Participants in the intervention group received a stipend of US $8 to offset internet-related expenses associated with participation in the intervention.

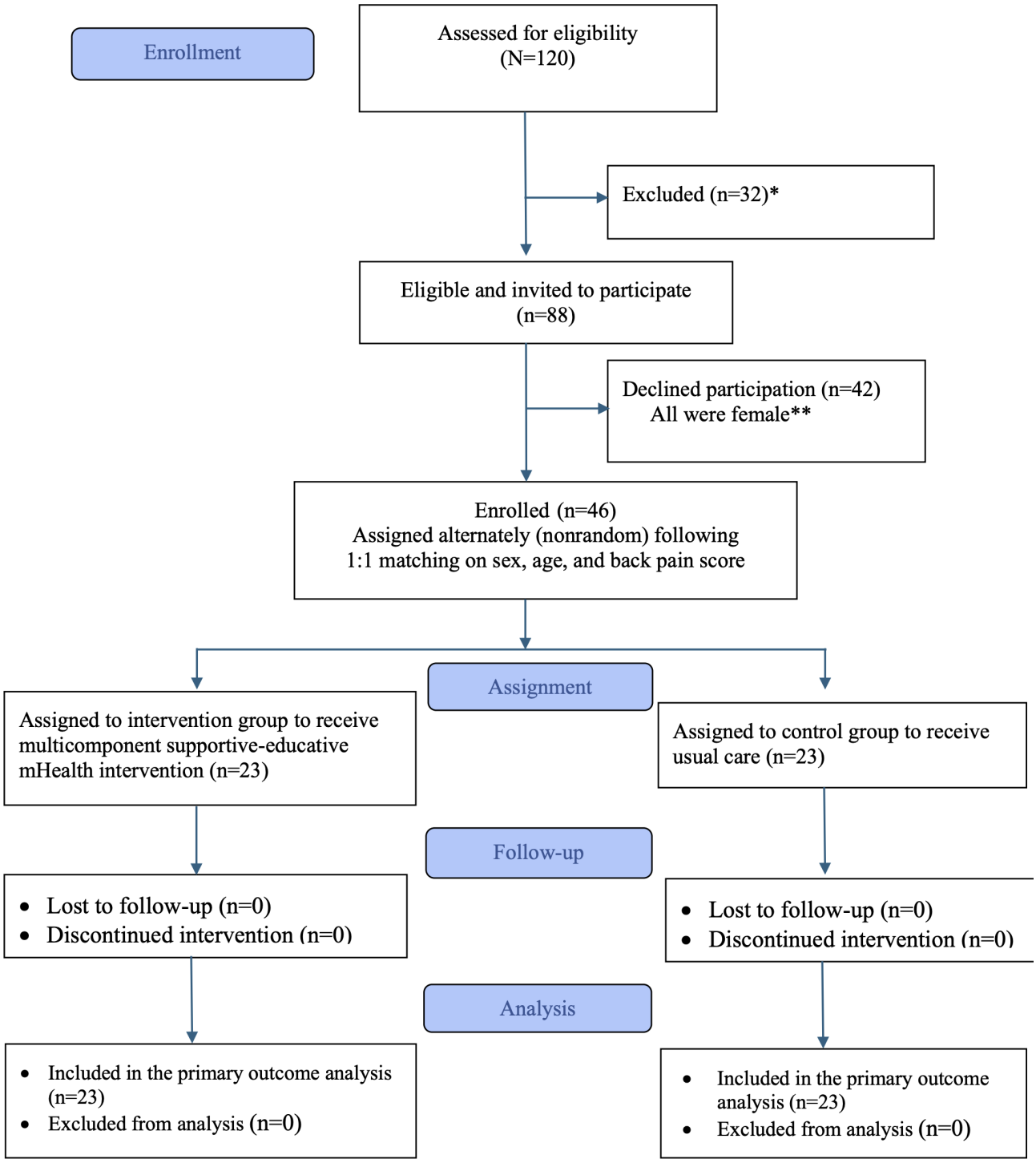
Results

Homogeneity Test for Participant Characteristics

A total of 46 participants were included in the study, with 23 (50%) in the intervention group and 23 (50%) in the comparison group (Figure 1). No participants dropped out of the program. The test assessing the homogeneity of participant characteristics and dependent variable (self-care) showed no significant differences between the 2 groups, indicating that they were comparable at baseline (Table 2).

XSL•FO
RenderX

Figure 1. Participant flow diagram. *Reasons for exclusion: severe pain (n=18), scheduled surgery (n=9), and hearing impairments (n=5). **Reasons for declined participation: time constraints or being busy (n=28), family-related inconvenience (eg, relative unavailable; n=9), and having already sought information independently (eg, reading materials or video clips; n=5). Some individuals reported more than one reason for declining participation. mHealth: mobile health.



XSL•FO
RenderX

Table . Homogeneity test of participant characteristics and dependent variables.

Characteristics	Intervention group (n=23)	Comparison group (n=23)	*P* value
Sex, n (%)			>.99[a]
Female	20 (87.0)	20 (87.0)	
Male	3 (13.0)	3 (13.0)	
Age (y), mean (SD; range)	71.04 (8.5; 56-88)	70.43 (8.9; 56-87)	.81[b]
Marital status, n (%)			>.99[c]
Single or divorced	4 (17.4)	3 (13.0)	
Married	19 (82.6)	20 (87.0)	
Educational level, n (%)			.23[a]
Secondary school or lower	12 (52.2)	7 (30.4)	
Bachelor's degree or higher	11 (47.8)	16 (69.6)	
Monthly income (US $), n (%)			.77[a]
≤700	13 (56.5)	14 (60.9)	
>700	10 (43.5)	9 (39.1)	
Health insurance scheme, n (%)			.20[a]
Civil servant medical benefit scheme	18 (78.3)	14 (60.9)	
Self-pay	5 (21.7)	9 (39.1)	
Duration of osteoporosis (y), n (%)			.13[a]
<5	16 (69.6)	11 (47.8)	
5-10	7 (30.4)	12 (52.2)	
Underlying diseases, n (%)			.67[a]
Hypertension	6 (26.1)	5 (21.7)	
Bone and joint diseases	10 (43.5)	13 (56.5)	
Others	7 (30.4)	5 (21.7)	
BMI (kg/m^2), mean (SD; range)	22.81 (3.48; 18.02-33.33)	23.24 (3.32; 17.48 - 33.31)	

[a]Chi-square test.

[b]Independent-sample *t* test.

[c]Fisher exact test.

Between-Group Differences in Postintervention Self-Care

ANCOVA was conducted to compare postintervention self-care mean scores between the intervention and control groups while controlling for baseline self-care mean scores. The results showed that the intervention group had significantly higher postintervention self-care mean scores than the control group (adjusted mean difference 0.97, 95% CI 0.88-1.06; $F_{1,43}$=459.29; P<.001; partial η^2=0.914; Table 3).

MANCOVA was conducted to examine differences in the 3 domains of self-care between the intervention and comparison groups after controlling for baseline domain mean scores (medication adherence, diet and physical activity, and safety performance). The overall multivariate group difference was statistically significant, as indicated by the Pillai trace (V=0.939; $F_{3,39}$=198.45; P<.001; partial η^2=0.939), suggesting significant differences between groups for all 3 domains after controlling for baseline domain mean scores. Compared with the comparison group, the intervention group demonstrated significantly higher postintervention mean scores for diet and physical activity (adjusted mean difference 1.02, 95% CI 0.89-1.15; $F_{1,41}$=245.04; P<.001; partial η^2=0.857), safety performance (adjusted mean difference 1.04, 95% CI 0.88-1.20; $F_{1,41}$=171.20; P<.001; partial η^2=0.807), and medication adherence (adjusted mean difference 0.81, 95% CI 0.68-0.95; $F_{1,41}$=144.54; P<.001; partial η^2=0.779; Table 3).

XSL•FO
RenderX

Table . Results of analysis of covariance (ANCOVA) and multivariate ANCOVA (MANCOVA) comparing postintervention outcomes between groups after controlling for baseline mean scores.

Domains	Baseline (covariate), raw mean (SD)			Postintervention time point, raw mean (SD)		Postintervention time point, adjusted mean (SE)		Adjusted mean difference (95% CI)	*F* test (*df*)	*P* value	Effect size, partial η^2
	Intervention group	Control group	*P* value	Intervention group	Control group	Intervention group	Control group				
ANCOVA											
Total mean score (1-5)	3.18 (0.59)	2.94 (0.28)	.10[a]	4.20 (0.33)	3.12 (0.22)	4.14 (0.03)	3.18 (0.03)	0.97[b] (0.88-1.06)	459.29 (1,43)	<.001[c]	0.914
MANCOVA (Pillai trace=0.939)									198.45 (3,39)	<.001[c]	0.939
Medication adherence (range 1-5)	3.16 (0.86)	3.03 (0.45)	.52[a]	4.23 (0.44)	3.33 (0.30)	4.19 (0.05)	3.38 (0.05)	0.81[b] (0.68-0.95)	144.54 (1,41)	<.001[c]	0.779
Diet and physical activity (range 1-5)	3.22 (0.69)	2.89 (0.28)	.04[a,c]	4.17 (0.34)	3.02 (0.28)	4.11 (0.05)	3.09 (0.05)	1.02[b] (0.89-1.15)	245.04 (1,41)	<.001[c]	0.857
Safety performance (range 1-5)	3.13 (0.69)	2.97 (0.42)	.36[a]	4.23 (0.46)	3.11 (0.32)	4.19 (0.06)	3.16 (0.06)	1.04[b] (0.88-1.20)	171.20 (1,41)	<.001[c]	0.807

[a]Independent-sample *t* test.

[b]Difference calculated as the adjusted mean of the intervention group minus that of the comparison group.

[c]Statistically significant ($P<.05$).

Discussion

Main Outcomes

The findings indicate that the multicomponent supportive-educative mHealth intervention was associated with higher postintervention self-care mean scores among individuals with osteoporosis. Higher adjusted mean scores were observed across all domains of self-care (medication adherence, diet and physical activity, and safety performance), with the largest between-group difference observed for diet and physical activity.

Interpretation

The observed between-group difference may reflect the combined contribution of several intervention components. These components included the provision of clear and reliable disease-related information, weekly nurse-participant interactions, family support, sustained attention, and digital reinforcement. Disease-related knowledge may support self-care by enhancing individuals' understanding of their condition and increasing their confidence in managing it. This is consistent with the supportive-educative nursing system by Orem et al [13], which emphasizes teaching and guidance to strengthen self-care agency. In addition, the weekly nurse-participant interactions provided opportunities for participants to discuss concerns, receive feedback, and clarify misunderstandings, whereas family involvement and digital reinforcement may have strengthened self-care behaviors and sustained participants' engagement throughout the intervention.

Family support may also have played a role in these observed differences. When family members learn alongside individuals with osteoporosis, this facilitates shared review and reinforcement of the content and supports engagement in self-care. This reflects the emphasis by Orem et al [13] on providing an environment that supports personal development and self-care. Although evidence on family support intervention for osteoporosis management is limited, findings from exploratory studies suggest that family involvement may support self-care [17]. When one person misunderstands or forgets information, others can provide reminders or exchange ideas, thereby supporting sustained self-care practices. In many countries, including Thailand, extended family living and close family ties are common.

Although higher postintervention self-care mean scores were observed in the intervention group than in the control group, both groups showed increases in self-care mean scores over time. Therefore, the observed between-group differences should be interpreted cautiously as factors other than the intervention may also have influenced the findings. Participants in the control group received usual care, and repeated completion of the

XSL•FO
RenderX

self-care assessments may have increased awareness of self-care behaviors. To reduce potential confounding, participants were matched on sex, age, and back pain score, factors previously associated with self-care [6,9]. In addition, baseline self-care mean scores were controlled for in the ANCOVA and MANCOVA analyses. However, the results should be interpreted as indicating an association between the multicomponent supportive-educative mHealth intervention and higher postintervention self-care mean scores rather than as definitive evidence of a causal effect.

The multicomponent supportive-educative intervention delivered partly through Facebook provided accessible and flexible learning. Video-based content allowed for repeated viewing, whereas remote interaction with nurses enabled timely guidance and reinforcement. This continuous connection supported engagement beyond clinical settings and facilitated shared learning with family members. However, this finding contrasts with those of previous studies of app-based interventions that found no significant differences on physical activity, calcium intake, or vitamin D intake, possibly because of the lack of a personalized and interactive approach [16].

Participants' experience provides insights into how the intervention influenced self-care. Participants found dietary guidance easy to understand and reported actively implementing fall prevention strategies; in contrast, exercise was less frequently adopted. In this study, both groups reported rarely engaging in exercise before the intervention. After the intervention, however, the intervention group reported sometimes engaging in exercise, whereas the control group continued to report rarely engaging in exercise. Although exercise can increase bone mineral density, enhance muscle strength, and improve body balance, thereby reducing the risk of falls, adherence remains challenging. Previous studies have reported that exercise adherence among individuals with osteoporosis varies from 59% to 100% even when expert coaching or accessible facilities are provided [20]. These findings highlight the need for additional strategies to promote exercise adherence.

Implications for Practice and Research

This study has implications for practice. The observed between-group differences suggest that the multicomponent supportive-educative intervention delivered partly through Facebook may warrant evaluation in randomized, attention-controlled studies before clinical implementation. Nurses may consider using digital platforms to extend care beyond clinical settings and engage family members as active partners in care while recognizing that the effectiveness of this approach requires confirmation in more rigorous studies.

Limitations

This study has several limitations. As an exploratory pilot study with nonrandom allocation, the findings should be interpreted with caution because factors other than the intervention, including the unequal intensity of support between groups, may have been associated with the observed between-group differences. The substantial effect sizes should also be interpreted cautiously as the relatively small sample size and reliance on self-reported measures may have influenced the magnitude of the estimated associations. In addition, recruitment from a single hospital may limit the generalizability of the findings. The medication adherence domain demonstrated low internal consistency at baseline (Cronbach α=0.56), which may have reduced the precision of the estimated associations for this domain. Self-care was assessed using self-reported measures, which may be subject to response bias. Participants and researchers were not blinded to group assignment, and contamination between groups may have occurred because participants were recruited from the same setting. Finally, the short follow-up period did not allow for evaluation of the long-term sustainability of the observed improvements. Future studies using randomized controlled designs, larger and more diverse samples, objective measures of self-care, and longer follow-up periods are needed to confirm these findings.

Conclusions

In conclusion, the multicomponent supportive-educative mHealth intervention was associated with higher self-care mean scores across multiple domains among individuals with osteoporosis. The observed between-group differences may reflect the integration of educational content, family involvement, ongoing support, and digital delivery. However, they should be interpreted with caution given the exploratory pilot study and relatively small sample size. Further rigorous randomized controlled trials are needed to confirm the efficacy of this intervention.

Acknowledgments

The authors would like to thank the participants for their involvement in the study. The authors also express their appreciation to Professor Daniela D'Angelo, National Center for Clinical Excellence, Healthcare Quality, and Safety, Rome, Italy, for permission to use the Self-Care of Osteoporosis Scale. Generative AI (ChatGPT; OpenAI) was used for English-language editing. AI was not used in the generation of the manuscript text and tables.

Funding

This study was supported by the Faculty of Nursing, Chulalongkorn University (grant 1/2569; January 9, 2026).

Data Availability

All data included in this study are available on request by contacting the corresponding author.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Intervention protocol.
[DOCX File, 18 KB - nursing_v9i1e101025_app1.docx]

Multimedia Appendix 2
Analysis of covariance and multivariate analysis of covariance statistical output.
[DOCX File, 2333 KB - nursing_v9i1e101025_app2.docx]

Checklist 1
TREND statement checklist.
[PDF File, 267 KB - nursing_v9i1e101025_app3.pdf]

References

1. Salari N, Ghasemi H, Mohammadi L, et al. The global prevalence of osteoporosis in the world: a comprehensive systematic review and meta-analysis. J Orthop Surg Res 2021 Oct 17;16(1):609. [doi: 10.1186/s13018-021-02772-0] [Medline: 34657598]
2. Fahimfar N, Noorali S, Yousefi S, et al. Prevalence of osteoporosis among the elderly population of Iran. Arch Osteoporos 2021 Jan 21;16(1):16. [doi: 10.1007/s11657-020-00872-8] [Medline: 33475880]
3. Kanis JA, Norton N, Harvey NC, et al. SCOPE 2021: a new scorecard for osteoporosis in Europe. Arch Osteoporos 2021 Jun 2;16(1):82. [doi: 10.1007/s11657-020-00871-9] [Medline: 34080059]
4. Farshbaf-Khalili A, Malekmirzaei E, Babaie S, Pakpour V. The determinants of self-care behaviors and calcium-vitamin D intake in elderly women: a comparative study of osteoporosis and healthy counterparts. Sci Rep 2025 Jul 26;15(1):27290. [doi: 10.1038/s41598-025-12755-4] [Medline: 40715503]
5. LeBoff MS, Greenspan SL, Insogna KL, et al. The clinician's guide to prevention and treatment of osteoporosis. Osteoporos Int 2022 Oct;33(10):2049-2102. [doi: 10.1007/s00198-021-05900-y] [Medline: 35478046]
6. Paşali Ki˙li˙t T, Onbaşi K, Özyi˙ği˙t F. Treatment adherence levels and factors affecting adherence in patients receiving osteoporosis treatment. Clin Exp Health Sci 2020;10(1):40-45. [doi: 10.33808/clinexphealthsci.533942]
7. Cittadini N, D'Angelo D, Basilici Zannetti E, et al. Development and testing of a new instrument to measure self-care in patients with osteoporosis: the Self-Care of Osteoporosis Scale. Int J Bone Frag 2021;1(1):28-33. [doi: 10.57582/IJBF.210101.028]
8. Malekmirzaei E, Farshbaf-Khalili A, Pakpour V. The self-care behaviors and health literacy can play important preventive roles in older female osteoporosis. BMC Womens Health 2023 Aug 11;23(1):429. [doi: 10.1186/s12905-023-02546-2] [Medline: 37568152]
9. Ansari A, Fahimfar N, Noruzi A, et al. Health information-seeking behavior and self-care in women with osteoporosis: a qualitative study. Arch Osteoporos 2021 May 5;16(1):78. [doi: 10.1007/s11657-021-00923-8] [Medline: 33954873]
10. Fallahi A, Valiee S, Chan SW. Needs of women with osteoporosis in disease self-management: a qualitative study. Health Scope 2019;8(2):e57234. [doi: 10.5812/jhealthscope.57234]
11. Gai QY, Lv H, Li YP, Fu QM, Li P. Education intervention for older adults with osteoporosis: a systematic review. Osteoporos Int 2020 Apr;31(4):625-635. [doi: 10.1007/s00198-019-05166-5] [Medline: 31828364]
12. Verdonck C, Willems R, Borgermans L. Implementation and operationalization of Integrated People-Centred Health Services delivery strategies in integrated osteoporosis care (IOC) initiatives: a systematic review. Osteoporos Int 2023 May;34(5):841-865. [doi: 10.1007/s00198-023-06678-x] [Medline: 36695826]
13. Orem DE, Taylor SG, Renpenning KM. Nursing: Concepts of Practice, 6th edition: Mosby; 2001.
14. Educating the elderly can help reduce the risk of recurrent fractures [Article in Thai]. Hfocus News Agency. 2018. URL: https://www.hfocus.org/content/2018/10/16469 [accessed 2026-08-14]
15. Chitsongboon K. Factors associated on one-year mortality rate after osteoporotic hip fracture patients at Phetchabun Hospital [Article in Thai]. Mahasarakham Hosp J 2018;15(2):13-22 [FREE Full text]
16. Alhussein G, Hadjileontiadis L. Digital health technologies for long-term self-management of osteoporosis: systematic review and meta-analysis. JMIR Mhealth Uhealth 2022 Apr 21;10(4):e32557. [doi: 10.2196/32557] [Medline: 35451968]
17. Zubick P, Dahlke S. Family/caregiver influence on osteoporosis management for older people: an integrative review. Osteoporos Int 2024 Jul;35(7):1153-1163. [doi: 10.1007/s00198-024-07081-w] [Medline: 38622263]
18. Des Jarlais DC, Lyles C, Crepaz N, TREND Group. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. Am J Public Health 2004 Mar;94(3):361-366. [doi: 10.2105/ajph.94.3.361] [Medline: 14998794]
19. McCaffery M, Beebe A. Pain: Clinical Manual for Nursing Practice: Mosby; 1989.

XSL•FO
RenderX

20. Liu Y, Zhang C, Zhang L, Ma X, Dong Z. Promoting and impeding factors for exercise in patients with osteopenia or osteoporosis: a systematic review. Int J Nurs Stud Adv 2025;10:100456. [doi: 10.1016/j.ijnsa.2025.100456] [Medline: 41551858]

Abbreviations

ANCOVA: analysis of covariance
MANCOVA: multivariate analysis of covariance
mHealth: mobile health
SCOS: Self-Care of Osteoporosis Scale
TREND: Transparent Reporting of Evaluations With Nonrandomized Designs

Edited by J Sarvestan; submitted 11.May.2026; peer-reviewed by B Cedeno-Veloz, C Suntayakorn; revised version received 01.Aug.2026; accepted 06.Aug.2026; published 20.Aug.2026.

Please cite as:
Pangutha N, Jittimanee S, Saksaen P
Multicomponent Supportive-Educative mHealth Intervention for Self-Care Among Individuals With Osteoporosis in Thailand: Exploratory Pretest-Posttest Pilot Study
JMIR Nursing 2026;9:e101025
URL: https://nursing.jmir.org/2026/1/e101025
doi:10.2196/101025

© Natthida Pangutha, Sirinapha Jittimanee, Porramat Saksaen. Originally published in JMIR Nursing (https://nursing.jmir.org), 20.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI-Based Structured Information Extraction From Synthetic Nursing Handover Transcripts: Comparative Evaluation of Large Language Models

Aaron Conway[1,2,3], BN, RN, PhD; Adriana Hada[3], RN, PhD; Jessica Schluter[3], RN, PhD; Hui (Grace) Xu[1,2], NP, PhD; Dan Lowden[4], MBBS; Tim Miller[5], PhD; Ken Donald[6], MBBS, PhD; Andrew Teodorczuk[3,7,8], BSc, MD, MBChB

[1]Centre for Healthcare Transformation, Queensland University of Technology, Park Rd, Kelvin Grove, Brisbane, QLD, Australia

[2]School of Nursing, Queensland University of Technology, Brisbane, Australia

[3]The Prince Charles Hospital, Metro North Health, Brisbane, Australia

[4]Caboolture Hospital, Brisbane, Australia

[5]School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, Australia

[6]School of Medicine & Dentistry, Griffith University, Gold Coast, Queensland, Australia

[7]The University of Queensland Northside Clinical Unit, The University of Queensland, Brisbane, Australia

[8]Royal Australian and New Zealand College of Psychiatrists, Melbourne, Australia

Corresponding Author:
Aaron Conway, BN, RN, PhD
Centre for Healthcare Transformation, Queensland University of Technology, Park Rd, Kelvin Grove, Brisbane, QLD, Australia

Abstract

Background: Clinical handover is the process during which responsibility and accountability for care are transferred between clinicians. AI has the potential to improve the reliability and completeness of clinical handover by helping clinicians detect predefined content areas that have been communicated, identify explicit information gaps, and prompt clarification before responsibility is transferred.

Objective: This study evaluated the performance of several large language models and prompt optimization strategies for structured information extraction of synthetic nursing handover transcripts.

Methods: Two registered nurses independently annotated a dataset of 203 synthetic handover transcripts to produce consensus labels for information extraction tasks. Tasks included (1) labeling spans of text into SBAR (Situation, Background, Assessment, Recommendation) categories, (2) content detection to determine if specific pieces of information were communicated, and (3) labeling spans of text that communicated information using uncertain terms that included a subtask for identifying unknown facts. Baseline and Genetic-Pareto (GEPA)–optimized prompts were compared for the GPT-5.2, GPT-5-nano, and MedGemma 27B large language models. Additionally, the LangExtract framework was evaluated for span-extraction tasks.

Results: The GPT-5.2–optimized model achieved a micro–F_1-score of 0.85 (95% CI 0.83 - 0.88) for content detection, an absolute improvement of +0.08 compared with the matched baseline. GPT-5-nano also performed better after optimization for content detection (micro–F_1-score 0.81, 95% CI 0.78 - 0.84), suggesting that this structured task was not limited to the highest-capacity model. For SBAR span extraction, GPT-5.2 with prompt optimization achieved a micro–F_1-score of 0.76 (95% CI 0.72 - 0.79), improving by +0.24 compared with baseline and exceeding LangExtract; GPT-5-nano also improved to a micro–F_1-score of 0.69 (95% CI 0.66 - 0.72). Broad uncertainty-span extraction remained comparatively weak despite prompt optimization (micro–F_1-score 0.41, 95% CI 0.33 - 0.48; absolute improvement +0.06). In contrast, explicit unknown-fact extraction was more accurate with GPT-5.2 (micro–F_1-score 0.84, 95% CI 0.63 - 1.00), GPT-5-nano (micro–F_1-score, 0.84 95% CI 0.63 - 1.00), and MedGemma 27B (micro–F_1-score 0.80, 95% CI 0.63 - 1.00). Genetic-Pareto–optimized prompts outperformed the LangExtract approach across each span-extraction task.

Conclusions: Prompt optimization improved matched-model point estimates, with the highest performance observed for predefined content detection and SBAR span extraction. Broad uncertainty extraction remained less accurate than the narrower unknown-fact task. These technical results do not establish clinical effectiveness, safety, or readiness for real-time use. Validation using authentic nursing handover communication and prospective evaluation in clinical workflows are required before clinical application.

(JMIR Nursing 2026;9:e106133) doi:10.2196/106133

XSL•FO
RenderX

KEYWORDS

artificial intelligence; clinical handover; nursing; patient handoff; natural language processing; large language models; prompt engineering; information extraction; patient safety; machine learning

Introduction

Clinical handover is the process during which responsibility and accountability for some or all aspects of a patient's care are transferred to another clinician or clinical team [1]. In nursing, shift-to-shift and transfer handovers support continuity of surveillance, care priorities, pending actions, and the inclusion of patient and family concerns. Inaccurate, incomplete, or misinterpreted communication at this transition can contribute to delays, duplicated work, and preventable harm [2,3]. International and Australian patient-safety standards therefore prioritize structured clinical handover, particularly at shift changes, transfers, and discharge [1,4].

Structured handover processes aim to standardize minimum content and the format of exchange so that critical information is predictably conveyed, acted upon, and auditable. In Australia, the Communicating for Safety Standard emphasizes structured clinical handover while retaining opportunities for questions, clarification, and confirmation [1,5]. One widely used framework is SBAR (Situation, Background, Assessment, Recommendation), which organizes clinician-to-clinician handover content into 4 information categories [6]. Recent systematic review evidence suggests that structured handoff protocols may improve some safety outcomes, but the certainty and implementation fidelity vary by protocol and setting; evidence specific to SBAR remains low certainty [7].

Even where structured tools are mandated or encouraged, their enactment in nursing handovers remains shaped by ward-level organizational and cultural conditions, time demands, interruptions, and the need to communicate patient-specific information for continuity of care [8,9]. This variability is amplified in bedside nursing handover where patient/family involvement is increasingly emphasized, yet research highlights tensions between standardization (predictability) and tailoring (patient-centeredness), along with barriers related to confidentiality and clinician concerns [10,11]. Reviews of handover mnemonics similarly emphasize local validation, clarification, and readback rather than assuming that every element is universally relevant [6].

Traditional supervised information-extraction methods can be highly effective for stable tasks with sufficiently large, task-specific labeled datasets. However, there are several potential advantages of using large language models (LLMs) in this context. For example, the same instruction-driven model can be configured for heterogeneous, context-dependent tasks. This may be useful where annotated nursing handover data are limited and communication is conversational and nonlinear. Prior work has demonstrated that few-shot clinical information extraction with LLMs can be accurate, while also showing that performance depends materially on prompt design and task framing [12,13].

Recent advances in LLMs create an opportunity to support clinical handover by analyzing information as it is communicated. Other studies have examined AI-based approaches to support clinical handover, although evaluations of these technologies remain limited [14]. For example, a recent multihospital study used an LLM to pregenerate content to support the preparation for handover [15]. An alternative application of AI yet to be investigated is to support the verbal clinician-to-clinician exchange itself. This spoken exchange remains central to transferring responsibility of care in clinical settings and provides opportunities to question, clarify, and confirm information [1,9]. Accurate extraction of structured information from the spoken exchange is a prerequisite for developing this form of communication support. The aim of this study was to evaluate the accuracy of LLM structured information extraction from synthetic clinical handover transcripts.

Methods

Study Design

This study used a model-evaluation design, corresponding to the "LLM evaluation" research-design category in TRIPOD-LLM (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis–Large Language Models) [16] (Checklist 1). This category covers studies that assess existing LLMs for their accuracy or suitability for a specific health care task. We compared selected LLMs and prompt configurations on predefined multilabel classification and span-extraction tasks using completed synthetic nursing handover transcripts and registered nurse annotations. The study was not an evaluation in a health care setting because it did not test workflow integration or clinical, administrative, or workforce outcomes.

Tasks related to clinical handover communication that were considered potentially augmentable with AI assistance were identified by the researchers using a co-design process with clinicians and consumers, which will be reported separately. The tasks were as follows:

- Extracting spans of text from transcripts that aligned with the SBAR framework for structuring clinical handover. In this study, SBAR span extraction was operationalized as a sequence-labeling task to identify contiguous spans of text in handover transcripts that corresponded to each SBAR category. This task evaluated whether models could map conversational handover text to SBAR-labeled spans, rather than whether the transcripts themselves followed a clean sequential SBAR structure.
- Identifying if key elements were communicated in handover transcripts as a content detection task. In this study, this task was operationalized as a checklist of items that are recommended to be addressed during nursing clinical handover, which were developed as part of quality-improvement processes at the researchers' institution. The checklist items are summarized in Table 1.

XSL•FO
RenderX

- Identifying spans of text in transcripts that were communicated using uncertain terms. In this study, this task was operationalized as span annotation of utterances during handover that conveyed incomplete knowledge, vague or hedged wording, imprecise timing, second-hand sourcing, unclear procedures, or unclear responsibility for follow-up actions. These forms of uncertainty were treated as potentially clinically important because they may indicate information that requires clarification or verification by the receiving clinician. The uncertainty categories and example guidance provided to annotators are summarized in Table 2.

Table . Checklist items used to operationalize identification of key concepts and entities in handover transcripts.

Category	Checklist item
Patient involvement	• Clinician introductions • Introduction of clinicians to the patient or carer • Invitation for the patient or carer to participate in handover
Identification	• Verification of 3 patient identifiers
Situation	• Primary diagnosis or reason for admission • Significant events or complications • Current status, including pending tests/procedures and interim plans/orders
Background	• Relevant clinical and social history, including comorbidities • Falls risk • Pressure injury risk • Allergies • Advance care planning
Assessment	• Observations, deterioration score, and recent escalations • Pain management • Devices, lines, and vascular access • Critical monitoring and alarms • Nutrition and dietary restrictions • Fluid balance and fluid restrictions • Infusions • Medication chart review, including high-risk medicines • Pathology results or pending investigations • Mobility and use of aids • Skin integrity and related interventions
Recommendation	• Discharge plan • Critical actions required • Follow-up care plan or pathway actions • Patient or carer goals and preferences

Table . Uncertainty categories used to support annotator identification of uncertainty-related spans in handover transcripts.

Category	Definition/when to use	Example from handover speech
Hedge/probability language	The speaker indicates partial confidence or doubt about information.	“I think ENT reviewed him”; “He should be going to theatre soon.”
Vague/qualitative expression	Information is described using imprecise or subjective language.	“He looks fine now”; “Seems okay.”
Unknown fact/explicit lack of knowledge	The speaker openly states missing knowledge or incomplete data.	“Not sure if consent’s been signed”; “I don’t know his allergies.”
Indefinite timing	Timing or schedule for an event is vague or lacks precision.	“Later today”; “After the round.”
Source uncertainty	Information relies on a second-hand or unverifiable source.	“ENT said he’s on the list”; “Night nurse told me.”
Procedural uncertainty	The next step in care is unclear or the plan is not explicitly stated.	“You might want to check his IV.”
Responsibility uncertainty	A required task or follow-up is mentioned, but it is unclear who is responsible for performing it.	“Bloods to be checked later”; “Needs review this afternoon.”

XSL•FO
RenderX

Data Sources

We used the publicly available National Information and Communications Technology Australia (NICTA) Synthetic Nursing Handover Dataset, which contains synthetic recordings of clinical handovers delivered by a registered nurse based on patient profiles with cardiovascular, neurological, renal, and respiratory conditions [17]. In this dataset, handover monologs were generated from comprehensive patient profiles that included information such as the patient's name, age, admission history, inpatient duration, and the familiarity between the nurses giving and receiving the handover. The nurse was instructed to simulate a bedside shift-to-shift handover within a medical ward setting [17].

For our study, we used 100 handover samples from the training partition of the NICTA dataset. First, audio recordings from the NICTA dataset were transcribed using the OpenAI Whisper speech-to-text model. Second, 3 videos depicting conversational nursing shift-to-shift handovers were transcribed in a similar manner to provide examples of interactive handover dialogue. These videos were developed at the authors' institution for educational purposes to demonstrate best practices for clinical handovers that are used in the undergraduate nursing program. Transcripts from the educational videos were then used as few-shot examples within the DSPy framework [18] using the BootstrapFewShot optimizer to guide the transformation of the 100 NICTA monolog transcripts into 2-sided conversational handovers so that the dataset would better reflect real-world contemporary clinical handover interactions. To broaden the range of clinical contexts represented in the dataset, we further synthesized 103 additional handover examples using the GPT-5 model interactively in a chat interface. These scenarios included inter- and intrahospital transfers, postprocedural handovers, emergency department transitions, and handovers that involved patients with complex mental health care needs. The final dataset comprised 203 synthetic handover transcripts used for subsequent annotation and model development.

Ethical Considerations

Ethics approval was not sought because this was not human research as defined by the National Statement on Ethical Conduct in Human Research [19]. The study involved only synthetic handover transcripts generated for model evaluation, with no recruitment, observation, or testing of human participants, and no use of patient, clinician, clinical-record, personal, identifiable, or potentially reidentifiable data. Accordingly, the study did not require human ethics review, and a formal exemption from ethics review was not applicable.

Annotation

Two annotators, who are experienced clinically active registered nurses, independently annotated transcripts using a structured rubric. Labeled annotations that met consensus between reviewers were used as the reference labels for downstream model development and evaluation.

Annotation was performed using the Prodigy annotation software using a custom interface that presented each transcript in 3 components [20]. First, annotators highlighted relevant spans of text and assigned labels corresponding to the SBAR framework together with markers of communicative uncertainty, including vagueness, hedging, unknown facts, indefinite timing, source uncertainty, procedural uncertainty, and uncertainty regarding responsibility. Overlapping span labels were permitted where a passage served more than 1 communicative function. Second, annotators completed a multiple-response checklist indicating whether predefined handover elements were present in the transcript.

Both annotators reviewed all transcripts independently within the same annotation environment and were supported by written guidance and examples to promote consistent interpretation of the coding framework. For the creation of the reference standard, a consensus dataset was derived by retaining only those span annotations and checklist items for which both annotators agreed. In practical terms, this meant that only text segments assigned the same label by both reviewers, and only checklist items selected by both reviewers, were carried forward for downstream model development and evaluation.

LLM Prompt Optimization Methods

Annotated handover transcripts were first partitioned deterministically into optimization (75%) and evaluation (25%) subsets using a fixed random split. Prompt optimization was performed within DSPy [18], which is a Python framework that can be used for prompt optimization using feedback from model outputs to improve task performance. For each task, we defined a task-specific DSPy signature that specified the transcript as input and a constrained structured output. Checklist prediction was formulated as multilabel classification, where the task was to return a list of the items from the checklist that were covered in the transcript. The SBAR and uncertainty-related tasks were formulated as span extraction requiring the model to return verbatim text segments from the transcript together with the appropriate label. The uncertainty-related tasks were further subdivided into a broad uncertainty-span extraction task (which included all uncertainty categories) and a more specific unknown-fact extraction task (which included only spans labeled as unknown facts).

Across the baseline DSPy evaluations, Genetic-Pareto (GEPA)–based DSPy optimization experiments, and LangExtract experiments, we purposively selected 3 underlying language models with different deployment profiles: OpenAI GPT-5.2 as the higher-capacity proprietary model, OpenAI GPT-5-nano as a smaller lower-cost proprietary model [21], and MedGemma 27B as a 27-billion-parameter open-weight model developed for medical tasks [22]. This panel was intended to examine matched prompt effects across contrasting model profiles, not to provide an exhaustive leaderboard of all contemporary LLMs. For DSPy baseline and GEPA runs, the same task model was used before and after optimization so that differences reflected the prompt configuration rather than a change in the underlying model. GPT-5.2 and GPT-5-nano were accessed through cloud-hosted OpenAI API end points; MedGemma 27B was run locally on institutional high-performance computing infrastructure.

GEPA was used to optimize the task prompts [23]. It iteratively evaluated candidate task instructions, combined numeric task scores with natural-language error feedback, and used GPT-5.2

XSL•FO
RenderX

as a separate reflection model to propose revisions. Optimization used 576 scoring calls. Checklist optimization targeted multilabel agreement, while span-task optimization rewarded same-label text overlap using intersection over union (IoU) so that closer boundaries received higher scores. A transcript with neither a reference span nor a predicted span for a target label was treated as a correct negative. The final compiled prompts were fixed before evaluation on the held-out partition. For reporting, span detection performance and boundary overlap were presented separately.

In addition to DSPy prompt optimizations, we conducted separate evaluations using the LangExtract framework, as a prompt-based few-shot structured information extraction approach for the SBAR, uncertainty-span, and unknown-fact span tasks [24]. These experiments used task-specific prompt descriptions together with annotated in-context examples derived from the reference data. We used 10 annotated examples from the training partition as few-shot exemplars, and inference was then performed on the full held-out test partition for each task.

Data Analysis

Interrater agreement was assessed with Cohen κ and mean IoU among matched spans. Agreement CIs were calculated with 2000 bootstrap resamples of transcript pairs using a fixed random seed.

Performance was measured for the checklist content detection task at the level of individual labels with counts of true positives, false positives, false negatives, and true negatives, together with precision, recall, and F_1-score. Aggregate performance was summarized using micro-averaged (pooled), macro-averaged (unweighted mean), and support-weighted precision, recall, and F_1-score across labels.

For span-extraction tasks, precision, recall, and F_1-score were calculated from predicted spans as binary detection measures, so that these statistics reflected the model's ability to identify the correct labeled spans. Span-boundary agreement was reported separately using the mean IoU across matched pairs. We additionally calculated per-label descriptive metrics including the number of reference spans, the number of predicted spans, matched-span precision, recall, F_1-score, and mean IoU.

Sampling uncertainty was summarized with 95% CIs calculated using nonparametric bootstrap resampling. For each result, we resampled transcripts with replacement 2000 times using a fixed random seed and recalculated the relevant metric from the pooled counts in each resample. Confidence limits are reported as the 2.5th and 97.5th percentiles of the bootstrap distribution.

Results

Preconsensus Interrater Agreement

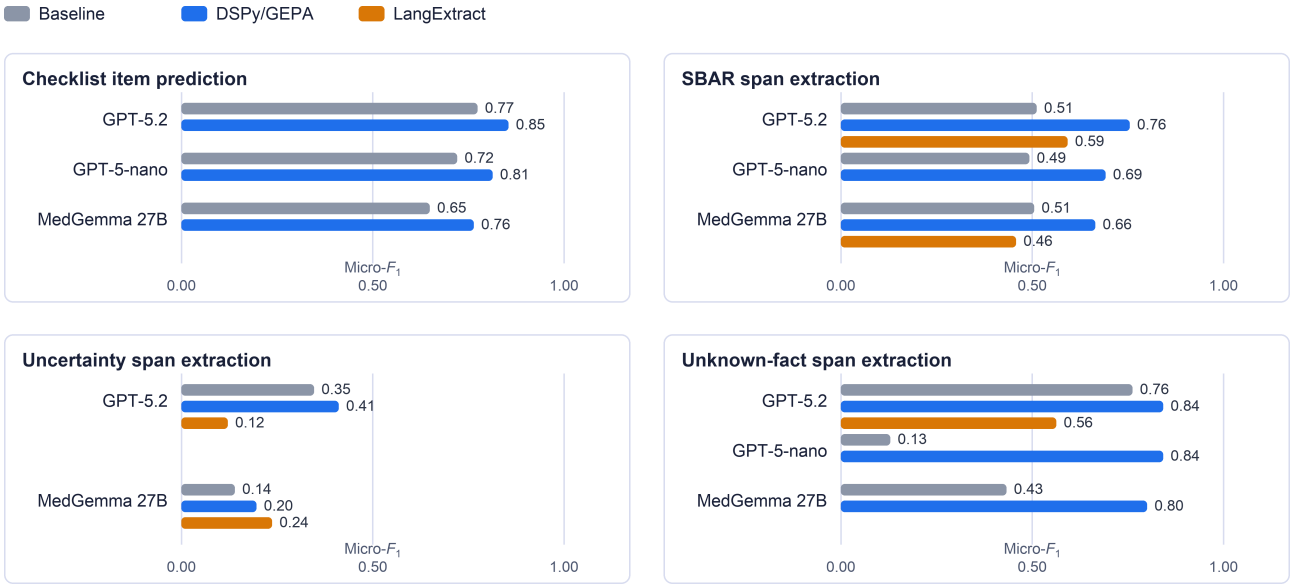
Cohen κ was 0.75 (95% CI 0.73 - 0.76) for checklist decisions, 0.70 (95% CI 0.68 - 0.72) for pooled SBAR token-by-label decisions, 0.12 (95% CI 0.08 - 0.15) for broad uncertainty, and 0.40 (95% CI 0.18 - 0.64) for unknown facts. Among overlapping same-label spans, the mean IoU was 0.86 (95% CI 0.85 - 0.87) for SBAR, 0.71 (95% CI 0.61 - 0.81) for broad uncertainty, and 0.78 (95% CI 0.56 - 0.98) for unknown facts. The relatively high matched-span IoU indicates similar boundaries when both nurses identified the same span type, with disagreement in uncertainty annotations arising mainly over whether and how to label an expression.

Key Findings

Within-model comparisons showed consistent point-estimate gains with DSPy/GEPA over matched unoptimized-prompt baselines. For GPT-5.2, micro–F_1-score increased from 0.77 (95% CI 0.74 - 0.80) to 0.85 (95% CI 0.83 - 0.88) for checklist prediction, from 0.51 (95% CI 0.47 - 0.55) to 0.76 (95% CI 0.72 - 0.79) for SBAR span extraction, from 0.35 (95% CI 0.28 - 0.41) to 0.41 (95% CI 0.33 - 0.48) for uncertainty span extraction, and from 0.76 (95% CI 0.50 - 1.00) to 0.84 (95% CI 0.63 - 1.00) for unknown-fact extraction. Across span tasks, LangExtract generally had lower point estimates than the corresponding highest DSPy/GEPA configuration. Among matched span predictions, the mean IoU exceeded 0.8 for the highest-performing GPT-5.2 span-extraction configurations. Figure 1 summarizes matched within-model comparisons across tasks for the 3 evaluated models.

XSL•FO
RenderX

Figure 1. Micro–F_1-scores for baseline, Genetic-Pareto (GEPA)–optimized, and LangExtract strategies for each model within each task.



SBAR Span Extraction

Among SBAR configurations, the highest overall score was achieved by DSPy/GEPA-optimized GPT-5.2 (micro–F_1-score 0.76, 95% CI 0.72 - 0.79), followed by DSPy/GEPA-optimized GPT-5-nano (micro–F_1-score 0.69, 95% CI 0.66 - 0.72) and LangExtract GPT-5.2 (micro–F_1-score 0.59, 95% CI 0.56 - 0.62). Table 3 provides label-level results for the best-performing GPT-5.2 DSPy/GEPA SBAR configuration.

Table . Per-label SBAR (Situation, Background, Assessment, Recommendation) metrics for the best-performing GPT-5.2 DSPy/Genetic-Pareto (GEPA) model.

Label	Reference spans	Predicted spans[a]	Recall (95% CI)	Precision (95% CI)	Mean IoU[b] (95% CI)	F_1-score (95% CI)
ASSESSMENT	194	192	0.77 (0.72 - 0.82)	0.78 (0.72 - 0.83)	0.77 (0.72 - 0.81)	0.77 (0.73 - 0.81)
BACKGROUND	47	50	0.72 (0.61 - 0.84)	0.68 (0.57 - 0.79)	0.66 (0.56 - 0.75)	0.70 (0.60 - 0.80)
RECOMMENDATION	113	124	0.76 (0.69 - 0.84)	0.69 (0.62 - 0.78)	0.78 (0.73 - 0.83)	0.73 (0.67 - 0.78)
SITUATION	73	52	0.68 (0.57 - 0.81)	0.96 (0.88 - 1.00)	0.82 (0.75 - 0.88)	0.80 (0.72 - 0.89)

[a]Predicted spans: model-generated spans mapped back to the transcript.

[b]IoU: intersection over union for matched span boundaries.

Within this GPT-5.2 SBAR comparison, macro-precision increased from 0.41 (95% CI 0.37 - 0.44) to 0.78 (95% CI 0.73 - 0.82), macro-recall from 0.69 (95% CI 0.63 - 0.75) to 0.73 (95% CI 0.69 - 0.78), and macro–F_1-score from 0.49 (95% CI 0.46 - 0.53) to 0.75 (95% CI 0.71 - 0.79). Span-boundary agreement among matched predictions was strongest for SITUATION and RECOMMENDATION, as shown by the label-level mean IoU estimates in Table 3.

Checklist Task

For checklist prediction, the best overall result was achieved by DSPy/GEPA-optimized GPT-5.2 (micro–F_1-score 0.85, 95% CI 0.83 - 0.88; macro–F_1-score 0.73, 95% CI 0.63 - 0.76; and support-weighted F_1-score 0.85, 95% CI 0.82 - 0.88). DSPy/GEPA-optimized GPT-5-nano also performed competitively (micro–F_1-score 0.81, 95% CI 0.78 - 0.84), while MedGemma 27B reached a micro–F_1-score of 0.76 (95% CI 0.74 - 0.79). Tables 4 and 5 present grouped per-label estimates for accuracy, precision, recall, and F_1-score for the best-performing GPT-5.2 checklist model.

XSL•FO
RenderX

Table . Grouped per-label checklist performance for the best-performing GPT-5.2 DSPy/Genetic-Pareto (GEPA) model.

Checklist item	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)
Identification				
ID check of 3 patient identifiers	1.00	1.00	1.00	1.00
Situation				
Primary diagnosis | reason for admission	0.96 (0.90 - 1.00)	0.96 (0.89 - 1.00)	1.00	0.98 (0.94 - 1.00)
Current status (awaiting tests/procedures, on interim orders/plan)	0.65 (0.51 - 0.78)	0.73 (0.57 - 0.88)	0.71 (0.55 - 0.87)	0.72 (0.58 - 0.84)
Significant events or complications	0.84 (0.73 - 0.94)	0.62 (0.25 - 1.00)	0.50 (0.17 - 0.83)	0.56 (0.20 - 0.80)
Background				
Alerts-allergies	0.92 (0.84 - 0.98)	0.89 (0.73 - 1.00)	0.89 (0.71 - 1.00)	0.89 (0.76 - 0.98)
Relevant clinical and social history | comorbidities	0.98 (0.94 - 1.00)	0.94 (0.81 - 1.00)	1.00	0.97 (0.90 - 1.00)
Alerts-falls risk	0.98 (0.94 - 1.00)	1.00	0.67	0.80
Alerts-pressure injury risk	0.98 (0.94 - 1.00)	1.00	0.67	0.80
Advanced care planning	1.00	1.00	1.00	1.00
Assessment				
Observations | Q-ADDS | recent escalations	0.92 (0.84 - 0.98)	0.95 (0.87 - 1.00)	0.95 (0.87 - 1.00)	0.95 (0.89 - 0.99)
Medication chart | flag high risk meds	0.90 (0.82 - 0.98)	0.86 (0.71 - 0.97)	0.96 (0.87 - 1.00)	0.91 (0.81 - 0.98)
Devices | lines | vascular access	0.94 (0.88 - 1.00)	0.88 (0.75 - 1.00)	1.00	0.94 (0.86 - 1.00)
Mobility | aids	0.90 (0.80 - 0.98)	0.80 (0.61 - 0.95)	0.94 (0.80 - 1.00)	0.86 (0.72 - 0.97)
Pain management	0.94 (0.86 - 1.00)	0.89 (0.71 - 1.00)	0.94 (0.81 - 1.00)	0.91 (0.80 - 1.00)
Infusions	0.84 (0.73 - 0.94)	0.89 (0.67 - 1.00)	0.53 (0.27 - 0.79)	0.67 (0.40 - 0.86)
Pathology	0.88 (0.78 - 0.96)	0.74 (0.53 - 0.93)	0.93 (0.79 - 1.00)	0.82 (0.67 - 0.94)
Nutrition | restrictions	0.88 (0.78 - 0.96)	0.70 (0.50 - 0.89)	1.00	0.82 (0.67 - 0.94)
Fluid balance | restrictions	0.90 (0.80 - 0.98)	0.70 (0.38 - 1.00)	0.78 (0.44 - 1.00)	0.74 (0.44 - 0.94)
Skin integrity | interventions	0.90 (0.80 - 0.98)	0.55 (0.22 - 0.83)	1.00	0.71 (0.36 - 0.91)
Critical monitoring | alarms	0.96 (0.90 - 1.00)	0.00 (0.00 - 0.00)	0.00 (0.00 - 0.00)	0.00 (0.00 - 0.00)

XSL•FO
RenderX

Table . Grouped per-label checklist performance for the best-performing GPT-5.2 DSPy/Genetic-Pareto (GEPA) model (continued)[a].

Checklist item	Accuracy[b] (95% CI)	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)
Recommendation				
Care plan/pathway actions to follow up	0.90 (0.80 - 0.98)	0.90 (0.80 - 0.98)	1.00	0.95 (0.89 - 0.99)
Asked patient/carer about goals and preferences	0.80 (0.67 - 0.90)	0.00 (0.00 - 0.00)	0.00 (0.00 - 0.00)	0.00 (0.00 - 0.00)
Discharge plan	0.98 (0.94 - 1.00)	0.80 (0.33 - 1.00)	1.00	0.89 (0.50 - 1.00)
Critical actions required	0.92 (0.84 - 0.98)	0.20 (0.00 - 0.67)	1.00	0.33 (0.00 - 0.80)
Patient involvement				
Introduction of clinicians involved in handover to patient/carer	0.84 (0.73 - 0.94)	0.71 (0.55 - 0.87)	1.00	0.83 (0.71 - 0.93)
Invitation for patient/carer to participate in handover	0.82 (0.69 - 0.92)	0.65 (0.44 - 0.83)	0.94 (0.78 - 1.00)	0.77 (0.59 - 0.89)

[a]95% CIs are not provided for labels with few positive examples in the test set.

[b]Accuracy: (true positives + true negatives)/all evaluated transcripts for that checklist item.

Uncertainty and Unknown-Fact Span Extraction

Broad uncertainty-span extraction included all annotated uncertainty categories (hedging, vagueness, unknown facts, indefinite timing, source uncertainty, procedural uncertainty, and responsibility uncertainty) and had the lowest micro–F_1-score of the evaluated tasks. The highest DSPy/GEPA point estimate used GPT-5.2 and achieved a precision of 0.32 (95% CI 0.26 - 0.39), a recall of 0.56 (95% CI 0.44 - 0.67), a micro–F_1-score of 0.41 (95% CI 0.33 - 0.48), and a mean IoU of 0.83 (95% CI 0.77 - 0.90). The highest unoptimized-prompt baseline had a micro–F_1-score of 0.35 (95% CI 0.28 - 0.41), and the highest LangExtract uncertainty result had a micro–F_1-score of 0.24 (95% CI 0.17 - 0.31).

The narrower unknown-fact subtask had higher point estimates. DSPy/GEPA reached a micro–F_1-score of 0.84 (95% CI 0.63 - 1.00) with GPT-5.2, compared with the highest baseline micro–F_1-score of 0.76 (95% CI 0.50 - 1.00) and the LangExtract GPT-5.2 micro–F_1-score of 0.56 (95% CI 0.29 - 0.72). These results concern explicitly stated lack of knowledge, not facts absent from the transcript. Table 6 summarizes the uncertainty and unknown-fact span results used for these comparisons.

XSL•FO
RenderX

Table . Uncertainty and unknown-fact span extraction.

Task, model, and approach	Precision (95% CI)	Recall (95% CI)	F_1-score (95% CI)	Mean IoU[a] (95% CI)
Uncertainty				
GPT-5.2				
Baseline	0.24 (0.19 - 0.29)	0.65 (0.53 - 0.75)	0.35 (0.28 - 0.41)	0.62 (0.53 - 0.70)
DSPy[b]/GEPA[c]	0.32 (0.26 - 0.39)	0.56 (0.44 - 0.67)	0.41 (0.33 - 0.48)	0.83 (0.77 - 0.90)
LangExtract	0.35 (0.21 - 0.55)	0.07 (0.03 - 0.13)	0.12 (0.05 - 0.20)	0.35 (0.28 - 0.43)
MedGemma 27B				
Baseline	0.12 (0.07 - 0.18)	0.17 (0.10 - 0.23)	0.14 (0.08 - 0.20)	0.56 (0.39 - 0.70)
DSPy/GEPA	0.19 (0.12 - 0.31)	0.20 (0.13 - 0.28)	0.20 (0.13 - 0.27)	0.72 (0.55 - 0.88)
LangExtract	0.22 (0.15 - 0.29)	0.26 (0.18 - 0.36)	0.24 (0.17 - 0.31)	0.69 (0.57 - 0.80)
Unknown fact				
GPT-5.2				
Baseline	0.73 (0.44 - 1.00)	0.80 (0.56 - 1.00)	0.76 (0.50 - 1.00)	0.86 (0.66 - 0.99)
DSPy/GEPA	0.89 (0.71 - 1.00)	0.80 (0.56 - 1.00)	0.84 (0.63 - 1.00)	0.92 (0.74 - 0.99)
LangExtract	0.41 (0.17 - 0.59)	0.90 (0.79 - 1.00)	0.56 (0.29 - 0.72)	0.68 (0.45 - 0.84)
GPT-5-nano				
Baseline	0.08 (0.00 - 0.17)	0.40 (0.00 - 1.00)	0.13 (0.00 - 0.27)	0.74 (0.00 - 0.97)
DSPy/GEPA	0.89 (0.71 - 1.00)	0.80 (0.56 - 1.00)	0.84 (0.63 - 1.00)	0.89 (0.79 - 0.97)
MedGemma 27B				
Baseline	0.30 (0.09 - 0.48)	0.80 (0.56 - 1.00)	0.43 (0.16 - 0.61)	0.90 (0.72 - 0.98)
DSPy/GEPA	0.80 (0.67 - 1.00)	0.80 (0.56 - 1.00)	0.80 (0.63 - 1.00)	0.90 (0.72 - 0.98)

[a]IoU: intersection over union for matched span boundaries.

[b]DSPy: framework for declarative language model programming.

[c]GEPA: Genetic-Pareto.

Discussion

Principal Results

This study identified that prompt optimization with GEPA outperformed baseline evaluations across all of the comparisons, with higher aggregate performance for checklist prediction than for SBAR span extraction. This pattern is expected because using generative AI to perform clinical natural language processing tasks is sensitive to framing, label definitions, and example selection. GEPA was explicitly supplied with task-specific scores and error feedback that could align instructions more closely with the annotation scheme [13,23]. Optimization was particularly useful for improving targeted span extraction across the SBAR task.

The difference in accuracy that we identified between checklist prediction and SBAR span extraction provides an important insight to consider for designing AI applications to support clinical handover. Checklist prediction is a simpler structured-output task where each item is a bounded present/absent judgment. By contrast, SBAR extraction requires the model to locate clinically relevant text, assign a communicative category, and reproduce appropriate span boundaries. Structured handover tools aim to make minimum content predictable and reduce omissions, while still requiring adaptation to local clinical workflow [25-27]. As such, bounded present/absent outputs may be useful when the intended function is to prompt review of missing elements or support audit and quality monitoring. It is fortunate, then, that the task with the strongest performance in our study was the one most closely aligned with a recurring mechanism of handover-related harm, namely omitted or incomplete information [2,26,28]. Although checklist prediction achieved high accuracy, the consequences of both false positives and false negatives should be considered. A false negative would usually create review burden by prompting clarification of an item that was already communicated, but a false positive could create more serious false reassurance that a safety-critical element was conveyed when it was not. In addition, a checklist item not detected in the captured transcript is not equivalent to clinically necessary information having been omitted. A safer near-term design could therefore present AI outputs as source-linked prompts for clinician review, with clear distinctions between detected in transcript, not detected, and requires verification. Evidence has indicated that effective clinical decision support is most useful when integrated into a workflow as actionable and readily available support at the time and place of decision-making, while also recognizing the risk of automation bias when clinicians overrely on system outputs [29-31].

XSL•FO
RenderX

Registered nurse annotation agreement was comparatively high for checklist and SBAR annotations but very low for the broad uncertainty taxonomy. The latter finding indicates that vague, hedged, or context-dependent communication was difficult to operationalize consistently even with written guidance. However, the combination of low κ and relatively high matched IoU scores for broad uncertainty indicates that the principal difficulty was deciding whether uncertainty was present, rather than locating its boundaries once identified. As all model optimization and evaluation used the final consensus ratings between annotators, the results should therefore be interpreted as being conservative.

This study did not compare LLMs with non–LLM-based methods of structured information extraction. It therefore provides evidence about relative prompt optimization approaches and model configurations, not the superiority of LLMs over conventional information extraction. Nevertheless, evaluating LLM-based approaches is a logical next step in this handover-specific research program. In the original NICTA benchmark, a feature-engineered conditional random field achieved a macro–F_1-score of 0.702 across 35 handover-form categories, but performance was markedly uneven. F_1-score was 0.217 for the more abstract "other observations" category and 0.496 for future-care goals, tasks, and expected outcomes, and the error analysis identified clinically relevant missed and misclassified information [17]. These limitations were concentrated in categories requiring greater contextual differentiation. Instruction-tuned LLMs therefore warranted evaluation as a potentially more flexible approach to context-dependent handover information, particularly where prompt optimization can adapt extraction behavior using a modest labeled set.

The optimized SBAR extraction model achieved high span-boundary overlap among matched spans in our study. It should be considered, though, that SBAR is an idealized communication framework, while real clinical handovers often move nonlinearly, revisit information, distribute relevant details across the conversation, or place content between categories. For implementation, this suggests that extracted spans could be displayed with a link to their source transcript context and should support clinician review of what was said, rather than automatically transforming the handover into a definitive structured note. This preserves the benefits of structured communication while avoiding over-compression of clinically meaningful narrative context [25,32,33].

Implementation Considerations

The appropriate balance between recall and precision may differ by use case when considering the implementation of the tasks we evaluated in this study into an AI tool to support clinical handover. For real-time clinician support, the intervention should be framed as shaping safe handover dialogue rather than only producing a structured output after the exchange has ended. Higher recall may be appropriate where the system displays nondetected checklist items under an "Items to check" heading and allows the receiving clinician to mark each item as addressed, not relevant, or requiring clarification. This presentation would avoid implying that the information was definitely omitted while supporting clarification before responsibility transfers. For retrospective audit or compliance monitoring, precision becomes more important because false positives could overestimate handover quality and obscure residual safety risks. The present checklist results, with high recall but nontrivial false positives in several categories, therefore support cautious separation of 2 implementation pathways: real-time clinician-facing gap prompts and separately validated audit/reporting workflows [28,29,34].

Gold standard handovers include active verification and shared situational awareness rather than passive transfer of uncertain information [33,35-37]. Evaluation results of the broad uncertainty-span performance indicate that even a GEPA-optimized prompt with a state-of-the-art proprietary LLM is not yet reliable for detecting the full range of ambiguous, hedged, or context-dependent communication. This should be interpreted against the clinical reality that uncertainty is also difficult for humans to recognize and act on consistently, particularly under time pressure or when concern is communicated indirectly. Many forms of clinically important uncertainty are implicit and embedded in shared team understanding: phrases such as "he seems off today" or "not quite themself" can convey concern through context, trajectory, and prior knowledge rather than through explicit wording that a model can reliably extract. By contrast, the comparatively more accurate unknown-fact extraction results suggest a potentially useful role for identifying explicit information gaps that can be converted into clarification prompts for the receiving clinician. As foundation models continue to improve, accuracy on this task may also improve, although reliable use in safety-critical handover settings will still require ongoing empirical validation.

Relatedly, it is highly likely that further gains could be realized after implementation independent of general improvements in underlying LLM capability. If clinicians review AI-generated handover outputs and corrections are retained as labeled examples, subsequent optimization cycles using the same GEPA prompt optimization pipeline we used in this study would plausibly improve performance over time. Prior clinical natural language processing active-learning studies have demonstrated that selectively labeled examples can improve data efficiency for both clinical text classification and clinical named-entity recognition [38,39]. In practice, any such learning cycle should be treated as a controlled quality-improvement process, with monitoring and re-evaluation before revised prompts or models are released into clinical use [40].

Finally, it should be noted that this study measured extraction and classification performance against annotated labels but did not evaluate whether AI outputs changed clinician questioning, closed-loop communication, task completion, escalation, near-miss detection, or patient outcomes. Technical metrics are necessary but insufficient for determining whether a clinical AI tool improves safety in practice [31,37,41]. Recent commentary and discursive work on nursing automation similarly argue that evaluation should move beyond time saved to examine how AI redistributes nursing work toward review and verification, whether tools meet end user–defined use thresholds, whether performance is equitable across linguistic and workforce groups,

XSL•FO
RenderX

and how automation can be integrated without undermining professional values or relational care [42,43]. Even accurate prompts may create new risks if clinicians overtrust them, ignore nonhighlighted issues, experience alert fatigue, or redirect attention away from direct patient/carer engagement. Human factors testing should therefore assess reliance, trust calibration, interruption burden, and the usability of source-linked evidence before deployment [30,36,44].

Limitations

The synthetic conversational handovers used in this study may not fully represent those performed in actual clinical practice where interruptions, time pressure, environmental noise, nonverbal cues, and local team dynamics can affect what is communicated and how it is interpreted. External validation on real handover audio/transcripts from intended deployment settings is therefore required [33,35,44]. External validity may be particularly limited for complex areas such as mental health, multimorbidity, and other contexts where clinical risk is often communicated through narrative nuance, formulation, behavioral change, staff concern, or tone of interaction rather than discrete data points. These contexts strengthen the need for AI outputs to remain clinician-reviewed, source-linked supports rather than autonomous clinical documentation.

It should be noted that several checklist items had low prevalence in the evaluation partition. Future evaluation should oversample or otherwise specifically test these items [26,31]. Furthermore, the reference standard retained only annotations agreed by both reviewers. This creates a conservative and reproducible benchmark but may exclude ambiguous or contested communication.

The evaluation corpus included synthetic transcripts generated in part with GPT-5. Although the uses were distinct, information extraction tasks were performed on synthesized language from a model from the same LLM provider, which is potentially susceptible to model familiarity. Synthetic transcripts may also be more predictable than authentic handover communication. In addition, the highest-performing models evaluated in this study were hosted proprietary systems that are updated and managed externally. Model updates, infrastructure variability, or configuration changes could alter performance over time in ways that are not fully transparent or controllable, with implications for reproducibility, governance, and consistency in clinical settings where reliability is critical. Locally hosted models may offer greater control over model versioning and deployment conditions but can require substantial compute and may be less responsive to rapid improvements in frontier model capability.

Our evaluation focused on downstream extraction from transcripts and did not separately quantify transcription accuracy, speaker attribution, or the effect of noisy audio. In a real-time handover system, errors introduced before the LLM stage could alter checklist detection, SBAR span extraction, and uncertainty identification. End-to-end evaluation should therefore include audio capture, transcription, diarization, and transcript-to-output performance under realistic clinical conditions [2,44]. We used a checklist that was developed at a local institution based on local priorities for communication during clinical nursing handover. Other wards, specialties, transfer types, or jurisdictions may prioritize different minimum datasets or use different terminology. Implementation would therefore require local mapping of labels to existing handover policy, audit tools, escalation pathways, and documentation workflows, followed by local validation rather than direct transfer of the reported performance estimates [26,34,44].

Conclusions

In this evaluation of AI-assisted clinical handover tasks, prompt optimization consistently improved matched model performance. Accuracy was the highest for bounded checklist prediction and SBAR span extraction. Broad uncertainty detection remained insufficiently reliable, while explicit unknown-fact extraction suggests a narrower but important role for prompting clarification when missing knowledge is directly expressed. Before clinical deployment, these approaches require external validation on real handover audio and transcripts; end-to-end testing of transcription and diarization effects; local calibration of checklist definitions, human-factors evaluation of trust, reliance, workflow burden, and downstream safety outcomes; and structured assessment of organizational readiness for AI implementation in nursing care [45].

Acknowledgments

The authors used OpenAI ChatGPT/Codex during manuscript preparation to assist with editing, formatting, and generation of tables and figures to present results. The authors reviewed, verified, and approved all analytic decisions, references, and final manuscript text.

Funding

This work was supported by the Prince Charles Hospital Foundation Collaboration Grant. The funder had no role in study design, data collection, analysis, interpretation, manuscript preparation, and the decision to submit.

Data Availability

The synthetic handover transcripts, annotation schema, evaluation outputs, and analysis code are available in a public repository on figshare [46]. The NICTA Synthetic Nursing Handover Dataset is publicly available from its original source.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: AC, AT
Data curation: AC
Formal analysis: AC
Investigation: AC, AH, JS
Methodology: AC, AH, AT, DL, KD
Project administration: AC
Software: AC
Visualization: AC
Writing – original draft: AC
Writing – review and editing: AC, AH, JS, GX, DL, TM, KD, AT

Conflicts of Interest

None declared.

Checklist 1
TRIPOD-LLM checklist.
[PDF File, 93 KB - nursing_v9i1e106133_app1.pdf]

References

1. Further information on communicating for safety. Australian Commission on Safety and Quality in Health Care. 2026. URL: https://www.safetyandquality.gov.au/national-standards/nsqhs-standards/communicating-safety-standard/further-information-communicating-safety [accessed 2026-09-22]
2. Ong MS, Coiera E. A systematic review of failures in handoff communication during intrahospital transfers. Jt Comm J Qual Patient Saf 2011 Jun;37(6):274-284. [doi: 10.1016/s1553-7250(11)37035-3] [Medline: 21706987]
3. Manias E, Geddes F, Watson B, Jones D, Della P. Perspectives of clinical handover processes: a multi-site survey across different health professionals. J Clin Nurs 2016 Jan;25(1-2):80-91. [doi: 10.1111/jocn.12986] [Medline: 26415923]
4. The Joint Commission, and Joint Commission International. Communication during patient hand-overs. : World Health Organization; 2007 URL: https://cdn.who.int/media/docs/default-source/patient-safety/patient-safety-solutions/ps-solution3-communication-during-patient-handovers.pdf [accessed 2026-09-22]
5. Hada A, Jones LV, Jack LC, Coyer F. Translating evidence-based nursing clinical handover practice in an acute care setting: a quasi-experimental study. Nurs Health Sci 2021 Jun;23(2):466-476. [doi: 10.1111/nhs.12836] [Medline: 33797197]
6. Yung AHW, Pak CS, Watson B. A scoping review of clinical handover mnemonic devices. Int J Qual Health Care 2023 Sep 8;35(3):mzad065. [doi: 10.1093/intqhc/mzad065] [Medline: 37616494]
7. McCarthy S, Motala A, Lawson E, Shekelle PG. Use of structured handoff protocols for within-hospital unit transitions: a systematic review from Making Healthcare Safer IV. BMJ Qual Saf 2025 Sep 18;34(10):680-690. [doi: 10.1136/bmjqs-2024-018385] [Medline: 40306923]
8. Moyo P, Anderson J, Francis K, Biles J. Exploring the experiences and perceptions of the utilisation of structured clinical handover frameworks by nurses working in acute care settings: a scoping review. J Clin Nurs 2024 Nov;33(11):4297-4313. [doi: 10.1111/jocn.17430] [Medline: 39287216]
9. Chien LJ, Slade D, Goncharov L, et al. Implementing a ward-level intervention to improve nursing handover communication with a focus on bedside handover-a qualitative study. J Clin Nurs 2024 Jul;33(7):2688-2706. [doi: 10.1111/jocn.17107] [Medline: 38528438]
10. Tobiano G, Bucknall T, Sladdin I, Whitty JA, Chaboyer W. Patient participation in nursing bedside handover: a systematic mixed-methods review. Int J Nurs Stud 2018 Jan;77:243-258. [doi: 10.1016/j.ijnurstu.2017.10.014] [Medline: 29149634]
11. Anshasi H, Almayasi ZA. Perceptions of patients and nurses about bedside nursing handover: a qualitative systematic review and meta-synthesis. Nurs Res Pract 2024;2024:3208747. [doi: 10.1155/2024/3208747] [Medline: 38716049]
12. Agrawal M, Hegselmann S, Lang H, Kim Y, Sontag D. Large language models are few-shot clinical information extractors. Proc 2022 Conf Empir Methods Nat Lang Process 2022:1998-2022. [doi: 10.18653/v1/2022.emnlp-main.130]
13. Sivarajkumar S, Kelley M, Samolyk-Mazzanti A, Visweswaran S, Wang Y. An empirical evaluation of prompting strategies for large language models in zero-shot clinical natural language processing: algorithm development and validation study. JMIR Med Inform 2024 Apr 8;12:e55318. [doi: 10.2196/55318] [Medline: 38587879]
14. Agha-Mir-Salim L, Alberto IR, Alberto NR, et al. Technological solutions to improve inpatient handover in the era of artificial intelligence: scoping review. J Med Internet Res 2025 Jul 31;27(July):e70358. [doi: 10.2196/70358] [Medline: 40743446]
15. Chen RJ, Wu MS, Tsai LW, Chang SS, Shen Hsiao ST, Lo YS. Integrating a large language model to streamline nursing handover documentation across multiple hospitals in Taiwan: development and implementation study. J Med Internet Res 2026 Mar 12;28:e81604. [doi: 10.2196/81604] [Medline: 41819121]

XSL•FO
RenderX

16. Gallifant J, Afshar M, Ameen S, et al. The TRIPOD-LLM reporting guideline for studies using large language models. Nat Med 2025 Jan;31(1):60-69. [doi: 10.1038/s41591-024-03425-5] [Medline: 39779929]
17. Suominen H, Zhou L, Hanlen L, Ferraro G. Benchmarking clinical speech recognition and information extraction: new data, methods, and evaluations. JMIR Med Inform 2015 Apr 27;3(2):e19. [doi: 10.2196/medinform.4321] [Medline: 25917752]
18. Khattab O, Singhvi A, Maheshwari P, et al. DSPy: compiling declarative language model calls into self-improving pipelines. Presented at: The 12th International Conference on Learning Representations (ICLR 2024); May 7-11, 2024 URL: https://openreview.net/pdf?id=sY5N0zY5Od [accessed 2026-09-22]
19. Australian Research Council (ARC); Universities Australia. National statement on ethical conduct in human research 2025. : National Health and Medical Research Council (NHMRC); 2025 URL: https://www.nhmrc.gov.au/sites/default/files/documents/attachments/publications/National-Statement-on-Ethical-Conduct-Human-Research-2025.pdf [accessed 2026-09-22]
20. Prodigy. URL: https://prodi.gy [accessed 2026-09-22]
21. Singh A, Fry A, Perelman A, et al. OpenAI GPT-5 system card. arXiv. Preprint posted online on Dec 19, 2025. [doi: 10.48550/arXiv.2601.03267]
22. Sellergren A, Kazemzadeh S, Jaroensri T, et al. MedGemma technical report. arXiv. Preprint posted online on Jul 7, 2025. [doi: 10.48550/arXiv.2507.05201]
23. Agrawal LA, Tan S, Soylu D, et al. GEPA: reflective prompt evolution can outperform reinforcement learning. Presented at: 14th International Conference on Learning Representations (ICLR 2026); Apr 23-27, 2026 URL: https://openreview.net/pdf?id=RQm2KQTM5r [accessed 2026-09-22]
24. Goel A. LangExtract v1.7.0. Zenodo. 2026. URL: https://doi.org/10.5281/zenodo.17015089 [accessed 2026-09-22]
25. Haig KM, Sutton S, Whittington J. SBAR: a shared mental model for improving communication between clinicians. Jt Comm J Qual Patient Saf 2006 Mar;32(3):167-175. [doi: 10.1016/s1553-7250(06)32022-3] [Medline: 16617948]
26. Riesenberg LA, Leitzsch J, Cunningham JM. Nursing handoffs: a systematic review of the literature. Am J Nurs 2010 Apr;110(4):24-34. [doi: 10.1097/01.NAJ.0000370154.79857.09] [Medline: 20335686]
27. Bukoh MX, Siah CJR. A systematic review on the structured handover interventions between nurses in improving patient safety outcomes. J Nurs Manag 2020 Apr;28(3):744-755. [doi: 10.1111/jonm.12936] [Medline: 31859377]
28. Manser T, Foster S. Effective handover communication: an overview of research and improvement efforts. Best Pract Res Clin Anaesthesiol 2011 Jun;25(2):181-191. [doi: 10.1016/j.bpa.2011.02.006] [Medline: 21550543]
29. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ 2005 Apr 2;330(7494):765. [doi: 10.1136/bmj.38398.500764.8F] [Medline: 15767266]
30. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc 2012;19(1):121-127. [doi: 10.1136/amiajnl-2011-000089] [Medline: 21685142]
31. Challen R, Denny J, Pitt M, Gompels L, Edwards T, Tsaneva-Atanasova K. Artificial intelligence, bias and clinical safety. BMJ Qual Saf 2019 Mar;28(3):231-237. [doi: 10.1136/bmjqs-2018-008370] [Medline: 30636200]
32. Rosenbloom ST, Denny JC, Xu H, Lorenzi N, Stead WW, Johnson KB. Data from clinical notes: a perspective on the tension between structure and flexible documentation. J Am Med Inform Assoc 2011;18(2):181-186. [doi: 10.1136/jamia.2010.007237] [Medline: 21233086]
33. Cohen MD, Hilligoss PB. The published literature on handoffs in hospitals: deficiencies identified in an extensive review. Qual Saf Health Care 2010 Dec;19(6):493-497. [doi: 10.1136/qshc.2009.033480] [Medline: 20378628]
34. Redley B, Waugh R. Mixed methods evaluation of a quality improvement and audit tool for nurse-to-nurse bedside clinical handover in ward settings. Appl Nurs Res 2018 Apr;40(April):80-89. [doi: 10.1016/j.apnr.2017.12.013] [Medline: 29579504]
35. Patterson ES, Roth EM, Woods DD, Chow R, Gomes JO. Handoff strategies in settings with high consequences for failure: lessons for health care operations. Int J Qual Health Care 2004 Apr;16(2):125-132. [doi: 10.1093/intqhc/mzh026] [Medline: 15051706]
36. Leonard M, Graham S, Bonacum D. The human factor: the critical importance of effective teamwork and communication in providing safe care. Qual Saf Health Care 2004 Oct;13(Suppl 1):i85-i90. [doi: 10.1136/qhc.13.suppl_1.i85] [Medline: 15465961]
37. Starmer AJ, Spector ND, Srivastava R, et al. Changes in medical errors after implementation of a handoff program. N Engl J Med 2014 Nov 6;371(19):1803-1812. [doi: 10.1056/NEJMsa1405556] [Medline: 25372088]
38. Figueroa RL, Zeng-Treitler Q, Ngo LH, Goryachev S, Wiechmann EP. Active learning for clinical text classification: is it better than random sampling? J Am Med Inform Assoc 2012;19(5):809-816. [doi: 10.1136/amiajnl-2011-000648] [Medline: 22707743]
39. Chen Y, Lasko TA, Mei Q, Denny JC, Xu H. A study of active learning methods for named entity recognition in clinical text. J Biomed Inform 2015 Dec;58:11-18. [doi: 10.1016/j.jbi.2015.09.010] [Medline: 26385377]
40. Feng J, Phillips RV, Malenica I, et al. Clinical artificial intelligence quality improvement: towards continual monitoring and updating of AI algorithms in healthcare. NPJ Digit Med 2022 May 31;5(1):66. [doi: 10.1038/s41746-022-00611-y] [Medline: 35641814]

XSL•FO
RenderX

41. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med 2019 Oct 29;17(1):195. [doi: 10.1186/s12916-019-1426-2] [Medline: 31665002]
42. Ronquillo CE. Beyond time saved: implementation, equity, and the utility threshold for nursing AI scribes. J Med Internet Res 2026 May 27;28:e101190. [doi: 10.2196/101190] [Medline: 42202251]
43. Pepito JA, Acaso NJ, Merioles R, Ismael J. Opportunities, challenges, and future directions for the integration of automation in nursing practice: discursive study. JMIR Nurs 2025 Aug 14;8:e72674. [doi: 10.2196/72674] [Medline: 40811767]
44. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. Qual Saf Health Care 2010 Oct;19 Suppl 3(Suppl 3):i68-i74. [doi: 10.1136/qshc.2010.042085] [Medline: 20959322]
45. Seibert K, Domhoff D, Altona J, et al. Readiness assessment for AI in nursing care projects: multimethods study. JMIR Nurs 2026 Jun 2;9:e84148. [doi: 10.2196/84148] [Medline: 42228895]
46. Large language models for structured information extraction in artificial intelligence-assisted clinical handover: evaluation study. Figshare. URL: https://figshare.com/articles/dataset/Large_Language_Models_for_Structured_Information_Extraction_in_Artificial_Intelligence-Assisted_Clinical_Handover_Evaluation_Study/32658084?file=65497158 [accessed 2026-09-22]

Abbreviations

GEPA : Genetic-Pareto
IoU : intersection over union
LLM : large language model
NICTA : National Information and Communications Technology Australia
SBAR : Situation, Background, Assessment, Recommendation
TRIPOD-LLM: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis–Large Language Models

Edited by E Borycki; submitted 02.Jul.2026; peer-reviewed by B Galatzan, C Zhang; revised version received 03.Aug.2026; accepted 07.Sep.2026; published 05.Oct.2026.

Please cite as:
Conway A, Hada A, Schluter J, Xu H(, Lowden D, Miller T, Donald K, Teodorczuk A
AI-Based Structured Information Extraction From Synthetic Nursing Handover Transcripts: Comparative Evaluation of Large Language Models
JMIR Nursing 2026;9:e106133
URL: https://nursing.jmir.org/2026/1/e106133
doi:10.2196/106133

© Aaron Conway, Adriana Hada, Jessica Schluter, Hui (Grace) Xu, Dan Lowden, Tim Miller, Ken Donald, Andrew Teodorczuk. Originally published in JMIR Nursing (https://nursing.jmir.org), 5.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

From Academic Integrity to Professional Formation in Generative AI–Supported Nursing Education: Integrative Review

Peter Hartin*, PhD; Paul Baretta*, MN

College of Healthcare Sciences, James Cook University, 1 James Cook Drive, Douglas, Queensland, Australia

*all authors contributed equally

Corresponding Author:
Peter Hartin, PhD
College of Healthcare Sciences, James Cook University, 1 James Cook Drive, Douglas, Queensland, Australia

Abstract

Background: The rapid integration of generative AI (GenAI) into nursing education has prompted widespread concern about academic integrity and assessment validity. However, this dominant discourse may obscure a more fundamental challenge: preparing nursing students for clinical practice environments where AI already mediates triage, diagnostics, documentation, and clinical decision-making.

Objective: This integrative review examines how GenAI is conceptualized within the nursing education literature and investigates the extent to which current discourse addresses professional formation rather than academic compliance.

Methods: Following Whittemore and Knafl's integrative review methodology, a systematic search of PubMed, Google Scholar, and SciSpace (Typeset) was conducted for literature published between January 2022 and February 2026. Two reviewers independently screened 456 records, with 119 papers meeting the inclusion criteria. The methodological quality of empirical studies was assessed using the Mixed Methods Appraisal Tool (MMAT). Corpus-level descriptive mapping quantified dominant themes and conceptual patterns, while an interpretive synthesis examined how GenAI was framed and the extent to which professional formation elements, including identity, moral agency, and clinical judgment, were addressed.

Results: Of the 119 unique studies analyzed, 84% (n=100) predominantly framed GenAI as a content-generation tool and academic integrity threat. Only 35.3% (n=42) conceptualized AI as a pedagogical tool, and a critical gap was identified in professional formation: identity formation was addressed in only 10% (n=12) of papers and moral agency in just 5% (n=6). Explicit acknowledgment that AI already routinely mediates frontline clinical decisions (eg, sepsis alerts and fall risk algorithms) in current practice settings was notably absent. The discourse emphasizes governance and faculty literacy, but sustained curricula addressing nurse identity in AI-mediated care remain underdeveloped.

Conclusions: Current nursing education discourse predominantly frames AI as a writing tool and an academic integrity concern, failing to prepare students for the algorithmic practice environments they will inhabit. Nursing education must reframe AI as a professional formation issue, explicitly addressing how algorithms reshape clinical judgment, moral agency, and professional identity.

(JMIR Nursing 2026;9:e99284) doi:10.2196/99284

KEYWORDS

AI; generative AI; nursing education; professional formation; academic integrity; clinical judgment; algorithmic practice; student identity

Introduction

The rapid emergence of generative AI (GenAI), particularly large language models such as ChatGPT (OpenAI), has prompted one of the most significant pedagogical disruptions experienced by higher education in recent decades. Since ChatGPT became publicly available in late 2022, nurse educators have grappled with questions surrounding academic integrity, assessment validity, authorship, and the appropriate boundaries of AI use in student learning [1,2]. These concerns have generated a rapidly expanding body of literature examining the educational implications of GenAI and have prompted widespread institutional responses, including revised assessment practices, AI policies, and faculty development initiatives [3-5].

Alongside these educational developments, health care itself is undergoing rapid technological transformation. AI is increasingly embedded within contemporary clinical practice through technologies such as clinical decision support systems, predictive analytics, automated documentation, risk stratification algorithms, and workflow optimization tools. Unlike GenAI, which generates novel content in response to user prompts, clinical decision support systems primarily support clinicians' decision-making using predefined rules, predictive models, or

XSL•FO
RenderX

machine learning algorithms embedded within health care workflows [6,7]. Consequently, graduates entering the nursing workforce will not simply encounter AI as an educational technology but will practice within health care environments where algorithmic systems increasingly influence clinical reasoning, decision-making, and patient care. Preparing nurses for this reality requires educational approaches that extend beyond managing AI use during assessment to considering how future practitioners develop professional judgment, ethical reasoning, and accountability within algorithmically mediated practice.

Recent reviews have made important contributions to understanding GenAI in nursing education by examining issues such as academic integrity, ethical considerations, educational applications, faculty preparedness, student perceptions, and pedagogical innovation [3,8-11]. Collectively, this work has established a valuable evidence base describing how GenAI is being adopted, the opportunities it presents, and the challenges it poses for teaching and learning. However, these reviews have largely conceptualized AI as an educational tool or educational problem. Comparatively little attention has been directed toward how the literature frames the broader educational purpose of preparing nurses for professional practice in health care systems where AI is already integrated into routine clinical care.

This distinction is important because educational discourse is not neutral. The ways in which emerging technologies are framed influence curriculum priorities, assessment practices, faculty development, and ultimately the capabilities that nursing programs seek to cultivate. When AI is positioned primarily as a threat to academic integrity, educational responses naturally focus on governance, surveillance, and assessment redesign. While these responses are appropriate and necessary, they risk narrowing the educational conversation by privileging academic compliance over the broader question of professional formation. Understanding how the literature conceptualizes the educational challenge posed by GenAI is therefore essential, as these conceptualizations shape what is valued, taught, and assessed within nursing education.

Professional identity formation provides a useful theoretical lens through which to examine this issue. Professional identity formation describes the developmental process through which learners progressively integrate the values, behaviors, knowledge, and responsibilities of the profession into their sense of self, ultimately learning to “think, act, and feel” like nurses [12,13]. In nursing, this process encompasses the development of clinical judgment, ethical reasoning, patient advocacy, relational care, accountability, and moral agency [14]. As AI becomes increasingly embedded within health care delivery, the contexts in which professional identity develops are also changing. Consequently, understanding how nursing education conceptualizes AI extends beyond questions of technology adoption to encompass how future nurses are prepared to exercise judgment, maintain professional responsibility, and provide person-centered care alongside increasingly sophisticated algorithmic systems.

Against this background, this review adopts a different perspective from previous syntheses. Rather than asking how GenAI is being used in nursing education or whether its use should be encouraged or restricted, this review examines how the nursing education literature conceptualizes the educational challenge itself. Specifically, it explores whether contemporary discourse remains predominantly centered on academic integrity and educational governance or whether it also addresses professional formation for algorithmically mediated clinical practice.

The review was guided by the following research question: how is GenAI conceptualized within the nursing education literature and to what extent does the current discourse address professional formation?

Professional identity formation therefore served as the conceptual lens underpinning the review, informing both the analysis of the literature and the interpretation of how nursing education currently positions students as future professionals within increasingly AI-enabled health care environments.

Methods

Design

This study used an integrative review methodology guided by Whittemore and Knafl’s [15] framework, which supports the systematic synthesis of diverse forms of evidence, including empirical research, reviews, theoretical papers, commentaries, editorials, and professional guidance documents. This approach was selected because scholarship related to GenAI in nursing education remains methodologically heterogeneous and conceptually emergent, making conventional systematic review approaches that focus exclusively on empirical effectiveness studies insufficient to address the research question.

Unlike reviews designed to evaluate intervention effectiveness, this review sought to examine how GenAI is conceptualized across contemporary nursing education literature. Accordingly, the review combined systematic literature identification with qualitative interpretive synthesis to explore the dominant narratives, assumptions, and conceptual framings underpinning current scholarship.

Conceptual Orientation

Professional identity formation served as the conceptual lens underpinning the review. Rather than cataloging educational applications of GenAI, this review examined how nursing education scholarship constructs the educational challenges associated with AI and the extent to which these constructions address preparation for algorithmically mediated clinical practice.

This orientation recognizes that literature not only reports evidence but also shapes professional priorities by privileging particular educational concerns while rendering others less visible. Consequently, the review adopted a discourse-oriented approach, seeking to identify both dominant conceptual framings and important conceptual absences within the literature.

Search Strategy

A systematic literature search was undertaken to identify publications examining GenAI within nursing education.

XSL•FO
RenderX

Searches were conducted across PubMed, Google Scholar, and SciSpace (Typeset), capturing literature published between January 1, 2022, and February 28, 2026. The search period commenced immediately prior to the public release of ChatGPT to capture the emergence of contemporary GenAI scholarship.

The database selection reflected complementary strengths. PubMed was selected because it comprehensively indexes biomedical and health professions literature. Google Scholar provides broad scholarly coverage across disciplines, while SciSpace facilitates the retrieval of recently published and emerging AI-related scholarship in this rapidly evolving field.

The search terms combined controlled vocabulary (where available) and keywords relating to GenAI, nursing education, and professional formation. Boolean operators were used to combine concepts, with database-specific adaptations applied as required. The complete search strategies for all databases are presented in Multimedia Appendix 1.

Following database searches, the reference lists of included reviews and key papers were manually examined to identify additional eligible studies not retrieved through electronic searches. Textbox 1 summarizes the eligibility criteria.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria:

- Papers addressing generative AI in nursing education contexts
- Published or preprinted between 2022 and 2026
- Focus on undergraduate or graduate nursing education
- Empirical studies, theoretical papers, position statements, or reviews

Exclusion criteria:

- Papers focusing solely on nongenerative AI (eg, predictive analytics and diagnostic imaging AI)
- Papers addressing AI in clinical practice without educational implications
- Conference abstracts without full text
- Duplicate publications

Study Selection

All search results were imported into a reference management program, and duplicate records were removed prior to screening.

Study selection followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 framework (Checklist 1). Titles and abstracts were independently screened by 2 reviewers against the predefined eligibility criteria. Potentially eligible studies subsequently underwent full-text assessment.

Disagreements at either stage were resolved through discussion until consensus was reached. Where uncertainty remained regarding eligibility, studies were retained for full-text review to maximize sensitivity during screening.

The study selection process is presented in the PRISMA flow diagram in the *Results* section.

Quality Appraisal

Methodological quality was assessed for all empirical studies using the Mixed Methods Appraisal Tool (MMAT), which allows consistent appraisal across qualitative, quantitative, and mixed methods research designs.

Because this review also included reviews, theoretical papers, editorials, and professional commentaries, these publications were appraised according to their conceptual relevance, clarity of argument, and contribution to the understanding of GenAI within nursing education rather than being evaluated with empirical quality criteria.

Consistent with integrative review methodology, studies were not excluded on the basis of methodological quality. Instead, quality appraisal informed the interpretation of the findings by providing greater weight to higher-quality empirical evidence while recognizing the important conceptual contributions of theoretical literature.

Data Extraction

A structured data extraction framework was developed prior to the analysis and was applied consistently across all included publications.

For each study, the following information was extracted: bibliographic details, year of publication, country of origin, publication type, study design, educational context, primary focus of the AI application, principal findings, and relevance to professional formation.

To support the review's conceptual focus, each paper was additionally coded for the presence or absence of 7 predefined professional formation constructs derived from the literature. Coding was based on explicit discussion of each construct within the publication. Where constructs were not directly named, studies were coded as present only when the underlying concept was clearly described or operationalized in relation to nursing education. The predefined professional formation constructs

XSL•FO
RenderX

comprised professional identity formation, ethical reasoning, moral agency, clinical judgment, digital professionalism, algorithmic bias and power awareness, and preparation for algorithmically mediated clinical practice.

Studies were further coded according to publication characteristics, dominant AI framing, educational focus, and the presence of the predefined professional formation constructs. Descriptive coding captured publication characteristics and educational context, while conceptual coding identified how GenAI was framed within the literature and whether professional formation constructs were explicitly or implicitly represented. Coding categories were initially informed deductively by the professional identity formation literature and subsequently refined through iterative familiarization and engagement with the corpus.

The complete extraction framework is presented in Multimedia Appendix 2.

Data Analysis

Analysis occurred in 2 complementary phases.

Phase 1: Corpus-Level Descriptive Mapping

The first phase involved corpus-level descriptive mapping across all included publications.

Descriptive analyses quantified publication trends, dominant conceptualizations of GenAI, and the frequency with which professional formation constructs appeared across the literature. These analyses provided an overview of the prevailing discourse within contemporary nursing education scholarship.

Phase 2: Interpretive Synthesis

The second phase used interpretive synthesis to examine how the identified patterns collectively constructed the discourse surrounding GenAI within nursing education.

Rather than synthesizing findings study-by-study, the interpretive analysis examined relationships across the corpus to identify dominant narratives, underlying assumptions, conceptual tensions, and important omissions. Constant comparison was used throughout the analysis, with emerging interpretations repeatedly tested against the complete dataset to ensure that they adequately represented the breadth of the included literature.

Professional identity formation served as the interpretive framework through which dominant conceptualizations were examined, enabling the identification of both explicit discussions and notable silences relating to identity development, moral agency, clinical judgment, and preparation for algorithmically mediated practice.

The integration of corpus-level descriptive mapping with interpretive synthesis enabled the review to move beyond describing what has been published toward understanding how the field currently conceptualizes the educational implications of GenAI.

To enhance analytical rigor, all included publications were independently coded by 2 reviewers using the predefined coding framework. Coding decisions were subsequently compared, with any discrepancies discussed and resolved through consensus. Throughout the analytical process, category definitions and interpretations were iteratively refined to ensure consistency across the corpus while remaining grounded in the included literature. Given the interpretive nature of this integrative review, formal intercoder reliability statistics were not calculated. Instead, analytical trustworthiness was enhanced through independent coding, iterative discussion, and consensus-based refinement of the coding framework, consistent with established integrative review methodology.

Results

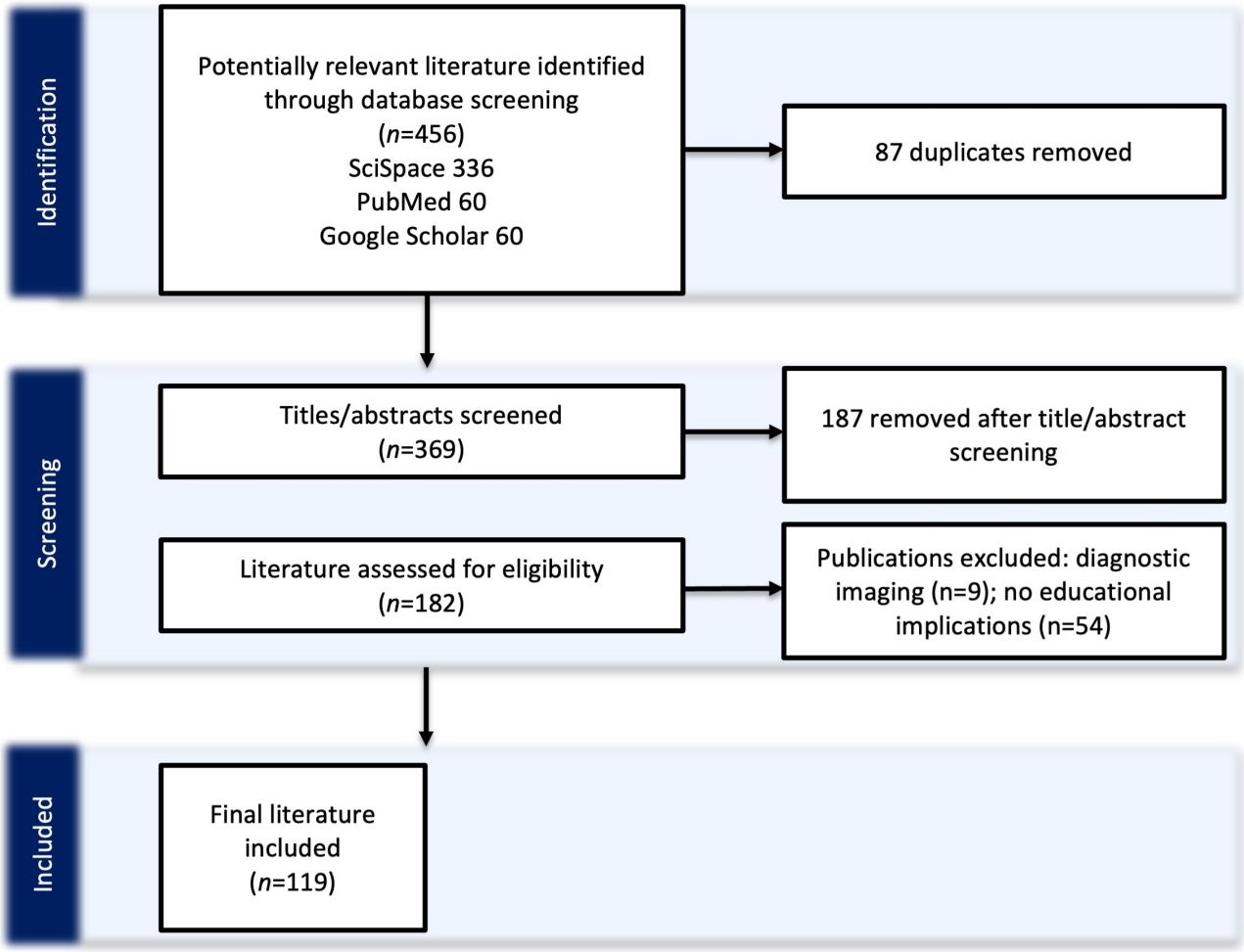
Study Characteristics

Overview of the Included Literature

The study selection process is presented in Figure 1. Database searches identified 456 records, of which 87 duplicates were removed. Following title and abstract screening, 250 records were excluded, leaving 119 publications that met the eligibility criteria and were included in the review.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.



Of the 119 publications, 64 (53.8%) were empirical studies, 31 (26.1%) were scoping or systematic reviews, 13 (10.9%) were position papers or commentaries, and 11 (9.2%) were theoretical or framework papers. Publications originated from multiple geographical regions, with the strongest representation from the United States, China, Australia, and Europe. Publication volume increased rapidly over the review period, rising from 8 publications during 2022 - 2023 to 43 in 2024 and 68 in 2025 - 2026, reflecting the rapid expansion of scholarly interest following the public release of GenAI technologies (Table 1).

Table . Characteristics of included studies (n=119).

Characteristic	Values, n (%)
Empirical studies	64 (53.8)
Scoping/systematic reviews	31 (26.1)
Position papers/commentaries	13 (10.9)
Theoretical/framework papers	11 (9.2)
Publications by year	
2022 - 2023	8 (6.7)
2024	43 (36.1)
2025 - 2026	68 (57.1)

Rather than presenting isolated themes, the findings are organized around the dominant ways in which GenAI is conceptualized within the nursing education literature. Corpus-level mapping identified recurring conceptual framings across the included publications, while interpretive synthesis examined how these framings positioned students, educators, and professional preparation for algorithmically mediated practice. Figure 2 presents the resulting conceptual model that informed the thematic synthesis.

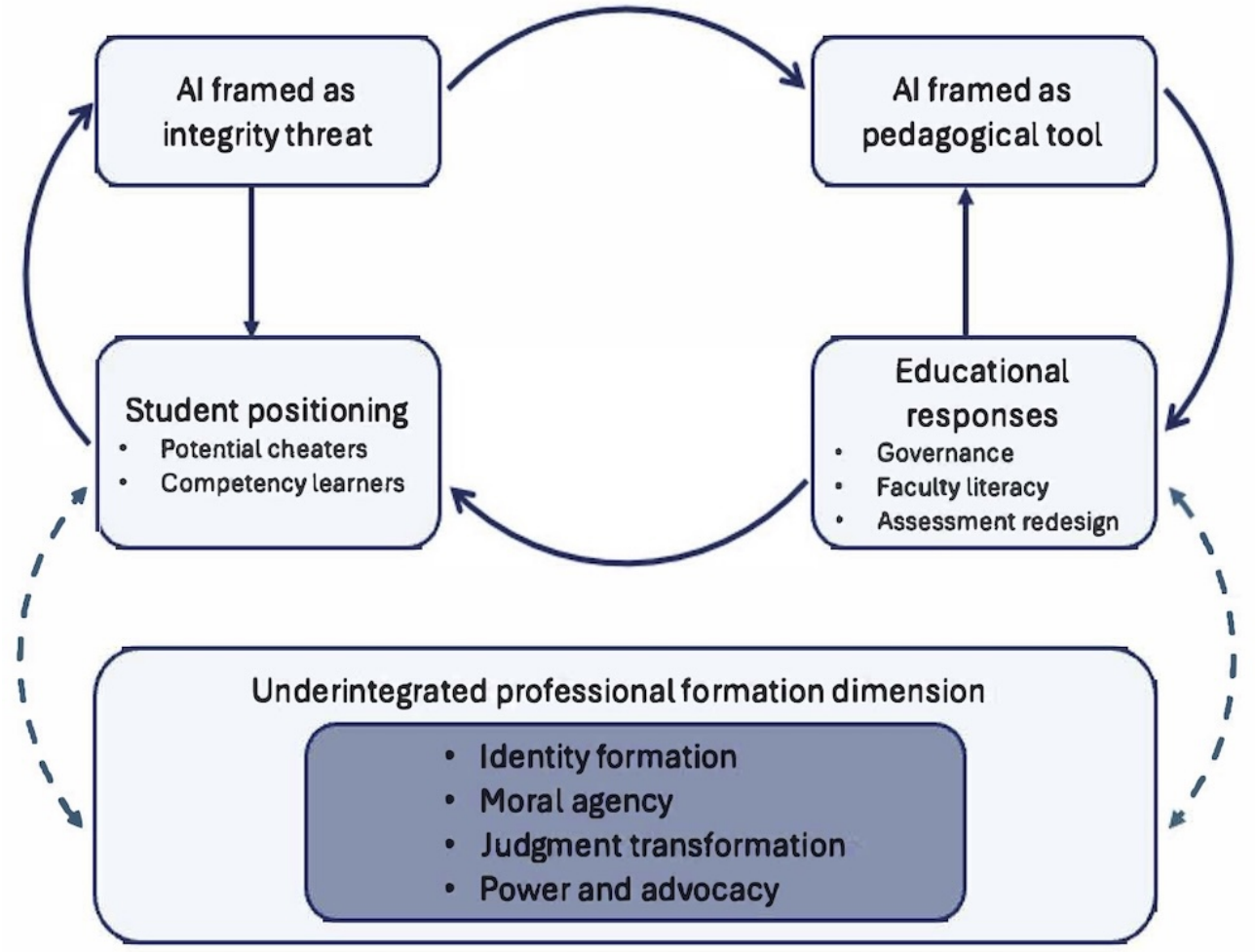
The included literature comprised empirical studies (n=64), scoping and systematic reviews (n=31), position papers and commentaries (n=13), and theoretical or framework papers (n=11). Geographic representation was global, with significant

XSL•FO
RenderX

contributions from the United States, China, Australia, and Europe. Publication frequency increased sharply over time, with 8 papers in 2022 - 2023, 43 in 2024, and 68 in 2025 - 2026, reflecting rapid scholarly response following the public release of GenAI technologies.

Figure 2 presents a conceptual model derived from the integration of corpus-level mapping and interpretive synthesis. The model illustrates a governance-pedagogy cycle anchored in integrity discourse, wherein AI is framed primarily as an integrity threat or a pedagogical instrument, generating governance and literacy-based responses. Professional formation in algorithmically mediated practice remains comparatively underintegrated within this cycle. The themes that follow elaborate on each component of this discourse cycle.

Figure 2. Conceptual model of generative AI discourse in nursing education. Synthesis revealed a dominant tool-oriented discourse cycle and comparatively limited integration of professional formation in algorithmically mediated practice.



Theme 1: AI as a Writing Tool and an Integrity Threat

The dominant discourse positioned GenAI primarily as a content-generation tool with implications for academic integrity, assessment validity, and authorship authenticity. Across the corpus, AI was most frequently discussed in relation to plagiarism, inappropriate assistance with assessments, and concerns regarding the integrity of student work, rather than its influence on professional preparation or future clinical practice.

Shorey et al [3] described ChatGPT as raising concerns regarding plagiarism, academic dishonesty, and the authenticity of student work, while Siah et al [4] similarly framed authenticity and academic integrity as the central educational challenges requiring investigation. Across the literature, 3 recurring subthemes characterized this integrity-focused discourse.

The first concerned assessment validity. Many publications questioned whether AI-generated work undermined assessment as a measure of student learning and clinical competence, often equating undisclosed AI use with plagiarism or other forms of academic misconduct [3-5]. The second focused on detection and governance, with considerable attention devoted to AI detection software, institutional policies, disclosure requirements, and faculty responsibilities for regulating AI use [5,16]. Finally, numerous publications advocated assessment redesign, recommending oral examinations, in-person assessments, authentic assessment tasks, and reflective activities that were perceived to be less vulnerable to AI-assisted completion [3,4,17].

Collectively, these studies positioned academic integrity as the principal educational problem associated with GenAI. Although this framing acknowledged legitimate concerns regarding

XSL•FO
RenderX

assessment authenticity, professional formation received comparatively limited attention within this dominant discourse.

Theme 2: AI as a Pedagogical Tool

Alongside the dominant integrity discourse, a smaller but growing body of literature conceptualized GenAI as a pedagogical resource capable of enhancing learning rather than threatening assessment integrity. These publications shifted attention from regulation and surveillance toward constructive educational integration.

Several pedagogical applications recurred throughout the corpus. AI was described as supporting clinical reasoning through the generation of patient scenarios, care plans, and differential diagnoses that students were required to critically evaluate [18-20]. Other studies emphasized personalized learning, describing AI systems that provided tailored feedback, adaptive learning opportunities, and conversational support for evidence-based practice [21,22]. Additional publications highlighted opportunities to improve accessibility, particularly for students with language barriers or diverse learning needs, while others positioned AI literacy itself as an emerging professional capability required for contemporary nursing practice [1,21,22].

Despite representing a more constructive framing than the integrity-focused literature, this discourse remained predominantly educational in orientation. AI was generally conceptualized as a teaching and learning resource rather than as a feature of contemporary clinical environments requiring explicit preparation for professional practice.

Theme 3: The Missing Discourse on Professional Formation in Algorithmic Practice

While integrity and pedagogy dominated the literature, a different pattern also emerged. Across the corpus, comparatively little attention was directed toward how AI might reshape the formation of nurses as professionals practicing within algorithmically mediated health care systems. Seven dimensions of professional formation were examined during the analysis, including identity formation, ethical reasoning, moral agency, clinical judgment, digital professionalism, awareness of algorithmic power, and recognition of AI within contemporary clinical practice.

Identity Formation, Moral Agency, and Clinical Judgment

Professional identity formation was explicitly addressed in only 12 of the 119 (10%) included publications. When identity was discussed, students were typically positioned as learners developing AI-related competencies rather than professionals negotiating evolving clinical roles and responsibilities. References to professional identity were commonly embedded within discussions of digital professionalism, focusing on appropriate AI use and disclosure practices rather than the broader developmental process through which nurses learn to think, act, and feel as members of the profession.

Moral agency was even less visible. Only 6 (5%) publications explicitly discussed nurses' responsibility to question, challenge, or override algorithmic recommendations when patient welfare required it. Although professional accountability was occasionally acknowledged, few studies described educational strategies for cultivating the moral courage needed to exercise independent clinical judgment within AI-supported practice environments.

Clinical judgment appeared more frequently, with 38 (31.9%) publications referring to critical thinking or decision-making. However, these discussions largely focused on evaluating AI-generated outputs rather than examining how algorithmic systems may reshape clinical reasoning itself. AI was commonly positioned as a learning scaffold supporting judgment, rather than as a technology transforming the contexts in which professional judgment is exercised.

Ethical Awareness, Power, and the Algorithmic Present

AI Ethics and Clinical Preparedness

Ethical issues were identified in approximately 38% of the corpus and were most commonly related to algorithmic bias, privacy, transparency, and the reliability of AI-generated information [9,11,23,24]. These discussions generally emphasized vigilance, verification, and technical oversight, with relatively few publications describing structured educational approaches to developing ethical reasoning in AI-supported clinical practice.

Similarly, approximately one-third of publications acknowledged algorithmic bias, inequity, or opacity, yet few translated these concerns into curricular approaches that prepared students to recognize, challenge, or respond to inequitable algorithmic decision-making. Student concerns regarding AI bias were evident in empirical studies, although participants frequently reported limited preparation for managing these issues in practice.

A further finding was the limited recognition that AI is already embedded within many contemporary health care systems. Across much of the literature, AI was discussed as an emerging or future technology despite its established role in early warning systems, clinical decision support, risk stratification, and documentation technologies. Taken together, these findings suggest that while ethical and technological issues receive growing attention, professional formation within algorithmically mediated health care remains comparatively underdeveloped.

Theme 4: Student Identity Construction From Threat to Opportunity

The analysis identified 4 dominant ways nursing students were positioned within the literature. First, students were frequently portrayed as potential academic integrity risks, requiring surveillance and regulation to prevent inappropriate AI use. Second, students were described as digitally capable but requiring guidance, reinforcing the faculty's responsibility for establishing appropriate boundaries and institutional expectations. Third, students were positioned as future professionals developing AI-related competencies, although competency development was generally framed in technical rather than professional terms. Finally, empirical studies portrayed students as pragmatic users seeking to integrate AI

XSL•FO
RenderX

into their learning while navigating uncertainty regarding institutional expectations.

Across these constructions, relatively little attention was directed toward students as professionals-in-formation preparing to practice alongside algorithmic technologies. Consequently, the literature rarely articulated how future nurses might develop professional judgment, moral agency, and relational practice within AI-enabled health care environments.

Theme 5: Proposed Educational Responses Involving Governance, Literacy, and Pedagogy

Three broad categories of educational response emerged across the literature: governance and policy development, faculty AI literacy, and pedagogical innovation. These responses largely reflected the dominant conceptual framings identified throughout the corpus.

Governance-focused publications advocated institutional policies, disclosure requirements, academic integrity training, and faculty guidance to regulate AI use within assessment [5,16,25]. Faculty development initiatives primarily focused on improving educators' technological knowledge, assessment design, and confidence in integrating AI into teaching [1,10,26]. Pedagogical innovations included AI-generated patient scenarios, structured evaluation of AI outputs, personalized feedback systems, and collaborative learning approaches in which students interacted directly with AI technologies [18-22,27,28].

Although these responses demonstrated increasing educational sophistication, they generally remained situated within existing curricular structures. Few publications proposed sustained curriculum transformation centered on professional identity formation, clinical judgment, or moral agency within algorithmically mediated practice environments.

Discussion

Principal Findings

The central finding of this review is a significant misalignment between how GenAI is framed within the nursing education literature and the realities of contemporary clinical practice. Among the 119 publications, the discourse is dominated by concerns about academic integrity (n=100, 84%), whereas comparatively little attention is given to professional identity formation (n=12, 10%) or moral agency (n=6, 5%). This imbalance suggests that nursing education is responding to GenAI primarily as an assessment problem rather than as a transformation of professional practice environments. Interpreting these findings through the lens of professional identity formation highlights how current discourse emphasizes governance and competency while underengaging with the developmental processes through which nurses learn to exercise judgment, agency, and relational care in technologically mediated practice environments.

Across the corpus, GenAI is most frequently conceptualized as a writing tool that threatens authorship authenticity and assessment validity [3-5]. This framing has generated extensive discussion about detection tools, policy frameworks, and assessment redesign [3,4,16]. While these concerns are legitimate, the prominence of this discourse risks positioning students primarily as potential rule-breakers rather than as emerging professionals who must develop judgment and ethical agency within algorithmically mediated health care systems.

Together, these findings identify an algorithmic formation gap in nursing education discourse. While the literature actively debates whether students should use AI in academic work, comparatively little attention is directed toward how nurses learn to practice alongside algorithmic systems in contemporary health care.

The Integrity Discourse: Necessary but Insufficient

The prominence of academic integrity discourse is not surprising. Academic integrity represents a familiar institutional challenge for universities, with established responses including plagiarism detection, honor codes, and assessment redesign [3,4]. Within this framing, GenAI becomes a new technological manifestation of a longstanding concern about academic misconduct.

However, this regulatory orientation may also reflect what can be understood as problem displacement. Integrity is a manageable challenge that universities can address through policy development, surveillance technologies, and compliance mechanisms [5]. By contrast, the deeper challenge posed by AI involves reconsidering how professional education prepares students for clinical practice environments increasingly shaped by algorithmic decision-support systems.

Empirical research suggests that students themselves often approach GenAI pragmatically. Rather than using AI primarily to circumvent learning, many students report using it as a study aid while expressing uncertainty regarding institutional expectations and professional boundaries [29-31]. An educational response focused primarily on prevention and detection may therefore overlook an important opportunity: developing the professional judgment required to navigate AI responsibly within clinical practice.

Integrity remains an essential foundation of professional practice. However, integrity discourse alone cannot adequately address the broader transformation of health care environments in which nurses must interpret, question, and sometimes override algorithmic recommendations.

The Algorithmic Formation Gap

If professional education aims to develop individuals who "think, act, and feel" like members of the profession [12], then the integration of AI raises fundamental questions for professional formation in nursing.

Professional identity formation involves the integration of professional values, knowledge, and behaviors into an individual's sense of self [12,13]. When clinical practice environments incorporate algorithmic decision-support systems, automated risk stratification tools, and AI-assisted documentation technologies, the context in which nurses develop professional identities is fundamentally altered.

XSL•FO
RenderX

This review suggests that nursing education discourse has not yet fully engaged with this transformation. While the literature acknowledges ethical and technological challenges associated with AI, sustained attention to how nurses develop professional identities within algorithmically mediated care environments remains limited.

It is important to acknowledge that nursing is not alone in responding to the educational implications of AI. Medical education has similarly focused on AI literacy, ethical considerations, governance, and the integration of AI into clinical reasoning curricula. More recently, however, medical education scholarship has begun to extend beyond technical competency by considering human-AI collaboration, professional responsibility, and professional identity formation as integral components of AI-enabled practice [6]. Likewise, broader work examining clinical decision-support systems has highlighted how algorithmic technologies may reshape clinicians' autonomy, judgment, and professional identity [7]. These developments suggest that professional formation is becoming an increasingly recognized concern across the health professions. However, the present review indicates that this perspective remains comparatively underdeveloped within the nursing education literature, where discourse continues to be dominated by academic integrity, assessment validity, and technological competency. Rather than positioning nursing as uniquely deficient, the findings suggest that nursing is participating in a broader transition across health professions education while highlighting opportunities for the discipline to more explicitly integrate professional identity formation into AI-related curriculum development.

The implications of this formation gap can be understood across 3 interrelated dimensions: thinking, acting, and feeling as a professional nurse.

Thinking: Clinical Judgment in Algorithmic Contexts

Clinical judgment is a foundational capability of professional nursing practice. Tanner's model of clinical judgment emphasizes the integration of noticing, interpreting, responding, and reflecting in complex patient situations [32].

The literature reviewed here frequently emphasizes the importance of critical evaluation of AI-generated outputs [18,19]. These approaches are valuable in developing analytical reasoning skills. However, they primarily frame clinical judgment as a capacity applied to algorithmic recommendations rather than as a capability that may itself be reshaped by algorithmic practice environments.

Algorithmic decision-support systems can introduce risks such as automation bias, in which clinicians defer to automated recommendations even when conflicting evidence is present [33]. Similarly, excessive reliance on automated outputs may contribute to deskilling, whereby clinicians' independent pattern recognition and intuition gradually diminish [34].

Preparing students for algorithmically mediated care therefore requires more than the ability to critique AI outputs. It requires developing the capacity to integrate algorithmic information into clinical reasoning while maintaining independent professional judgment.

Acting: Moral Agency and Professional Courage

A second dimension of professional formation concerns moral agency as the ability to act on ethical conclusions even when doing so requires challenging institutional or technological authority.

Nursing has long emphasized patient advocacy as a central professional responsibility [35]. In algorithmic contexts, advocacy may involve questioning automated recommendations, identifying bias within predictive models, or prioritizing individualized patient needs over standardized algorithmic outputs.

Despite the ethical concerns identified across the literature, including bias, transparency, and privacy, few papers address how nurses develop the moral courage required to challenge algorithmic decision-support systems when patient welfare demands it [9,11,23,36]. Ethical awareness alone is insufficient if it does not translate into the capacity for moral action.

Professional formation in the age of AI must therefore include opportunities for students to practice exercising judgment and advocacy in situations where algorithmic recommendations conflict with clinical assessment.

Feeling: Relational Care in Technologically Mediated Environments

A third dimension of professional identity formation concerns the relational and affective aspects of nursing practice. Nursing has historically defined itself through commitments to compassion, presence, and holistic patient care.

The integration of algorithmic technologies introduces new forms of technological mediation between nurses and patients. Clinical dashboards, automated alerts, and AI-assisted documentation systems may shift the focus of clinical attention from the bedside to the digital interface.

This shift raises important questions regarding how nurses sustain relational care and moral responsibility when clinical decision-making becomes increasingly mediated by technological systems. Professional formation must therefore address how nurses maintain human connection, empathy, and patient advocacy within technologically complex health care environments.

Addressing the Algorithmic Present

A striking observation from this review is the limited acknowledgment that AI is already embedded within contemporary health care delivery. Across the literature, AI is frequently framed as an emerging or future technology rather than as an existing component of clinical infrastructure.

In reality, algorithmic systems already influence many aspects of clinical practice, including sepsis prediction models, early warning systems, risk stratification tools, and automated documentation support within electronic health records. These systems shape triage decisions, clinical workflows, and resource allocation across health care settings.

The relative absence of this "algorithmic present" within nursing education discourse may reflect disciplinary boundaries or

XSL•FO
RenderX

technological invisibility, where background computational systems are not recognized as AI by educators and clinicians. However, preparing students for a technological future while underengaging with the algorithmic realities of contemporary practice risks widening the formation gap identified in this review.

Addressing this gap requires stronger partnerships between clinical environments and nursing education to ensure that curricula reflect the technological infrastructures already shaping health care delivery.

Toward a Professional Formation Framework

Addressing the algorithmic formation gap requires reframing AI integration as a professional formation challenge rather than a technological or regulatory issue.

This reframing shifts the focus from detecting inappropriate AI use toward preparing nurses to practice ethically and effectively within algorithmically mediated health care systems.

Table 2 outlines 5 conceptual shifts that may support this reframing. These shifts involve moving from integrity-focused discourse to curricular approaches that emphasize identity formation, moral agency, and professional judgment in technological contexts.

Examples of such approaches may include simulation scenarios involving algorithmic failures, structured reflection on professional identity in technologically mediated practice, and case-based ethical discussions involving AI-supported clinical decision-making.

Rather than treating AI as a discrete technological topic, nursing education may need to integrate AI considerations longitudinally across curricula, linking technological engagement with professional values, clinical reasoning, and patient advocacy.

Table . Proposed conceptual shifts for AI in nursing education.

Integrity/tool discourse	Professional formation framework
AI as an integrity threat	AI as a professional practice context
Tool competency and literacy	Identity formation (think/act/feel)
Ethical awareness/principles	Moral agency and professional resistance
Critical evaluation of output	Transformation of clinical judgment
Future technological preparation	Engagement with current algorithmic reality

Implications for Nurse Educators

For nurse educators, the findings of this review suggest that GenAI should not be addressed solely through academic integrity policies or isolated technological skills training. Instead, AI must be integrated into curricula as part of professional formation. This includes creating learning experiences in which students critically engage with algorithmic decision-support systems, reflect on their evolving professional identities in technologically mediated environments, and practice exercising clinical judgment when AI recommendations conflict with patient needs. By situating AI within professional formation rather than academic compliance, educators can better prepare graduates for contemporary health care systems in which algorithmic tools increasingly shape clinical practice.

Implications for Nursing Education

The question facing nursing education is no longer whether students should use AI in their coursework. The more pressing question is how nurses will exercise judgment, advocacy, and relational care within health care systems increasingly shaped by algorithmic decision-making.

Preparing students for these environments requires moving beyond governance and literacy toward professional formation. This includes creating opportunities for students to engage critically with AI systems, reflect on their evolving professional identities, and practice exercising clinical judgment in technologically complex care environments.

By reframing AI integration as a professional formation challenge, nursing education can ensure that graduates possess not only the technical literacy to interact with algorithmic systems but also the ethical judgment, moral courage, and relational capacity required to advocate for patients within them.

Limitations

This review has several limitations. First, the rapid evolution of GenAI means that relevant literature published or in press after the search period may not have been captured. Given the pace of scholarly and technological developments in this field, findings should be interpreted as a time-sensitive snapshot of the discourse between 2022 and early 2026. Second, the search was limited to English-language publications and 3 databases (PubMed, Google Scholar, and SciSpace). Although these sources captured a broad and diverse corpus, relevant studies indexed elsewhere or published in other languages may have been missed. Third, the integrative review design intentionally included heterogeneous evidence types, such as empirical studies, reviews, theoretical papers, and commentaries. While this breadth enabled conceptual mapping of the field, the inclusion of nonempirical literature may limit direct comparability across studies and may amplify emerging opinion-based discourse. Fourth, the analysis relied on explicit reporting within published papers. The absence of discussion regarding professional identity formation, moral agency, or recognition of clinical AI infrastructure in the literature does not necessarily indicate the absence of these considerations in educational practice; rather, it reflects their limited articulation within the published discourse. Finally, the interpretive synthesis

XSL•FO
RenderX

involved researcher judgment in coding and thematic interpretation. To enhance rigor, independent coding and reviewer discussion were used to refine categories and resolve disagreements.

Conclusions

This integrative review identifies a fundamental algorithmic formation gap in nursing education discourse. The literature predominantly frames AI as a content-generation tool and academic integrity threat, with extensive attention to detection, prevention, and policy frameworks. While these concerns are legitimate, the disproportionate focus on integrity may obscure a more critical challenge: preparing nursing students for professional practice in algorithmically mediated health care environments.

The literature's limited engagement with professional identity formation, moral agency, clinical judgment transformation, and the reality of algorithmic clinical practice represents a critical gap. Nursing education cannot adequately prepare students for contemporary practice if it treats AI primarily as a writing tool to be regulated rather than as a defining context for professional formation.

The absence of professional formation discourse is not merely a gap to be filled with additional content or competencies. It represents an argument for fundamental curricular transformation. If AI already mediates triage, diagnostics, care pathways, and clinical decision-making in practice settings and the evidence from health care delivery suggests it does, then nursing education must explicitly address how students develop professional identity, clinical judgment, and moral agency within these algorithmic contexts.

This transformation requires moving beyond the question "how do we prevent students from cheating with AI?" to the more fundamental question "how do we prepare students to think, act, and feel like nurses in algorithmically mediated care environments?" It requires recognizing that professional formation in the age of AI is not about acquiring discrete competencies but about developing the capacity to exercise clinical judgment, moral agency, and relational practice alongside algorithmic systems.

The nursing profession has long defined itself through commitments to patient advocacy, holistic care, clinical judgment, and moral courage. As algorithms increasingly mediate health care delivery, these professional commitments are not obsolete; they are more essential than ever. However, they require explicit, sustained, and curricular attention to how they are enacted in algorithmic contexts.

Nursing education stands at a critical juncture. Continuing to treat AI primarily as an integrity threat leaves graduates unprepared for the algorithmic reality of modern health care. By reframing AI integration as a professional formation challenge, nursing education can ensure that the next generation of nurses possesses not only the literacy to use AI but also the judgment and agency to advocate for patients within it.

Generative AI Use

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) to assist with language editing, clarity, and structural refinement. All AI-assisted suggestions were critically reviewed by the authors, who retain full responsibility for the accuracy, originality, and integrity of the published work.

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 14 KB - nursing_v9i1e99284_app1.docx]

Multimedia Appendix 2
Data extraction table.
[DOCX File, 69 KB - nursing_v9i1e99284_app2.docx]

Checklist 1
PRISMA 2020 checklist.
[PDF File, 88 KB - nursing_v9i1e99284_app3.pdf]

References

1. Gonzalez L. Nursing education in the era of ChatGPT: implications and opportunities. Online J Issues Nurs 2024;29(3). [doi: 10.3912/OJIN.Vol29No03PPT76]

XSL•FO
RenderX

2. Topaz M, Peltonen LM, Michalowski M, et al. The ChatGPT effect: nursing education and generative artificial intelligence. J Nurs Educ 2025 Jun;64(6):e40-e43. [doi: 10.3928/01484834-20240126-01] [Medline: 38302101]
3. Shorey S, Mattar C, Pereira TLB, Choolani M. A scoping review of ChatGPT's role in healthcare education and research. Nurse Educ Today 2024 Apr;135:106121. [doi: 10.1016/j.nedt.2024.106121] [Medline: 38340639]
4. Siah CJR, Lim JV, Lyu MM, Levett-Jones T, Mattar I. Authenticity and academic integrity in generative artificial intelligence (GenAI) use among undergraduate nursing students. J Prof Nurs 2026;62:45-51. [doi: 10.1016/j.profnurs.2025.10.007] [Medline: 41679804]
5. Labrague LJ. A five-tier framework for guiding responsible AI use in nursing students' coursework: a faculty guide. Teach Learn Nurs 2026 Jan;21(1):62-66. [doi: 10.1016/j.teln.2025.09.012]
6. Fu C, Li J, Fan H, et al. AI literacy in undergraduate medical education: a competency-based interpretive framework for curriculum and assessment. Front Med (Lausanne) 2026;13:1871524. [doi: 10.3389/fmed.2026.1871524] [Medline: 42523801]
7. Ackerhans S, Huynh T, Kaiser C, Schultz C. Exploring the role of professional identity in the implementation of clinical decision support systems-a narrative review. Implement Sci 2024 Feb 12;19(1):11. [doi: 10.1186/s13012-024-01339-x] [Medline: 38347525]
8. Amankwaa I, Ekpor E, Cudjoe D, et al. Patterns, advances, and gaps in using ChatGPT and similar technologies in nursing education: a PAGER scoping review. Nurse Educ Today 2025 Oct;153:106822. [doi: 10.1016/j.nedt.2025.106822] [Medline: 40644844]
9. Hardie P, Darley A, Derwin R, et al. Applications, attitudes and ethical considerations of generative artificial intelligence (Gen AI) in nursing education: a scoping review. BMC Nurs 2026 Jan 16;25(1):148. [doi: 10.1186/s12912-025-04253-9] [Medline: 41545825]
10. Hinsche L, Hasseler M, Tischendorf T, Schaal T. Rapid review on GenAI in nursing education. Front Health Serv 2025;5:1725425. [doi: 10.3389/frhs.2025.1725425] [Medline: 41623519]
11. Luque Loor AH, Cedeño Orejuela JA, Melis Sosa A, et al. Implications of generative artificial intelligence on ethics in nursing education: state of the art review [Article in Spanish]. Rev Esp Edu Med 2025;6(4):676041. [doi: 10.6018/edumed.676041]
12. Cruess RL, Cruess SR, Boudreau JD, Snell L, Steinert Y. Reframing medical education to support professional identity formation. Acad Med 2014 Nov;89(11):1446-1451. [doi: 10.1097/ACM.0000000000000427] [Medline: 25054423]
13. Cruess RL, Cruess SR, Steinert Y. Amending Miller's pyramid to include professional identity formation. Acad Med 2016 Feb;91(2):180-185. [doi: 10.1097/ACM.0000000000000913] [Medline: 26332429]
14. Browne CA, Fetherston CM. How do we facilitate international clinical placements for nursing students: a cross-sectional exploration of the structure, aims and objectives of placements. Nurse Educ Today 2018 Jul;66:1-7. [doi: 10.1016/j.nedt.2018.03.023] [Medline: 29627757]
15. Whittemore R, Knafl K. The integrative review: updated methodology. J Adv Nurs 2005 Dec;52(5):546-553. [doi: 10.1111/j.1365-2648.2005.03621.x] [Medline: 16268861]
16. Chen SL, Lee LL. Strengths, weaknesses, opportunities, and threats analysis of integrating ChatGPT into nursing education. Hu Li Za Zhi 2024 Oct;71(5):7-13. [doi: 10.6224/JN.202410_71(5).02] [Medline: 39350704]
17. Stratton-Maher D, Lindsay J, Jacka L. The impact of generative AI on portfolio assessment in the bachelor of nursing program at the University of Southern Queensland. In: eBook of Shortened Peer Reviewed Papers: E-Portfolios Australia; 2024:17-22 URL: https://eportfoliosaustralia.wordpress.com/wp-content/uploads/2024/09/ebook-of-shortened-peer-reviewed-papers.pdf#page=17 [accessed 2026-10-15]
18. Simms RC. Work with ChatGPT, not against: 3 teaching strategies that harness the power of artificial intelligence. Nurse Educ 2024;49(3):158-161. [doi: 10.1097/NNE.0000000000001634] [Medline: 38502607]
19. Fung TCJ, Chan SL, Lam CFM, et al. Effects of generative artificial intelligence (GenAI) patient simulation on perceived clinical competency among global nursing undergraduates: a cross-over randomised controlled trial. BMC Nurs 2025 Jul 17;24(1):934. [doi: 10.1186/s12912-025-03492-0] [Medline: 40676632]
20. Hwang BY, Yu M. From film to simulation: mixed-methods study of learner-driven artificial intelligence-generated ethical dilemma scenarios and role-play to enhance ethical competence in nursing students. Nurse Educ Pract 2026 Mar;92:104773. [doi: 10.1016/j.nepr.2026.104773] [Medline: 41722254]
21. Go J, Lee Y. The effectiveness of AI-based conversational agents in nursing education: a systematic review. Nurse Educ Pract 2026 Feb;91:104728. [doi: 10.1016/j.nepr.2026.104728] [Medline: 41547257]
22. Yang S, Shi M, Qian Y, et al. A large language model-powered reflective AI agent for evidence-based nursing education: design and evaluation. Nurse Educ Pract 2026 Feb;91:104710. [doi: 10.1016/j.nepr.2026.104710] [Medline: 41520587]
23. Yuan Y, Fu J, Leng L, et al. The strengths, weaknesses, opportunities, and threats of generative artificial intelligence: a qualitative study of undergraduate nursing students. Front Public Health 2025;13:1672140. [doi: 10.3389/fpubh.2025.1672140] [Medline: 40977793]
24. Rodrigues D, Cruz-Correia R. Large language models in nursing education: state-of-the-art. Stud Health Technol Inform 2024 Aug 22;316:1024-1028. [doi: 10.3233/SHTI240584] [Medline: 39176964]

XSL•FO
RenderX

25. Choi MJ, Seo MH, Kim J, Kim S, Jeong SH. Research trends in generative artificial intelligence in nursing: a scoping review. J Korean Acad Nurs 2025 Aug;55(3):468-487. [doi: 10.4040/jkan.25006] [Medline: 40878571]
26. Gerdes M, Bayne A, Henry K, et al. Emerging artificial intelligence-based pedagogies in didactic nursing education: a scoping review. Nurse Educ 2025;50(1):E7-E12. [doi: 10.1097/NNE.0000000000001746] [Medline: 39383486]
27. Stratton-Maher D, Kelly J. Learning with AI, not from AI: ethical and pedagogical implications of ChatGPT in undergraduate education. Nurse Educ Today 2026 Mar;158:106961. [doi: 10.1016/j.nedt.2025.106961] [Medline: 41422773]
28. Vanderlaan J, Nicholas L, Leland N. Practical intelligence: generative AI toolkit for nurse education. : University of Nevada, Las Vegas (UNLV); 2025 URL: https://oasis.library.unlv.edu/nursing_fac_articles/424 [accessed 2026-10-15]
29. Sajja R, Sermet Y, Demir I, Chi NC. Evaluating artificial intelligence assisted nursing education: student perceptions, ethical concerns, and pedagogical implications. Teach Learn Nurs 2025;21:146-154. [doi: 10.1016/j.teln.2025.09.010]
30. Blomquist J, Alderden J, Llewellyn S, Haynes M, Dworak E. Student and faculty perceptions and use of generative artificial intelligence in nursing education. J Nurs Educ 2025 Aug;64(8):519-522. [doi: 10.3928/01484834-20250312-06] [Medline: 40801511]
31. Summers A, Haddad ME, Prichard R, Clarke KA, Lee J, Oprescu F. Navigating challenges and opportunities: nursing student's views on generative AI in higher education. Nurse Educ Pract 2024 Aug;79:104062. [doi: 10.1016/j.nepr.2024.104062] [Medline: 38996582]
32. Tanner CA. Thinking like a nurse: a research-based model of clinical judgment in nursing. J Nurs Educ 2006 Jun;45(6):204-211. [doi: 10.3928/01484834-20060601-04] [Medline: 16780008]
33. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc 2012;19(1):121-127. [doi: 10.1136/amiajnl-2011-000089] [Medline: 21685142]
34. Parasuraman R, Manzey DH. Complacency and bias in human use of automation: an attentional integration. Hum Factors 2010 Jun;52(3):381-410. [doi: 10.1177/0018720810376055] [Medline: 21077562]
35. Lachman VD. Moral courage in action: case studies. Medsurg Nurs 2007 Aug;16(4):275-277. [Medline: 17907703]
36. De Gagne JC, Hwang H, Jung D. Cyberethics in nursing education: ethical implications of artificial intelligence. Nurs Ethics 2024 Sep;31(6):1021-1030. [doi: 10.1177/09697330231201901] [Medline: 37803810]

Abbreviations

GenAI: generative AI
MMAT: Mixed Methods Appraisal Tool
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by E Borycki; submitted 24.Apr.2026; peer-reviewed by J Wang, JAC Santos, R Nguyen; revised version received 29.Jul.2026; accepted 05.Oct.2026; published 30.Oct.2026.

Please cite as:
Hartin P, Baretta P
From Academic Integrity to Professional Formation in Generative AI–Supported Nursing Education: Integrative Review
JMIR Nursing 2026;9:e99284
URL: https://nursing.jmir.org/2026/1/e99284
doi:10.2196/99284

© Peter Hartin, Paul Baretta. Originally published in JMIR Nursing (https://nursing.jmir.org), 30.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Thirty-Day Readmission Risk Stratification at Hospital Discharge: Single-Center Temporal Validation and Nursing Implementation Readiness Study

Hui-Wen Po[1], RN, MSc; Tzu-Yu Huang[1], RN, MSc; Ying-Chien Chu[1], RN; Hui-Chen Tsai[1], RN; Chung-Yu Chen[2], MD, PhD; Chien-Hsin Yang[3], PhD

[1]Department of Nursing, National Taiwan University Hospital Yunlin Branch, Douliou, Taiwan

[2]Department of Internal Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Cheng-Hsing Campus, No.1, University Road, Tainan, Taiwan

[3]Department of Industrial Engineering and Management, Chaoyang University of Technology, Taichung, Taiwan

Corresponding Author:
Chung-Yu Chen, MD, PhD
Department of Internal Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Cheng-Hsing Campus, No.1, University Road, Tainan, Taiwan

Abstract

Background: Discharge-planning nurses and case managers need to prioritize transitional care, but predictive ranking alone does not identify which patients will benefit from a specific intervention.

Objective: This study aimed to develop and temporally validate a local discharge time model for 30-day readmission; evaluate calibration, risk tiers, clinical utility, and subgroup performance; and describe historical nursing action cohorts only as non–model-triggered feasibility evidence for a future pathway.

Methods: This single-center observational study used 41,935 discharge episodes from 2017 to 2018. After excluding 2.3% (964/41,935) in-hospital deaths, 48.5% (19,890/40,971) of episodes from 2017 were used for development, and 51.5% (21,081/40,971) of episodes from 2018 were used for temporal validation. The outcome was the same-institution all-cause readmission within 30 days, defined as the union of readmissions on days 1 to 14 and 15 to 30. Logistic regression, random forest, and Extreme Gradient Boosting were calibrated using 5-fold out-of-fold predictions from 2017. The 95% CIs were obtained with 1000 stratified bootstrap replicates. The separate historical cohorts included shared decision-making (n=72), long-term care linkage (n=364), and remote monitoring (n=43). None was selected or routed by model output.

Results: The validation cohort had 20.7% (4361/21,081) readmissions. Random forest was selected from development cross-validation and achieved an area under the receiver operating characteristic curve (AUROC) of 0.816 (95% CI 0.809-0.823), area under the precision-recall curve (AUPRC) of 0.593 (95% CI 0.579-0.607), and Brier score of 0.119 (95% CI 0.116-0.121); the calibration intercept was –0.117, the slope was 0.974, and the observed-to-expected ratio was 0.951. Readmission occurred in 6.5% (661/10,120) of low-risk episodes, 18.5% (1160/6271) of moderate-risk episodes, 47.3% (1769/3739) of high-risk episodes, and 81.1% (771/951) of very high–risk episodes. The principal diagnosis group accounted for 84.6% of the random forest importance. After excluding index episodes labeled as planned admission or discharge and refitting, the AUROC decreased to 0.726, and the AUPRC decreased to 0.284. In historical cohorts, discharge direction changed in 41.7% (30/72) of cases; 82.7% (301/364) of patients used long-term care; 90.2% (268/297) of patients with valid dates started service within 7 days; and the uncontrolled remote monitoring cohort showed exploratory 3-month reductions in hospital days, emergency visits, and readmissions.

Conclusions: The recalculated local model showed useful temporal discrimination and near-unity calibration after probability calibration, but its performance depended heavily on diagnosis coding and potentially planned readmissions. External generalizability and model actionability remain unestablished. Prospective silent-run evaluation, local threshold review, subgroup monitoring, and controlled testing of the nurse-led response are required before clinical activation.

(JMIR Nursing 2026;9:e101335) doi:10.2196/101335

KEYWORDS

nursing informatics; discharge planning; transitional care; clinical decision support; hospital readmission; machine learning; calibration; case management

XSL•FO
RenderX

Introduction

Thirty-day readmission remains an important indicator of discharge quality and postacute care transitions, and reducing avoidable readmission is a continuing health system priority [1,2]. Discharge-planning nurses and case managers coordinate patient and caregiver education, medication reconciliation, postacute referrals, and postdischarge escalation. Recent international syntheses indicate that transitional care can reduce health care use, but effectiveness varies by population, intervention complexity, follow-up duration, and implementation fidelity [3,4]. Hospitals, therefore, need transparent ways to prioritize limited nursing resources rather than delivering the most intensive intervention to every patient.

Readmission models frequently achieve only modest discrimination and may not translate into useful bedside decisions [5,6]. Earlier transitional care reviews similarly showed heterogeneous intervention effects [7,8]. Recent work suggests that nursing data can contribute predictive information beyond administrative variables [9]. Predictive performance, however, is distinct from clinical actionability: a model may rank patients by risk without identifying who will respond to shared decision-making, long-term care linkage, remote monitoring, or another care component.

Our institutional program previously examined discharge-planning effectiveness, nursing care models, remote health monitoring, and shared decision-making with long-term care linkage [10-13]. These studies describe potentially feasible care components, but they do not constitute a model-triggered intervention pathway. The present study, therefore, treats those cohorts only as historical, illustrative evidence of feasibility and not as proof that the prediction model improves outcomes.

Prediction model guidance emphasizes transparent data sources, predictor timing, outcome definitions, validation, calibration, and applicability [14-16]. For nursing informatics, a further concern is that functional status, caregiver capacity, health literacy, self-management ability, symptom burden, and nursing complexity may be documented in structured assessments, standardized nursing terminologies, or narrative notes rather than administrative fields. Natural language processing and standardized nursing concepts offer possible methods for recovering this information [17-20].

This study aimed to (1) develop and temporally validate a single-center discharge time model for same-institution 30-day readmission; (2) evaluate probability calibration, risk tier separation, decision curve net benefit, subgroup performance, and sensitivity to outcome and cohort definitions; and (3) describe historical nursing action cohorts as illustrative evidence for designing a future, prospectively evaluated workflow. We did not test whether model-triggered transitional care improved patient outcomes.

Methods

Study Design and Setting

We conducted a multicomponent observational study at National Taiwan University Hospital Yunlin Branch. The primary component was the development of a prediction model using discharge episodes from 2017, with temporal validation on 2018 episodes. Historical shared decision-making, long-term care linkage, and remote monitoring datasets were analyzed separately to describe the available nursing services. The model was not visible to clinicians, did not alter care, and did not trigger enrollment in any historical cohort. The action cohorts were not used for model training, calibration, model selection, threshold derivation, or treatment effect estimation.

Prediction Cohort and Outcome

The source extract contained 41,935 discharge episodes. We excluded 964 episodes ending in in-hospital death, leaving 40,971 episodes for analysis: 19,890 in 2017 and 21,081 in 2018. All age groups were retained. The deidentified extract did not contain a stable patient identifier. Therefore, unique patient counts could not be reconstructed, and repeated discharge episodes were treated as independent observations (Table S1 in Multimedia Appendix 1).

The primary outcome was same-institution all-cause readmission within 30 days after discharge. The raw dataset contained separate indicators for readmission on days 1 to 14 and days 15 to 30; the cumulative 30-day outcome was defined as either indicator being positive. The extract did not capture readmissions to other hospitals and did not provide a reliable indicator distinguishing planned from unplanned subsequent admissions. Accordingly, potentially planned readmissions were included in the primary outcome.

During revision, an audit of the raw readmission fields identified that the previously reported 9.7% reflected only days 15 to 30 rather than cumulative 30-day readmission. The outcome was reconstructed as the logical union of the 2 windows, and all downstream analyses were rerun. The revised outcome definition, analysis code, and outputs were independently reviewed and verified by the authors and study statistician (Table S1 in Multimedia Appendix 1).

Postdischarge deaths were not treated as censoring events because the extract did not support time-to-event modeling with complete competing risk data. Among the analytic episodes, 233 deaths occurred within 30 days after discharge; 105 had a recorded readmission before death, and 128 did not. Episodes without a recorded readmission were classified as nonevents. Sensitivity analyses excluded these deaths and used a composite of readmission or death within 30 days.

Candidate Predictors and Missing Data

Candidate predictors were restricted to information available at discharge: age band, sex, length of stay, broad principal diagnosis group, and discharge destination group. Age was categorized as 0 to 17 years, 18 to 39 years, 40 to 59 years, 60 to 79 years, or 80 years or older. Principal diagnosis labels that changed between calendar years were harmonized into clinically broad groups, including a planned admission or discharge category. Destination indicators and raw disposition codes were combined into the following categories: home, facility, acute transfer, community service, other, or missing. The putative prior readmission field was not modeled because it was blank

XSL•FO
RenderX

in 99.9% (40,931/40,971) of the episodes and could not be interpreted reliably.

Age was missing in 0.0% (1/40,971) of episodes, length of stay was missing in 2.2% (887/40,971), and discharge destination was missing in 6.3% (2566/40,971); sex and diagnosis group were complete after harmonization. Categorical missingness was represented explicitly. Length of stay was capped at the 99th percentile in 2017 (54 days), transformed as log(1 + length of stay), and median imputed within the modeling pipeline. Variable definitions are reported in Table S2 in Multimedia Appendix 1, and period-level missingness is reported in Table S1 in Multimedia Appendix 1.

Model Development, Calibration, and Temporal Validation

We evaluated prespecified logistic regression, random forest, and Extreme Gradient Boosting models. Categorical variables were one-hot encoded. Categories occurring fewer than 5 times were grouped by the encoder. Hyperparameters were fixed before examining the 2018 validation outcomes and are reported in Table S3 in Multimedia Appendix 1. Five-fold stratified cross-validation within 2017 generated out-of-fold predictions. Logistic calibration (Platt scaling) was fitted to these out-of-fold predictions and then applied to predictions from each model refitted on all 2017 episodes.

The candidate model with the highest 2017 out-of-fold area under the precision-recall curve (AUPRC) was selected; Brier score and area under the receiver operating characteristic curve (AUROC) were secondary criteria. The untouched 2018 cohort was used for temporal validation. Performance was summarized using the AUROC, AUPRC, Brier score, calibration intercept, calibration slope, mean predicted risk, observed-to-expected ratio, and a decile calibration plot [14-16] (Table S4 and Figure S1 in Multimedia Appendix 1). The 95% CIs for the AUROC, AUPRC, and Brier score were calculated using 1000 stratified nonparametric percentile bootstrap replicates with seed 20260801. The existing institutional rule-based high-risk flag was retained as a comparator and evaluated in the temporal validation cohort.

Decision curve analysis estimated the net benefit across risk thresholds from 0.05 to 0.50. Because this was a retrospective decision analytic assessment and the score triggered no care, net benefit was interpreted as providing potential decision support rather than as demonstrating clinical effectiveness (Figure S2 in Multimedia Appendix 1).

Operational Risk Tiers

Four provisional tiers were derived from the calibrated 2017 out-of-fold random forest distribution: below the 50th percentile, from the 50th to below the 80th percentile, from the 80th to below the 95th percentile, and the 95th percentile or higher. The resulting cutoff points (0.1256, 0.3014, and 0.7746) were applied without re-estimation to 2018. Intervals were lower inclusive and upper exclusive except for the low tier, which included probabilities below 0.1256, and the very high tier, which included probabilities of 0.7746 or higher. The tiers were designed for workload planning, not as validated treatment thresholds (Table S5 in Multimedia Appendix 1).

Subgroup and Sensitivity Analyses

For the selected random forest, we calculated the AUROC, AUPRC, Brier score, observed-to-expected ratio, and calibration slope within age (<65 vs ≥65 years), sex, diagnosis group, and discharge destination subgroups when sample size and event counts permitted. Sensitivity analyses excluded postdischarge deaths within 30 days, used a readmission-or-death composite, restricted the cohort to adults, and excluded index episodes labeled as planned admission or discharge. The fixed-model analysis applied the full cohort model unchanged to the restricted 2018 subset; the refitted analysis re-estimated the random forest and its calibration using the corresponding restricted 2017 subset. This index diagnosis exclusion is an imperfect proxy for potentially planned readmissions, not a direct classification of whether a subsequent admission was scheduled. Random forest feature importance was summarized descriptively, with full subgroup and sensitivity results reported separately (Tables S6-S9 in Multimedia Appendix 1).

Historical Nursing Action Cohorts and Data Provenance

Three historical datasets were used only to illustrate candidate components for a future nurse-led pathway. No patient was selected or routed to these services by the prediction model. The supplied 72-case shared decision-making dataset and 364-case long-term care dataset match the sample sizes and key outcomes reported by Huang et al [13]. They are therefore treated as overlapping study arms rather than independent evidence. Their incremental contribution here is limited to corrected descriptive denominators and explicit placement within a proposed, not-yet-implemented workflow, consistent with guidance on overlapping publication [21] (Table S10 in Multimedia Appendix 1).

The supplied remote monitoring workbook contained 43 patients enrolled between November 12, 2020, and November 10, 2021, with 3-, 6-, and 12-month use windows. It appears to be an updated or overlapping program dataset compared with the previously published cohort [11], but no patient-level crosswalk was available to quantify the overlap. It is therefore presented as historical program data and not as an independent replication.

Remote monitoring outcomes were paired count measures and were compared using 2-sided Wilcoxon signed-rank tests. No multiplicity adjustment was applied. All *P* values are exploratory. The uncontrolled pretest-posttest design is vulnerable to regression to the mean, secular trends, and confounding by indication and does not support causal claims (Table S11 in Multimedia Appendix 1).

Proposed Future Nursing Workflow

The proposed pathway separates retrospective model evaluation from future clinical use. Phase 1 is a prospective silent run in which predictions remain hidden and do not change care. The hospital information system would log data availability, missingness, scores, tiers, outcomes, drift, subgroup performance, and projected nursing workload. Only after prespecified technical, calibration, equity, and operational review should activation be considered. In any subsequent activation, the score should prompt structured nursing

XSL•FO
RenderX

assessment and shared decision-making rather than automatically determining discharge placement or service enrollment [22] (Table S12 in Multimedia Appendix 1).

Statistical Analysis of Historical Cohorts

Continuous variables were summarized as means with SDs or medians with IQRs, as appropriate; categorical variables were summarized as counts and percentages. Long-term care initiation times were calculated only when service use and valid nonnegative discharge-to-service dates were available. Remote monitoring means were retained for comparability with the program dataset, while Wilcoxon signed-rank tests were used for paired inference.

Ethical Considerations

This study was approved by the Research Ethics Committee of National Taiwan University Hospital Yunlin Branch (202403148RINE). Because routinely collected clinical data and retrospectively assembled cohorts were used, the requirement for informed consent was waived. Patients and the public were not involved in the design, conduct, reporting, or dissemination planning of this retrospective analysis. Future prospective deployment should include patient and caregiver feedback on risk communication, shared decision-making, and the acceptability of remote monitoring.

Results

Cohort Flow and Characteristics

Of 41,935 source discharge episodes, 964 in-hospital deaths were excluded, leaving 40,971 analytic episodes. Mean age was 62.2 (SD 17.9; range 0-105) years, 57.2% (23,425/40,971) of the episodes involved male patients, and median length of stay was 5 (IQR 3-9) days. The cumulative same-institution 30-day readmission rate was 20.8% (8536/40,971), comprising 4465 events on days 1 to 14 and 4071 events on days 15 to 30. The 2017 development cohort contained 21.0% (4175/19,890) of the events. The 2018 validation cohort contained 20.7% (4361/21,081) of the events (Table S1 in Multimedia Appendix 1).

Model Performance and Calibration

Random forest achieved the highest development out-of-fold AUPRC and was selected before temporal validation. In 2018, it achieved an AUROC of 0.816 (95% CI 0.809-0.823), an AUPRC of 0.593 (95% CI 0.579-0.607), and a Brier score of 0.119 (95% CI 0.116-0.121; Table 1). Extreme Gradient Boosting had a slightly higher AUROC but lower AUPRC, while logistic regression was less discriminative. The existing rule-based flag had an AUROC of 0.527; readmission occurred in 24.5% (1181/4817) of flagged episodes and 19.6% (3180/16,264) of unflagged episodes (Table 1).

Table . Temporal validation performance of candidate models[a].

Models[b]	AUROC[c] (95% CI)	AUPRC[d] (95% CI)	Brier score (95% CI)
Random forest	0.816 (0.809-0.823)	0.593 (0.579-0.607)	0.119 (0.116-0.121)
Extreme Gradient Boosting	0.817 (0.810-0.824)	0.589 (0.575-0.604)	0.119 (0.116-0.121)
Logistic regression	0.806 (0.798-0.813)	0.541 (0.528-0.557)	0.121 (0.119-0.124)

[a]95% CIs for the 3 statistical models were calculated using 1000 stratified nonparametric percentile bootstrap replicates. The binary rule-based flag was included as an operational comparator; a bootstrap CI was not calculated for that fixed flag.

[b]Existing rule-based flag: AUROC=0.527, AUPRC=0.217, and Brier score=0.323.

[c]AUROC: area under the receiver operating characteristic curve.

[d]AUPRC: area under the precision-recall curve.

The random forest calibration intercept was –0.117, the calibration slope was 0.974, and the observed-to-expected ratio was 0.951; the mean predicted risk was 21.8% (SD 22.2%) compared with an observed rate of 20.7% (4361/21,081). The calibration plot is provided in Figure S1 in Multimedia Appendix 1. Decision curve analysis showed greater net benefit than treat-all and treat-none strategies across thresholds from 0.05 to 0.50, but this retrospective analysis does not establish benefit from actual clinical use (Figure S2 in Multimedia Appendix 1).

Operational Risk Tiers

Applying development-derived probability cutoff points to the 2018 cohort produced a steep readmission gradient (Table 2). The validation distribution differed modestly from the target 50%, 30%, 15%, and 5% development proportions, as expected in temporal validation. These tiers describe risk separation and projected workload; they were not used to assign care (Table S5 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Random forest risk tiers in the 2018 temporal validation cohort[a].

Tier	Validation[b], n (%)	Readmission, n/N (%)	Calibrated score interval	Illustrative future nursing action
Low	10,120 (48.0)	661/10,120 (6.5)	<0.1256	Standard education and routine follow-up
Moderate	6271 (29.7)	1160/6271 (18.5)	0.1256 to <0.3014	Nurse follow-up and medication reconciliation
High	3739 (17.7)	1769/3739 (47.3)	0.3014 to <0.7746	Case management review; assess SDM[c], home care, or long-term care
Very high	951 (4.5)	771/951 (81.1)	≥0.7746	Multidisciplinary review; assess intensive follow-up or facility needs

[a]Cutoff points were the 50th, 80th, and 95th percentiles of calibrated 2017 out-of-fold random forest predictions. Intervals are lower inclusive and upper exclusive except the low tier and the very high lower bound shown above. Actions are illustrative and were not delivered according to model output.

[b]Validation-cohort percentages use all 21,081 discharge episodes as the denominator (N=21,081).

[c]SDM: shared decision-making.

Feature Importance, Subgroups, and Sensitivity Analyses

The principal diagnosis group accounted for 84.6% of the total random forest impurity importance, followed by length of stay (8.6%), age band (4.5%), sex (1.2%), and discharge destination (1.2%). The planned admission or discharge indicator alone accounted for 62.5% of the total importance. In validation, 74.0% (1783/2411) of the episodes in the planned category had a recorded 30-day return (Table S6 in Multimedia Appendix 1).

Subgroup performance was heterogeneous. Random forest AUROC was 0.861 among patients younger than 65 years and 0.751 among those aged 65 years or older. It was 0.827 in female episodes and 0.807 in male episodes. By discharge destination, AUROC was 0.825 for home discharge but 0.539 for facility discharge. Diagnosis-specific AUROC values were generally modest despite substantial overall discrimination, consistent with strong separation between diagnosis categories rather than uniformly strong within-category ranking. Full subgroup results are reported in Tables S7 and S8 in Multimedia Appendix 1.

Excluding index episodes labeled as planned admission or discharge and refitting the random forest reduced the validation sample to 18,670 episodes with 2578 (13.8%) readmissions. AUROC decreased to 0.726, AUPRC decreased to 0.284, and the calibration slope was 0.936. Results were materially unchanged when postdischarge deaths were excluded, when readmission or death was used as a composite, or when analysis was restricted to adults (Table S9 in Multimedia Appendix 1).

Historical Nursing Action Cohort Outcomes

Selected historical nursing cohort outcomes not triggered by the model are listed in Table 3. In the overlapping shared decision-making dataset, discharge direction changed in 41.7% (30/72) of cases, the recommended and actual directions were concordant in 59.7% (43/72) of cases, and 15.3% (11/72) of patients had any readmission within 30 days. In the overlapping long-term care dataset, 82.7% (301/364) of patients used long-term care services. Among 297 users with valid nonnegative service dates, the median time to service initiation was 1 (IQR 0-5) day, 66.7% (198/297) initiated within 3 days, and 90.2% (268/297) initiated within 7 days (Table S13 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Selected historical nursing cohort outcomes not triggered by the model[a].

Historical cohort and outcome	Result[b]	Interpretation
Shared decision-making dataset		
Discharge direction changed	30/72 (41.7%)	Previously reported arm; illustrative only
Any 30-d readmission	11/72 (15.3%)	Union of 1-14–d and 15-30–d indicators
Long-term care dataset		
Used long-term care services	301/364 (82.7%)	Previously reported comparison arm
Long-term care users with valid dates		
Service started within 7 d	268/297 (90.2%)	Median 1 (IQR 0-5) d
Remote monitoring		
3-mo hospital days[c]	Mean of 9.49 (SD 14.78) before enrollment to mean of 3.51 (SD 9.63) after enrollment (*P*=.02)	Exploratory Wilcoxon comparison
3-mo emergency visits[c]	Mean of 0.65 (SD 0.69) before enrollment to mean of 0.19 (SD 0.55) after enrollment (*P*=.001)	Exploratory Wilcoxon comparison
3-mo readmissions[c]	Mean of 0.58 (SD 0.59) before enrollment to mean of 0.28 (SD 0.63) after enrollment (*P*=.02)	Exploratory Wilcoxon comparison

[a]The 72- and 364-case datasets overlap with the study by Huang et al [13]. The 43-patient remote monitoring dataset appears updated or overlapping with the study by Po et al [11], but patient-level overlap could not be quantified. No historical cohort was selected or routed using model scores.

[b]Results are presented as n/N (%) for the shared decision-making and long-term care datasets and as means, SDs, and *P* values for the remote monitoring dataset.

[c]Mean changes from before to after enrollment (SD) were −5.98 (17.57) for 3-month hospital days, −0.47 (0.83) for 3-month emergency visits, and −0.30 (0.77) for 3-month readmissions.

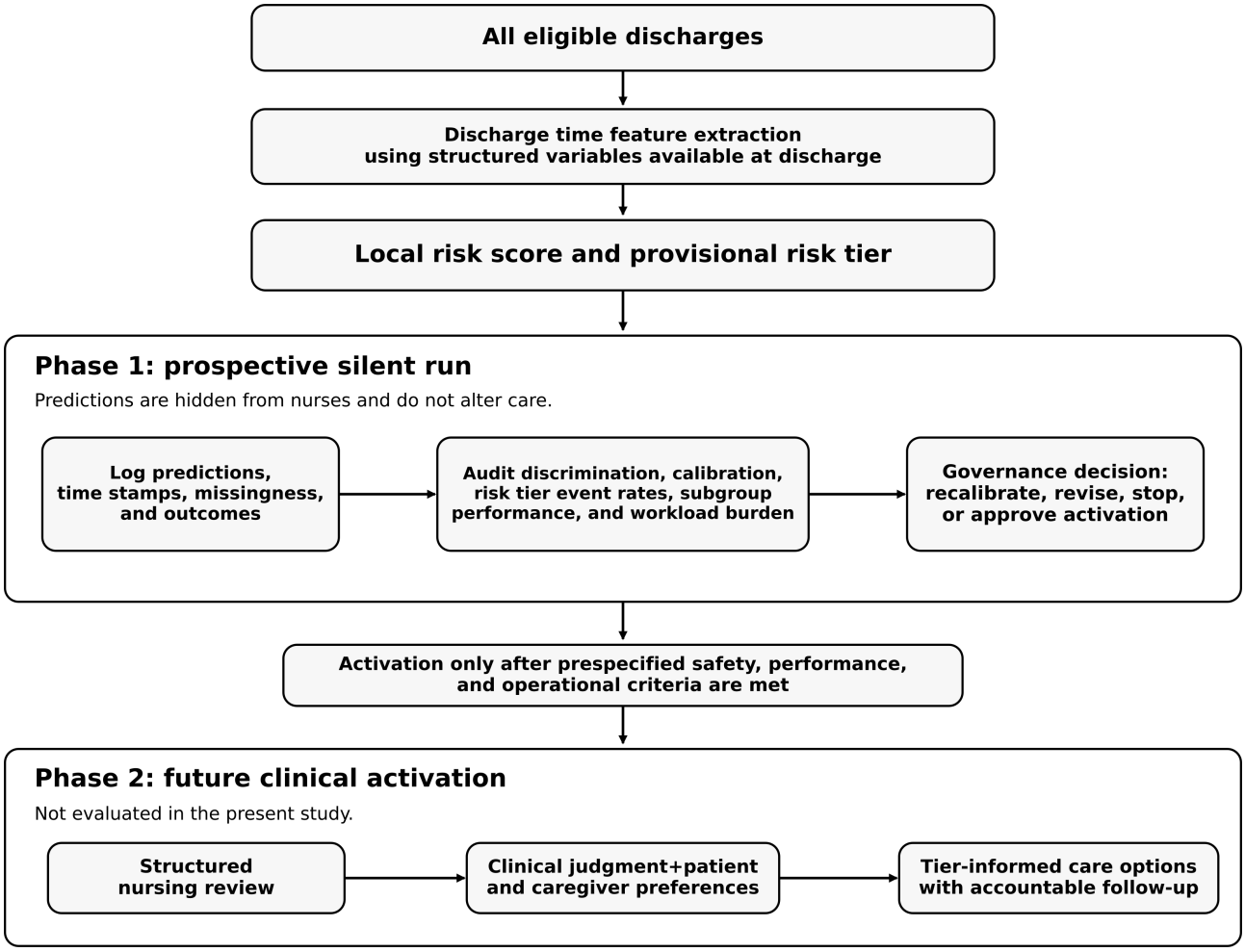
In the 43-patient remote monitoring dataset, 3-month hospital days decreased from a mean of 9.49 (SD 14.78) before enrollment to 3.51 (SD 9.63) after enrollment (mean change –5.98; *P*=.02), emergency visits decreased from a mean of 0.65 (SD 0.69) to 0.19 (SD 0.55) (mean change –0.47; *P*=.001), and readmissions decreased from a mean of 0.58 (SD 0.59) to 0.28 (SD 0.63) (mean change –0.30; *P*=.02), while outpatient visits increased from a mean of 0.88 (SD 1.42) to 1.60 (SD 1.56) (mean change +0.72; *P*=.02). At 6 months, emergency visits and readmissions remained lower, whereas hospital days did not differ significantly. At 12 months, the differences in emergency visits and hospital days were not statistically significant. Because no multiplicity correction or concurrent control group was used, these findings represent only exploratory associations (Table 3 and Table S11 in Multimedia Appendix 1).

Proposed Implementation Pathway

Figure 1 shows the proposed 2-phase pathway. A prospective silent run, technical audit, calibration, subgroup review, and nursing workload assessment precede any activation. If activation is approved later, model output should initiate structured nursing review and shared decision-making rather than automatically determining the discharge destination or service enrollment (Table S12 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 1. Proposed staged nursing-facing discharge risk workflow. Historical nursing action cohorts are illustrative and are not triggered by the model. Phase 1 is a prospective silent run and phase 2 represents future clinical activation only after prespecified safety, performance, equity, and operational criteria are met.



The historical shared decision-making, long-term care linkage, and remote monitoring cohorts in this study were illustrative.
They were not selected, routed, or treated according to model output and do not establish treatment benefit among model-defined risk groups.

Discussion

Principal Findings

This recalculated single-center study has 4 main findings. First, the corrected cumulative 30-day outcome occurred in approximately one-fifth of discharge episodes rather than the previously reported 9.7%, which represented only the window of days 15 to 30. Second, a calibrated random forest showed useful temporal discrimination, precision-recall performance, and near-unity calibration. Third, the overall performance and tier gradient were driven heavily by principal diagnosis, particularly index episodes labeled as planned admission or discharge. Discrimination decreased substantially when this category was excluded. Fourth, the historical nursing cohorts described available services but did not demonstrate that the model identified patients who benefited from those services.

A model can identify patients with a higher probability of readmission without identifying which intervention, if any, will reduce that probability. Demonstrating actionability requires the prospective linkage of model output to a predefined nursing assessment and care pathway, the measurement of treatment delivery and workload, and the comparison with an appropriate control strategy. The historical cohorts supply only feasibility information for designing that future evaluation.

Probability calibration was acceptable after Platt scaling, and decision curve analysis suggested potential net benefit across a broad range of thresholds. Nevertheless, the high readmission rate in the very high tier should not be interpreted as evidence that intensive transitional care will reduce events. The outcome included potentially planned returns, and the cutoff points were distribution based rather than derived from treatment effects or cost-effectiveness. Before activation, thresholds should be reviewed against contemporary event rates, nurse capacity, intervention eligibility, and the relative harms of missed and unnecessary escalation.

The dominance of the diagnosis group is both a performance signal and a vulnerability. It indicates that differences between broad diagnostic pathways explained much of the observed risk, but it also makes the model sensitive to coding changes, planned care patterns, and case mix shifts. The decrease in performance after excluding this index diagnosis category demonstrates sensitivity to index admission case mix; it does not establish

XSL•FO
RenderX

performance for confirmed unplanned subsequent readmissions. Lower performance among older patients and facility discharges also supports prospective subgroup monitoring and caution against a single unexamined threshold for all populations.

The current model omits many determinants of successful care transitions. Recent research on readmissions using nursing data suggests that nursing assessments may provide useful predictive information [9]. Narrative documentation may contain functional limitations, caregiver concerns, self-management barriers, symptom burden, education needs, and home care feasibility that are not visible in administrative coding. Natural language processing and terminology mapping can extract and standardize such activities, although methods and performance reporting remain heterogeneous [17,18].

Health literacy deserves particular attention. Prospective multicenter evidence indicates that inadequate health literacy is associated with readmissions and emergency department visits, especially among patients with high nursing complexity [19]. Standardized nursing diagnoses may also provide clinically interpretable predictors; a recent random forest study ranked individual nursing diagnoses by relevance to intensive care transfer [20]. Future refinement should evaluate structured nursing diagnoses, nursing complexity, health literacy, caregiver burden, functional dependency, discharge readiness, and medication self-efficacy.

No valid external validation was available, and generalizability beyond the development hospital has not been established. We removed the previously proposed University of California, Irvine (UCI) diabetes benchmark because its population, era, feature definitions, and outcome structure were too mismatched to support a clinically meaningful external validation claim. A future external study should use contemporaneous, clinically comparable data with harmonized outcome definitions, handling of planned readmissions, complete calibration assessment, and local recalibration as needed.

A future prospective study should begin with a silent run. Technical reliability, data latency, missingness, calibration, subgroup performance, case manager capacity, review time, alert burden, overrides, and care bundle completion should be measured before scores are shown to clinicians. Early-stage live evaluation should follow relevant clinical AI guidance, including DECIDE-AI (Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence) [23]. Any subsequent randomized evaluation of an AI-triggered care bundle should follow the CONSORT-AI (Consolidated Standards of Reporting Trials–Artificial Intelligence) [24] guidelines.

Limitations

This study has substantial limitations. It is a single-center retrospective development and temporal validation study without external validation. The outcome captured only same-institution returns and could not distinguish between planned and unplanned readmissions. Therefore, it is not equivalent to avoidable readmissions. Patient identifiers were unavailable, so repeated episodes were treated as independent, and patient-level clustering could not be modeled. Postdischarge deaths were retained as nonreadmissions when no prior return was recorded, although sensitivity analyses were stable. The feature set was parsimonious; diagnosis dominated; and excluded functional, social, caregiver, medication, physiological, and nursing assessment data. The risk tiers were based on score quantiles and have not been linked to treatment benefit or resource optimization. Subgroup results were retrospective, and some categories were small. The historical nursing datasets overlap with prior publications, are not linked to model scores, and do not constitute independent or causal evidence. The remote monitoring analysis was small, uncontrolled, and susceptible to regression to the mean and confounding by indication and involved multiple exploratory comparisons without multiplicity adjustment.

Conclusions

A calibrated discharge time random forest stratifies discharge episodes according to their predicted risk of readmission to the same hospital within 30 days after discharge, but the apparent performance depended heavily on diagnosis coding and potentially planned returns. Current evidence supports retrospective model evaluation and implementation planning but not external generalizability, automated care assignment, or benefit from model-triggered nursing interventions. A prospective silent run, explicit ascertainment of planned vs unplanned subsequent readmissions, subgroup and workload monitoring, local threshold review, and a later controlled evaluation of the nurse-led response are prerequisites for clinical activation.

Acknowledgments

The authors thank the discharge-planning, case management, shared decision-making, remote monitoring, and long-term care linkage teams of National Taiwan University Hospital Yunlin Branch for maintaining the clinical registries used in this study.

Generative AI Use

During revision, ChatGPT (GPT-6 Astra Pro; OpenAI) was used for language editing, document organization, and code-assisted verification and reanalysis of the deidentified datasets supplied by the authors. All variable definitions, code, outputs, interpretations, and manuscript changes were independently reviewed and verified by the authors and the study statistician. The AI tool did not make clinical decisions, determine authorship, or assume responsibility for the conclusions. The authors retain full responsibility for the accuracy and integrity of the work.

XSL•FO
RenderX

Funding

This research received no external funding.

Data Availability

The deidentified institutional datasets are not publicly available because they contain patient-level hospital information and are subject to institutional privacy and governance restrictions. Qualified investigators may request access to a governed analytical dataset from the corresponding author and host institution subject to institutional review and a data use agreement. Variable definitions, fixed hyperparameters, software versions, and random seed are documented in Tables S2 and S3 in Multimedia Appendix 1. A governed copy of the analysis code may be requested from the corresponding author subject to institutional review. Patient-level data are not embedded in the code.

Authors' Contributions

HWP and CYC conceived the study and drafted the manuscript. HWP, TYH, YCC, and HCT curated the data and interpreted the results. CHY performed the primary analyses and validation. CYC supervised the study. All authors critically revised the manuscript and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methods, tables, and figures describing 30-day readmission model development and temporal validation, subgroup and sensitivity analyses, historical nursing cohort outcomes, and a proposed implementation monitoring framework.
[DOCX File, 393 KB - nursing_v9i1e101335_app1.docx]

References

1. Hospital Readmissions Reduction Program. Centers for Medicare & Medicaid Services. URL: https://www.cms.gov/medicare/payment/prospective-payment-systems/acute-inpatient-pps/hospital-readmissions-reduction-program-hrrp [accessed 2026-08-01]
2. Hospital readmissions. Agency for Healthcare Research and Quality. URL: https://www.ahrq.gov/topics/hospital-readmissions.html [accessed 2026-08-01]
3. Tyler N, Hodkinson A, Planner C, et al. Transitional care interventions from hospital to community to reduce health care use and improve patient outcomes: a systematic review and network meta-analysis. JAMA Netw Open 2023 Nov 1;6(11):e2344825. [doi: 10.1001/jamanetworkopen.2023.44825] [Medline: 38032642]
4. Sakashita C, Endo E, Ota E, Oku H. Effectiveness of nurse-led transitional care interventions for adult patients discharged from acute care hospitals: a systematic review and meta-analysis. BMC Nurs 2025 Apr 7;24(1):379. [doi: 10.1186/s12912-025-03040-w] [Medline: 40197243]
5. Kansagara D, Englander H, Salanitro A, et al. Risk prediction models for hospital readmission: a systematic review. JAMA 2011 Oct 19;306(15):1688-1698. [doi: 10.1001/jama.2011.1515] [Medline: 22009101]
6. Yu MY, Son YJ. Machine learning-based 30-day readmission prediction models for patients with heart failure: a systematic review. Eur J Cardiovasc Nurs 2024 Oct 21;23(7):711-719. [doi: 10.1093/eurjcn/zvae031] [Medline: 38421187]
7. Kripalani S, Theobald CN, Anctil B, Vasilevskis EE. Reducing hospital readmission rates: current strategies and future directions. Annu Rev Med 2014;65:471-485. [doi: 10.1146/annurev-med-022613-090415] [Medline: 24160939]
8. Leppin AL, Gionfriddo MR, Kessler M, et al. Preventing 30-day hospital readmissions: a systematic review and meta-analysis of randomized trials. JAMA Intern Med 2014 Jul;174(7):1095-1107. [doi: 10.1001/jamainternmed.2014.1608] [Medline: 24820131]
9. Oh EG, Oh S, Cho S, Moon M. Predicting readmission among high-risk discharged patients using a machine learning model with nursing data: retrospective study. JMIR Med Inform 2025 Mar 5;13:e56671. [doi: 10.2196/56671] [Medline: 40106364]
10. Po HW, Lin FJ, Cheng HJ, et al. Factors affecting the effectiveness of discharge planning implementation: a case-control cohort study. J Nurs Res 2023 Jun 1;31(3):e274. [doi: 10.1097/jnr.0000000000000555] [Medline: 37167623]
11. Po HW, Chu YC, Tsai HC, Lin CL, Chen CY, Ma MH. Efficacy of remote health monitoring in reducing hospital readmissions among high-risk postdischarge patients: prospective cohort study. JMIR Form Res 2024 Sep 13;8:e53455. [doi: 10.2196/53455] [Medline: 39269747]
12. Po HW, Chu YC, Tsai HC, Chen CY, Chiu YW. Evaluate the differential effectiveness of the case management and primary nursing models in the implementation of discharge planning. J Clin Nurs 2025 Sep;34(9):3753-3775. [doi: 10.1111/jocn.17550] [Medline: 39528412]

XSL•FO
RenderX

13. Huang TY, Po HW, Chu YC, Tsai HC, Chen CY, Cheng BW. Shared decision making in discharge planning and long-term care linkage: a retrospective study. J Am Med Dir Assoc 2026 Jan;27(1):106003. [doi: 10.1016/j.jamda.2025.106003] [Medline: 41253289]
14. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating, 2nd edition: Springer; 2019. [doi: 10.1007/978-3-030-16399-0]
15. Collins GS, Moons KG, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
16. Moons KG, Damen JA, Kaul T, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ 2025 Mar 24;388:e082505. [doi: 10.1136/bmj-2024-082505] [Medline: 40127903]
17. Mitha S, Schwartz J, Hobensack M, et al. Natural language processing of nursing notes: an integrative review. Comput Inform Nurs 2023 Jun 1;41(6):377-384. [doi: 10.1097/CIN.0000000000000967] [Medline: 36730744]
18. Vanalli M, Cesare M, Cocchieri A, D'Agostino F. Natural language processing and string metric-assisted assessment of semantic heterogeneity method for capturing and standardizing unstructured nursing activities in a hospital setting: a retrospective study. Ann Ig 2023;35(1):3-20. [doi: 10.7416/ai.2022.2517] [Medline: 35403664]
19. Cocchieri A, Cristofori E, Nurchis MC, Damiani G, Cesare M. Nursing complexity and health literacy as determinants of patient outcomes: a prospective one-year multicenter cohort study. Nurs Rep 2025 Apr 17;15(4):135. [doi: 10.3390/nursrep15040135] [Medline: 40333082]
20. Cesare M, Nurchis MC, Damiani G, Cocchieri A. Ranking nursing diagnoses by predictive relevance for intensive care unit transfer risk in adult and pediatric patients: a machine learning approach with random forest. Healthcare (Basel) 2025 Jun 4;13(11):1339. [doi: 10.3390/healthcare13111339] [Medline: 40508953]
21. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals: overlapping publications. International Committee of Medical Journal Editors. 2026. URL: https://www.icmje.org/recommendations/browse/publishing-and-editorial-issues/overlapping-publications.html [accessed 2026-08-01]
22. Sittig DF, Wright A, Osheroff JA, et al. Grand challenges in clinical decision support. J Biomed Inform 2008 Apr;41(2):387-392. [doi: 10.1016/j.jbi.2007.09.003] [Medline: 18029232]
23. Vasey B, Nagendran M, Campbell B, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med 2022 May;28(5):924-933. [doi: 10.1038/s41591-022-01772-9] [Medline: 35585198]
24. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020 Sep;26(9):1364-1374. [doi: 10.1038/s41591-020-1034-x] [Medline: 32908283]

Abbreviations

AUPRC: area under the precision-recall curve
AUROC: area under the receiver operating characteristic curve
CONSORT-AI: Consolidated Standards of Reporting Trials–Artificial Intelligence
DECIDE-AI: Developmental and Exploratory Clinical Investigations of Decision Support Systems Driven by Artificial Intelligence
UCI: University of California, Irvine

Edited by E Borycki; submitted 14.May.2026; peer-reviewed by M Cesare, Y Yu; revised version received 14.Sep.2026; accepted 21.Oct.2026; published 09.Nov.2026.

Please cite as:
Po HW, Huang TY, Chu YC, Tsai HC, Chen CY, Yang CH
Thirty-Day Readmission Risk Stratification at Hospital Discharge: Single-Center Temporal Validation and Nursing Implementation Readiness Study
JMIR Nursing 2026;9:e101335
URL: https://nursing.jmir.org/2026/1/e101335
doi:10.2196/101335

© Hui-Wen Po, Tzu-Yu Huang, Ying-Chien Chu, Hui-Chen Tsai, Chung-Yu Chen, Chien-Hsin Yang. Originally published in JMIR Nursing (https://nursing.jmir.org), 9.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and

XSL•FO
RenderX

reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Double-Edged Sword Effect of AI Application Among Clinical Nurses Based on the Job Demands-Resources Model: Qualitative Study

Xiaoyan Zhang[1], PhD; Jiaxin Fang[2], MSN; Sihan Chen[2], MSN; Jiayin Luo[3], MSN

[1]Department of Vascular Surgery, Beijing Hospital, National Center for Gerontology; National Clinical Research Center for Gerontology; The Key Laboratory of Geriatrics of NHC; Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, P, Beijing, China

[2]School of Nursing, Beijing University of Chinese Medicine, Beijing, P.R. China, Beijing, China

[3]Department of Outpatient, Beijing Hospital, National Center for Gerontology; National Clinical Research Center for Gerontology; The Key Laboratory of Geriatrics of NHC; Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, No. 1, Dahuajie, Dongdan, Dongcheng District, Beijing, China

Corresponding Author:
Jiayin Luo, MSN
Department of Outpatient, Beijing Hospital, National Center for Gerontology; National Clinical Research Center for Gerontology; The Key Laboratory of Geriatrics of NHC; Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, No. 1, Dahuajie, Dongdan, Dongcheng District, Beijing, China

Abstract

Background: AI is rapidly transforming clinical nursing, promising administrative relief and decision support. However, the frontline reality presents a double-edged sword effect, where technological empowerment is frequently offset by novel occupational burdens and technostress.

Objective: This study aims to theoretically deconstruct the bidirectional impacts of AI application among clinical nurses and identify buffering conditions, using the Job Demands-Resources (JD-R) theoretical framework.

Methods: A descriptive qualitative study was conducted across multiple general hospitals in mainland China. Using maximum variation and purposive sampling, semistructured in-depth interviews were conducted with registered nurses who actively use clinical AI systems. Data were analyzed using directed qualitative content analysis guided by predefined JD-R constructs.

Results: The analysis revealed 3 overarching domains comprising 9 main themes and 24 subthemes. On the gain path (job resources), AI empowered nurses through a workflow efficiency leap, clinical cognitive empowerment, and professional capital appreciation. Conversely, along the drain path (job demands), hidden costs were exposed, conceptualized as cognitive impediment, relational attrition, and digital involution driven by performance inflation and competitive perfectionism. The interplay between these pathways was perceived to be buffered by contextual mechanisms, specifically nurses' proactive coping strategies, professional boundary demarcation, and the provision of a cohesive organizational support architecture.

Conclusions: The impact of AI integration in nursing is not technologically deterministic. While AI provides valuable cognitive and operational support, it concurrently generates novel digital demands. To prevent AI from devolving into an occupational hazard, health care administrators must establish multidimensional life cycle AI governance, cultivate comprehensive AI literacy, and safeguard the irreplaceable humanistic core of clinical care.

(JMIR Nursing 2026;9:e100916) doi:10.2196/100916

KEYWORDS

AI; clinical nursing; qualitative research; Job Demands-Resources model; double-edged sword

Introduction

The global health care landscape is currently undergoing a digital transformation, driven by the rapid expansion of AI technologies. To address the chronic global shortage of nursing personnel and the increasing complexity of patient care, AI has emerged as a strategic imperative [1,2]. In mainland China, this technological wave is particularly pronounced, propelled by robust national policies aimed at constructing "smart hospitals" [3]. Consequently, clinical nursing practice has witnessed the extensive deployment of diverse AI applications. These range from specialized clinical algorithms—such as big data-driven early warning scores, intelligent workforce scheduling, and clinical decision support systems—to generative AI tools increasingly used for administrative tasks like drafting documentation [4-6]. As the largest professional workforce within the health care ecosystem, clinical nurses are positioned

XSL•FO
RenderX

at the forefront of this digital revolution, inevitably serving as the core end users of these multifaceted AI systems [7,8].

In line with the optimistic view that AI alleviates administrative burdens and improves decision-making [9], the actual deployment of AI-based systems has led to measurable reductions in documentation transcription time and costs: outpatient clinic letter completion time fell by nearly 8 days, and transcription savings totaled Aus $16,903 (Aus $1=US $0.71 as of May 27, 2026) [10]. However, frontline nursing reality presents a starkly different narrative. Emerging evidence reveals a complex double-edged sword effect [11,12]. While these metrics underscore AI's capacity to successfully automate certain routine tasks, its integration has spawned novel occupational hazards. Frontline nurses report experiencing alarm fatigue from hypersensitive predictive models, technostress stemming from complex system maintenance, and a pervasive sense of autonomy deprivation induced by ubiquitous digital surveillance [11-13]. Alarmingly, the unprecedented accessibility of AI tools has implicitly elevated organizational performance expectations, triggering a phenomenon of efficiency inflation and subsequent hyper-competition among peers [14,15]. This juxtaposition of technological promise and clinical reality provokes a critical inquiry: are these advanced algorithms genuinely empowering the nursing workforce or inadvertently entrapping them within a rigid, machine-dictated ecosystem?

To theoretically make sense of this seemingly contradictory clinical phenomenon, the Job Demands-Resources (JD-R) model serves as an optimal analytical lens [16]. The JD-R model posits that all psychosocial work characteristics can be classified into 2 overarching categories: job demands and job resources (Multimedia Appendix 1). Viewed through this framework, AI is not a neutral artifact but an active agent dynamically reshaping nursing work. On the one hand, AI can function as a robust job resource—streamlining workflows or providing cognitive support—thereby initiating a motivational process that enhances clinical efficacy. Conversely, unreliable algorithms or heightened performance expectations can morph AI into a stringent job demand. The resulting cognitive overload and socioemotional labor trigger a health impairment process, ultimately culminating in professional burnout. Thus, the JD-R model provides a nuanced architecture to deconstruct the bidirectional impacts of AI [17,18].

Despite the theoretical use of the JD-R model, a critical review of the literature reveals significant methodological and contextual gaps. To date, research concerning AI in nursing has been predominantly quantitative, frequently relying on technology acceptance models to assess willingness to use or focusing on algorithmic diagnostic accuracy [19,20]. While recent qualitative scholarship has begun examining AI perception among nursing leaders and researchers [6,21], in-depth qualitative inquiries explicitly centered on clinical nurses' heterogeneous interactions with diverse AI applications remain limited. Standardized questionnaires inherently fail to capture psychological undercurrents, such as the cognitive friction of verifying AI hallucinations or the subtle erosion of human-machine trust [22]. Furthermore, within the uniquely high-intensity health care context of mainland China, the unintended sociocultural consequences of AI integration—specifically the phenomena of efficiency inflation and professional involution—remain largely unexplored.

To bridge this knowledge gap, this study aims to deeply explore the "double-edged sword" effect of AI applications among clinical nurses, using directed qualitative content analysis (DQCA) [23] guided by the JD-R model. The novelty of this study is to theoretically deconstruct the bidirectional impacts of AI on clinical nurses and to identify buffering conditions, using the JD-R theoretical framework. Specifically, this qualitative inquiry has three objectives: (1) explicate the gain path by mapping how AI functions as an empowering job resource; (2) delineate the drain path by identifying how AI simultaneously acts as a novel digital job demand; and (3) uncover the perceived buffering conditions that navigate the interplay between these dual pathways. Ultimately, this study endeavors to provide a robust qualitative empirical foundation for health care organizations to develop more sustainable, human-centric AI integration strategies.

Methods

Study Design

A descriptive qualitative study design was adopted to deeply explore the lived experiences of clinical nurses interacting with AI. Given that this study is theoretically anchored in the JD-R model, we used a DQCA approach. DQCA is particularly appropriate when prior theory exists about a phenomenon but would benefit from further description or contextualization. This approach allowed us to deductively map nurses' experiences onto the predefined JD-R constructs while retaining the inductive flexibility to capture emerging, culturally specific phenomena. This study is reported in accordance with the Standards for Reporting Qualitative Research (SRQR) [24] guidelines.

Study Setting

Study Setting and Participant Context

The study was conducted in mainland China between November 2025 and March 2026. In contemporary Chinese health care, the practice scope of registered clinical nurses is inherently multidimensional, encompassing direct bedside care, clinical administration/quality improvement, and practice-based continuing education and scholarly research. Consequently, frontline nurses interact with a heterogeneous spectrum of AI applications, including embedded clinical AI systems—institutionally integrated, algorithm-driven tools designed for direct patient care, risk monitoring, and operational management (eg, early warning scoring systems and clinical decision support systems)—and generative AI tools driven by large language models (LLMs; eg, DeepSeek [Hangzhou DeepSeek Artificial Intelligence Basic Technology Research Co, Ltd], Doubao [Beijing Chuntian Zhiyun Technology Co, Ltd], or ChatGPT [OpenAI]). Participants were recruited across 8 general hospitals (tertiary and secondary) and 1 affiliated primary community health center in 4 municipalities/provinces (Beijing, Shandong, Guangdong, and Sichuan). The community facility was included as an integrated primary-care arm of a

XSL•FO
RenderX

regional tertiary health alliance to capture grassroots AI adoption. In this study, clinical nurses were defined broadly as registered nurses providing direct patient care, including those working in primary care settings, to reflect the full spectrum of AI deployment across different tiers of the health care system.

Participant Eligibility, Sampling, and Recruitment

A purposive sampling strategy framed by a maximum variation matrix was adopted to ensure broad coverage across key demographic and clinical dimensions. Inclusion criteria were as follows: (1) registered nurses holding a valid license; (2) currently engaged in direct clinical care with at least 1 year of clinical experience; and (3) having active, hands-on experience with at least one AI-based system—such as embedded clinical AI systems or generative AI tools—in their daily workflow for ≥3 months. Exclusion criteria included nurses exclusively in administrative roles without direct AI clinical application. Participants were recruited via professional nursing association networks and hospital department announcements across participating institutions. A total of 37 eligible nurses were approached; 2 declined due to clinical shift conflicts and 1 withdrew before the interview for personal reasons, yielding a final sample of 34 (completion rate: 91.9%) participants. To construct the sampling matrix, we initially recruited nurses varying in age, professional titles, educational backgrounds, and hospital tiers, aiming to capture the widest possible spectrum of routine AI interaction experiences. Building upon this maximum-variation framework, we then deliberately applied intensified sampling to 2 subgroups of particular theoretical interest: tertiary hospital nurses (n=26, 76.5%) and male nurses (n=13, 38.2%). Tertiary teaching hospitals serve as the primary innovation hubs for advanced AI deployment in China, and oversampling this group allowed us to capture dense, firsthand accounts of complex AI system integration. Similarly, male nurses—a demographic minority within the profession—were intentionally oversampled to ensure that their perspectives on technostress and human-AI interaction were not marginalized in the final analysis. This layered sampling design prioritizes both the breadth of routine experiences and the depth of theoretically critical cases. Enrollment continued until data saturation was achieved, defined as the point where no new codes or themes emerged from the last 3 consecutive interviews.

Procedures

Interviews were conducted by 2 qualitative researchers (XZ and JF) trained in advanced qualitative content analysis. Neither interviewer had prior personal, clinical, or administrative relationships with the participants. To ensure reflexivity, both interviewers maintained reflexive journals throughout the study to acknowledge, document, and bracket preexisting assumptions regarding the JD-R model (the detailed interview guide is provided in Multimedia Appendix 2). A total of 14 interviews were conducted face-to-face in private, quiet hospital conference rooms, while 20 were conducted via a secure online video conferencing platform (Tencent Meeting [Tencent]) to accommodate geographically distant participants across provinces. All interviews were conducted in Mandarin Chinese and lasted between 20 and 60 minutes. Field notes were recorded during and immediately after each interview to capture nonverbal cues, vocal hesitations, and emotional responses, which were subsequently integrated into transcript familiarization.

Analysis

Audio recordings were transcribed verbatim in Mandarin within 48 hours. Selected verbatim quotes were translated into English by a bilingual nursing researcher and back-translated by an independent native English academic editor to maintain conceptual and cross-language equivalence. Data analysis was managed using NVivo 20 software (QSR International) and followed the DQCA systematic steps: (1) preparation and familiarization—2 researchers independently read the transcripts multiple times alongside field notes to achieve deep immersion in the data; (2) deductive coding—initial coding categories were generated deductively from predefined JD-R constructs, and transcripts were coded by assigning text passages to these predefined categories; (3) inductive coding—data that could not be categorized into the initial JD-R model were assigned new codes inductively using explicit rules; for instance, statements describing technology-driven peer hyper-competition and artificially escalated management performance baselines were inductively coded under digital involution. To ensure coding reliability, 2 coders independently coded an initial batch of 6 transcripts; (4) theme refinement, disagreement resolution, and rigor—the research team held regular consensus meetings to compare codes, resolve discrepancies, and abstract the codes into main themes and subthemes. Discrepancies were discussed until 100% agreement was reached. Theoretical saturation was audited through a cumulative codebook; code generation stabilized at Participant 31, and the final 3 consecutive interviews produced zero new codes or subthemes, confirming data saturation.

Ethical Considerations

Ethical approval for this study was obtained from the Ethics Committee of Beijing Hospital (approval 2024BJYYEC-KY203-02). Before data collection, all prospective participants were provided with a detailed information sheet outlining the study's purpose, procedures, and their rights. Written informed consent was obtained from each participant before the interviews commenced. Participants were explicitly informed that their participation was completely voluntary and that they had the right to withdraw from the study at any time without penalty. To address ethical and privacy concerns regarding technology use, the research team explicitly inquired about data security practices during the interviews. All participants who used noninstitutional, general-purpose AI tools reported manually executing strict deidentification protocols (eg, stripping patient names, identification numbers, specific dates, and institutional identifiers) prior to any text or data entry. Disclosures of unauthorized or unapproved AI practices were treated with strict research confidentiality, framed nonpunitively, and analyzed as systemic indicators of unmet operational efficiency demands rather than individual misconduct. To strictly protect participant confidentiality and anonymity, all audio recordings and transcripts were stored securely on password-protected devices. Each participant was assigned a pseudonym, and specific identifying details regarding their

XSL•FO
RenderX

employment were carefully modified or redacted to ensure they remained unidentifiable in the presentation of the findings.

Trustworthiness

To ensure methodological rigor, we adhered to Lincoln and Guba's evaluative criteria: (1) credibility—member checking was performed by returning a summary of the synthesized themes to 5 participants to verify whether the findings accurately reflected their realities. Additionally, peer debriefing was conducted with a senior qualitative researcher to challenge and refine the analytical interpretations. (2) Transferability—thick descriptions of the study setting, participant demographics, and contextual factors (eg, China's specific clinical environment) were provided to allow readers to assess the applicability of the findings to other contexts. (3) Reliability—a comprehensive audit trail was maintained, documenting all methodological decisions, coding dictionaries, and analytical memos. Two independent coders performed the analysis, and intercoder agreement was continuously negotiated. (4) Confirmability—the researchers practiced continuous reflexivity by keeping a reflexive journal to acknowledge and bracket their preexisting assumptions regarding AI and the JD-R model, ensuring that the findings were deeply grounded in the participants' narratives rather than the researchers' biases.

Results

Participant Characteristics

A total of 34 clinical nurses from various departments and hospital tiers were recruited through purposive sampling. The demographic profile of the participants is summarized in Multimedia Appendix 3. The participants' ages ranged from 23 to 43 years (mean age 30.65, SD 5.56 y), and their nursing experience spanned from 1 to 22 years (mean experience 8.12, SD 5.48 y). The sample consisted of 21 (61.8%) females and 13 (38.2%) males. Regarding educational background, 23 (67.6%) participants held an undergraduate degree, 7 (20.6%) held a postgraduate degree, and 4 (11.8%) had a junior college diploma or below. In terms of professional titles, the majority were intermediate-level nurses (n=19, 55.9%), followed by junior-level (n=11, 32.4%) and senior-level nurses (n=4, 11.8%). Participants were drawn from diverse clinical settings, including surgery (n=14, 41.2%), internal medicine (n=12, 35.3%), emergency (n=4, 11.8%), and intensive care units (n=4, 11.8%). Most participants (n=26, 76.5%) were recruited from tertiary hospitals, with the remainder from secondary (n=7, 20.6%) or community (n=1, 2.9%) health care facilities.

Across the sample, nurses interacted with a dual ecosystem of AI technologies, combining hospital-integrated clinical AI systems (eg, early warning scoring systems, clinical decision support systems, risk prediction systems, and smart infusion pumps) with personally accessed generative AI tools (eg, DeepSeek, Doubao, ChatGPT, and Kimi [Moonshot AI]). Hospital-sanctioned clinical AI systems were characterized by high daily workflow integration for direct bedside care following formal institutional onboarding, whereas generative AI tools were predominantly self-taught and used for administrative drafting, literature synthesis, and health education content creation (see Multimedia Appendix 4 for more details).

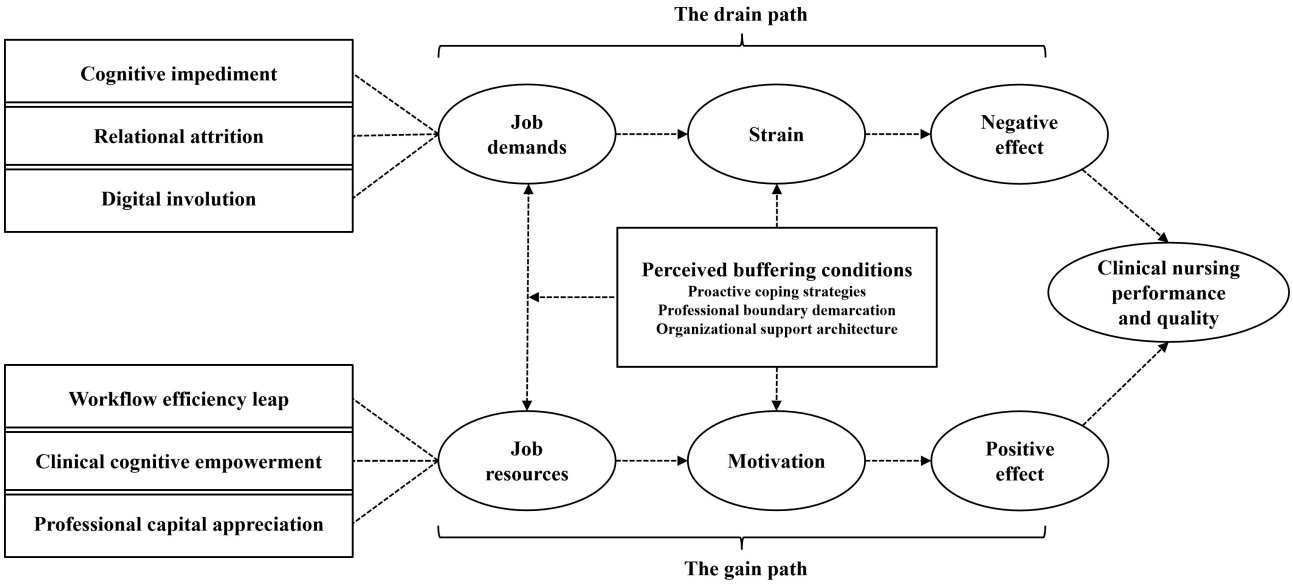
Overview of the Findings

Thematic Structure of the Findings

Through DQCA guided by the JD-R framework, we identified 3 overarching domains comprising 9 main themes and 24 subthemes that illustrate the complex double-edged sword effect of AI application in clinical nursing (Figure 1). Specifically, the gain path (job resources) demonstrates the empowering potential of AI by facilitating a workflow efficiency leap, providing clinical cognitive empowerment, and fostering professional capital appreciation. Conversely, the drain path (job demands) highlights the hidden and continuous costs of AI integration, conceptualized as perceived cognitive impediment, relational attrition that challenges humanistic empathy, and digital involution driven by performance inflation and competitive pressures. Furthermore, the interplay between these perceived gains and demands is contextualized by key perceived buffering conditions, specifically nurses' proactive coping strategies, strict professional boundary demarcation, and the presence of a supportive organizational support architecture. The detailed thematic coding tree is presented in Multimedia Appendix 5.

XSL•FO
RenderX

Figure 1. Qualitative conceptual map of the double-edged sword effect of AI application among clinical nurses under the Job Demands-Resources (JD-R) framework.



Domain 1: AI as a Job Resource (the Gain Path)

AI emerged as a robust job resource for clinical nurses, triggering a motivational process. By interacting with various AI systems, nurses extracted perceived empowerment, which manifested across 3 distinct themes: workflow efficiency leap, clinical cognitive empowerment, and professional capital appreciation.

Theme 1: Workflow Efficiency Leap

In this study, workflow efficiency leap is defined as structural optimization and operational acceleration of daily nursing duties through AI automation, enabling nurses to transition from administrative overextension to streamlined task execution. Participants frequently articulated how AI integration fundamentally remodeled their daily operations across 3 subthemes: documentation automation, task cycle reduction, and team synergy strengthening.

Subtheme 1: Documentation Automation

Nurses reported an alleviation of administrative burdens driven primarily by generative AI tools. Rather than expending energy on repetitive typing and formatting, nurses used AI to structure scattered data and polish professional communication. For instance, generative AI tools provided vital support for cross-lingual writing, refining the phrasing of workplace correspondence, and rapidly summarizing lengthy literature. This capability lowered the barrier to content creation and eased the cognitive friction associated with writer's block. As nurse N10 illustrated: "if I want to create an outline myself, it might take a bit longer. But if AI helps me, it only takes one or two minutes, or even just twenty or thirty seconds" (see Multimedia Appendix 5 for the full dataset).

Subtheme 2: Task Cycle Reduction

Both embedded clinical AI systems (eg, automated electronic health record data extraction) and generative AI tools acted as perceived accelerators across varied nursing duties. Participants noted a substantial reduction in task initiation time and overall completion cycles. Efficient data retrieval algorithms replaced traditional, time-consuming manual searches. By handling these low-value, repetitive tasks, AI provided notable psychological relief, reducing the subjective intensity and time pressure of administrative duties. As nurse N24 noted: "The main advantage is that it's extremely efficient. You just send it your key images, data, and requirements, and it generates the output for you. It's much more efficient than searching online yourself and analyzing or comparing things one by one; it's very fast and convenient. It has reduced the amount of time I spend" (see Multimedia Appendix 5 for the full dataset).

Subtheme 3: Strengthening Team Synergy

Beyond individual use, embedded clinical AI systems (such as intelligent shift scheduling software and task distribution platforms) functioned as collaborative catalysts. These tools facilitated transparent role assignments and optimized team task distribution. In scenarios requiring collective decision-making, generative AI tools enabled the rapid integration of multidisciplinary information. Automated scheduling systems enhanced team synchronization and facilitated perceived collective outputs. As nurse N3 stated: "During group discussions, we can use AI to organize the questions we need to answer and better establish our future direction" (see Multimedia Appendix 5 for the full dataset).

Theme 2: Clinical Cognitive Empowerment

This theme is conceptualized as the intellectual scaffolding provided by AI, which dynamically bridges individual knowledge gaps, reinforces clinical decision-making logic, and safeguards patient care in high-stakes clinical environments. In these complex settings, AI transcended its role as a mere physical tool to become an indispensable cognitive partner through instant knowledge replenishment, thinking logic scaffolding, and clinical safety safeguarding.

Subtheme 4: Instant Knowledge Replenishment

Nurses perceived generative AI tools and integrated clinical decision support systems as an on-demand knowledge repository

XSL•FO
RenderX

that was particularly crucial during emergencies or unfamiliar cases. The systems provided rapid access to specialized clinical knowledge, complex pharmacological data (eg, drug interactions), and evidence-based guidelines. This instant access effectively masked experience deficits among junior nurses. As nurse N6 noted: "I use AI to organize and recall anatomical positions and functions; after sorting it out, I have a general idea of the clinical purpose and direction" (see Multimedia Appendix 5 for the full dataset).

Subtheme 5: Thinking Logic Scaffold

When confronting cognitive bottlenecks, such as complex case analyses or practice-based academic reflection, generative AI tools served as a cognitive coach. They provided structural frameworks, established logical connections between disparate symptoms, and assisted in structuring care plans. Participants emphasized that AI stimulated secondary reflection and expanded thinking dimensions. As nurse N8 described: "When I'm stuck on how to approach something, I might need AI to point the way. For example, when writing a report for the first time, I needed AI to give me a logical pivot. It also told me how to structure the logic of the text" (see Multimedia Appendix 5 for the full dataset).

Subtheme 6: Clinical Safety Safeguarding

Embedded clinical AI systems (eg, early warning scoring models and closed-loop smart infusion pumps) acted as vigilant partners in direct patient care. By monitoring microphysiological trends, predictive models offered early warnings for clinical deterioration, effectively reducing continuous vigilance fatigue. Automated assistance in high-risk procedures was perceived to enhance clinical safety monitoring and foster technological trust. As nurse N23 noted: "The intelligent early warning system mainly monitors patients' vital signs... If any data is abnormal, it immediately alerts us so we can respond quickly" (see Multimedia Appendix 5 for the full dataset).

Theme 3: Professional Capital Appreciation

Professional capital appreciation refers to the long-term enrichment of nurses' psychological well-being, digital competencies, and occupational value, which fosters professional self-efficacy and enables a recentering on humanistic patient care. The use of AI generated these secondary benefits for nurses' career trajectories through emotional energy replenishment, cross-boundary skill expansion, and core value regression.

Subtheme 7: Emotional Energy Replenishment

Delegating routine tasks to both generative AI tools and embedded clinical AI systems generated significant positive emotional energy. Nurses reported a heightened sense of control over their work pace and a reduction in digital anxiety. The reliable, on-demand nature of AI assistance served as a psychological buffer during overwhelming workloads. As nurse N4 shared: "Even though it's complicated, we still have AI, right ?... It boosts my confidence in my work. With that confidence, I'm less likely to be bothered by difficulties" (see Multimedia Appendix 5 for the full dataset).

Subtheme 8: Cross-Boundary Skill Expansion

Generative AI tools functioned as accessible tutors that lowered thresholds for acquiring new digital competencies. Nurses used AI to bridge knowledge gaps, learn nursing research methodologies (eg, SPSS Statistics [Version 20.0; IBM Corp] operations and statistical workflows), and rapidly acquire digital content creation skills for patient education (eg, short video scripts). As nurse N33 shared: "It forced me to learn data analysis. Before, I just followed orders; now, I interpret AI-generated reports and analyze chronic disease trends in my district... It transformed me from an executor into a manager" (see Multimedia Appendix 5 for the full dataset).

Subtheme 9: Core Value Regression

A value-driven expectation emerged from the data. By absorbing trivial administrative and physical tasks, AI implicitly redirected nurses' finite attention back to the essence of nursing: direct patient care and complex emotional labor. Participants expressed future aspirations for advanced AI—such as intelligent ward-roaming robots or automated health education loops—to fully eradicate manual drudgery, thereby allowing them to reclaim their professional identity as empathetic caregivers. As nurse N1 stated: "It can take over more mechanical tasks, such as documentation or monitoring, allowing nurses to focus more on patient psychological care and complex clinical management" (see Multimedia Appendix 5 for the full dataset).

Domain 2: AI as a Job Demand (the Drain Path)

Viewed through the JD-R theoretical lens, the deployment of AI was not entirely benign. Instead, it introduced novel and complex job demands that triggered a potential health impairment process. Participants highlighted a perceived drain path, where the superficial alleviation of physical tasks was counterbalanced by subjective psychological, cognitive, and systemic burdens.

Theme 4: Cognitive Impediment

While AI promised cognitive relief, the unreliability and complex operational logic of current algorithms generated new dimensions of mental load, transforming nurses from care providers into algorithmic supervisors.

Subtheme 10: Verification Overload

A pervasive narrative was the persistent threat of algorithmic unreliability, which manifested distinct cognitive burdens across AI classes. For generative AI tools, nurses faced the threat of AI hallucinations (fabricated medical citations or inaccurate pharmacological data); for embedded clinical AI systems, nurses struggled with hypersensitive algorithms that caused frequent false alarms. Both mandated compulsory cross-verification, increasing cognitive fatigue and offsetting initial efficiency gains. As nurse N1 noted, "In individual diagnosis and treatment-related aspects, the results can vary significantly... We still need to cross-check with clinical experience or other case studies" (see Multimedia Appendix 5 for the full dataset).

Subtheme 11: Interaction Fatigue

The human-machine interaction associated with generative AI systems constituted a notable hidden demand. Participants described the high cognitive cost of prompt engineering,

XSL•FO
RenderX

continuously adjusting, refining, and debugging instructions to align AI outputs with complex clinical realities or institutional frameworks. This cycle of trial and error interrupted workflows. As nurse N5 illustrated: "Initially, when giving instructions to AI, it often didn't understand my intent, requiring constant refinement. Ambiguities inevitably arise, necessitating repeated questioning—this creates early-stage burdens"(see Multimedia Appendix 5 for the full dataset).

Subtheme 12: Mental Dependency

Long-term exposure to both generative AI tools and embedded clinical AI systems generated anxiety regarding professional deskilling. Nurses expressed deep concerns over cognitive blunting, fearing that outsourcing clinical reasoning, calculations, and diagnostic structuring to algorithms would atrophy their independent critical thinking and intuitive clinical judgment. As nurse N3 noted: "I unconsciously default to AI for task completion, bypassing independent thinking—this gradual reliance dulls cognitive engagement" (see Multimedia Appendix 5 for the full dataset).

Theme 5: Relational Attrition

The insertion of algorithms into the health care ecosystem disrupted traditional human dynamics, generating novel interpersonal conflicts and threatening the empathetic core of nursing practice.

Subtheme 13: Interpretive Labor

The democratization of consumer-accessible generative AI tools meant that patients and families frequently used these systems for self-diagnosis. Frontline nurses were forced to absorb substantial amounts of new interpretive labor—deconstructing AI-generated misinformation, managing patient anxiety, and rebuilding trust. As nurse N30 stated: "Patients and families create challenges by consulting unreliable online sources... This undermines our health education efforts" (see Multimedia Appendix 5 for the full dataset).

Subtheme 14: Care Dehumanization

Participants harbored deep apprehensions regarding technology-mediated alienation caused by standardized algorithmic care pathways and automated communication scripts. Nurses feared a loss of humanistic care, perceiving algorithmic reliance as an emotional barrier that reduced holistic nursing to rigid, machine-dictated transactions. As nurse N12 noted, "Compared to humans, it's deficient in humanistic care—a cornerstone of our profession" (see Multimedia Appendix 5 for the full dataset).

Subtheme 15: Professional Alienation

Embedded clinical AI systems (eg, performance-tracking algorithms and automated shift scheduling) and rapid technology iterations fostered professional insecurity. Senior staff facing the digital divide experienced anxiety regarding displacement and feeling managed by algorithms, which eroded their clinical agency. As nurse N23 described: "My greatest fear is AI wholly supplanting nursing roles, eliminating job opportunities" (see Multimedia Appendix 5 for the full dataset).

Theme 6: Digital Involution

In the highly competitive context of the modern health care system, digital involution is conceptualized as a sociotechnical paradox wherein technology-driven efficiency artificially inflates performance expectations, locking nurses into a cycle of hyper-competition, perfectionism, and invisible labor without proportional gains in well-being. Specifically, the integration of AI catalyzed unintended structural pressures across 3 interrelated dimensions: performance inflation, competitive perfectionism, and workload expansion.

Subtheme 16: Performance Inflation

The accessibility and speed of generative AI tools lowered the threshold for complex administrative and quality-improvement deliverables (eg, report synthesis or presentation design), which paradoxically triggered performance inflation. Management implicitly raised baseline standards, demanding perfect deliverables and forcing nurses into a cycle of continuous overperformance. As nurse N33 illustrated: "Management increasingly judges quality by numbers rather than substance. Quantity gradually overshadows quality focus, and performance benchmarks keep escalating." (see Multimedia Appendix 5 for the full dataset).

Subtheme 17: Competitive Perfectionism

The accessibility of AI-generated content fostered an environment of intensified peer comparison. Nurses felt compelled to produce flawlessly polished materials to maintain their professional standing, leading to a state of competitive perfectionism. This subtheme highlights how AI-driven high standards became the new minimum requirement, forcing individuals to invest excessive cognitive effort into refining AI outputs to avoid being outperformed by colleagues. As nurse N23 noted: "I have to spend time after work every day learning related knowledge, and I feel a lot of pressure. Before, when we wrote health education materials, 'good enough' was fine. But now with AI, the requirements keep getting higher, the quality keeps improving — it's very competitive" (see Multimedia Appendix 5 for the full dataset).

Subtheme 18: Workload Expansion

Contrary to labor-saving promises, findings reveal a paradoxical expansion of invisible digital labor across both technology classes. This includes troubleshooting embedded system errors, performing redundant dual-entry processes between legacy systems and AI platforms, and debugging generative AI outputs. As nurse N7 noted: "Nurses haven't had responsibilities reduced by AI; instead, we've gained the added duty of monitoring AI for malfunctions" (see Multimedia Appendix 5 for the full dataset).

Domain 3: Perceived Buffering Conditions for Navigating the Dual-Edged Sword Effect of AI

Within the JD-R model, the transition of AI from a potential job demand back to an empowering job resource is heavily contingent on specific buffering conditions. The analysis revealed that nurses actively and passively used various moderating mechanisms to navigate the double-edged nature of AI. These mechanisms, which mitigate the health impairment

XSL•FO
RenderX

process (ie, the strain path where high job demands deplete emotional and cognitive energy, leading to burnout) and enhance the motivational process (ie, the gain path where abundant job resources foster work engagement and self-efficacy), are conceptualized into 3 core themes: proactive coping strategies, professional boundary demarcation, and organizational support architecture.

Theme 7: Proactive Coping Strategies

Nurses did not remain passive recipients of technological outputs; instead, they engaged in highly proactive behavioral crafting to harness AI's use while neutralizing its risks.

Subtheme 19: Critical Verification Execution

To combat generative AI hallucinations and predictive model false alarms, nurses developed rigorous cross-verification habits. Rather than blindly trusting algorithmic outputs, participants used multisource cross-verification strategies, treating AI-generated content purely as a preliminary draft rather than definitive evidence. This active cognitive filtering served as a crucial buffer, preventing machine errors from translating into clinical incidents. As nurse N6 illustrated: "If findings contradict my foundational knowledge, I reevaluate and reassess AI's conclusions" (see Multimedia Appendix 5 for the full dataset).

Subtheme 20: Strategic Interaction Mastery

Nurses ameliorated interaction friction, specifically when using generative AI tools, by mastering prompt engineering, strategic data feeding (providing authentic clinical contexts), and task-technology calibration (delegating low-risk administrative drafting to AI while reserving complex clinical decisions for traditional human pathways). As nurse N15 shared: "Command refinement is key—iterative tweaks align outputs with needs" (see Multimedia Appendix 5 for the full dataset).

Theme 8: Professional Boundary Demarcation

A robust psychological and professional defense mechanism emerged as nurses deliberately fortified their occupational identity against technological encroachment.

Subtheme 21: Role Positioning Clarity

Participants conceptualized both generative AI tools and embedded clinical AI systems strictly as high-capacity subordinate assistants or information aggregators, vehemently rejecting AI as an autonomous decision-maker. Recognizing AI's lack of genuine empathy insulated nurses from replacement anxiety. As nurse N26 noted: "A textual assistant. For reference and support—nothing more" (see Multimedia Appendix 5 for the full dataset).

Subtheme 22: Clinical Agency Retention

Nurses preserved clinical autonomy by insisting on the irreplaceable nature of physical touch and empathetic care. In clinical decision-making, they adopted a "think first, consult second" paradigm, ensuring human reasoning framed the clinical strategy. As nurse N6 stated: "*I adhere to my own logic—AI won't override my reasoning. Critical thinking is non-negotiable. With independent thought, AI becomes a tool, not a puppeteer* "(see Multimedia Appendix 5 for the full dataset).

Theme 9: Organizational Support Architecture

The organizational environment served as a macroscopic macroregulator. The presence or absence of institutional backing dictated whether AI integration felt like an empowering resource or a competitive burden.

Subtheme 23: Systemic Resource Provision

Participants expressed a need for hospitals to transcend fragmented usage by deploying unified, institutionally sanctioned, vertically integrated medical AI platforms (eg, specialized medical LLMs trained on validated clinical guidelines). Centrally provided platforms lower cognitive thresholds and alleviate security concerns. As nurse N3 noted: "Hospitals could consolidate databases... Institution-led AI platforms would streamline retrieval" (see Multimedia Appendix 5 for the full dataset).

Subtheme 24: Capacity Building Support

Nurses emphasized the need for structured organizational training focusing on underlying algorithmic logic (for embedded clinical AI systems) and prompt structuring and error avoidance (for generative AI tools), backed by clear institutional policies and ethical boundaries. As nurse N3 described: "If hospitals adopt AI, they should develop structured teaching programs" (see Multimedia Appendix 5 for the full dataset).

Discussion

Summary of Principal Findings

Guided by the JD-R model, this qualitative study provides a nuanced qualitative empirical understanding of the double-edged sword effect of AI integration among frontline clinical nurses. Our findings demystify AI from being a unilaterally empowering artifact, revealing it instead as a highly dynamic agent that concurrently activates distinct gain and drain pathways. On the gain path, AI functions as a robust job resource that initiates a motivational process, characterized by a substantial workflow efficiency leap, critical clinical cognitive empowerment, and the subsequent appreciation of professional capital. Conversely, the drain path exposes the hidden sociotechnical costs where AI operates as a novel, multifaceted job demand. These demands manifest as cognitive friction, relational strain that threatens traditional empathetic care, and a pervasive systemic involution driven by technology-induced efficiency inflation. Crucially, our qualitative findings suggest that this technological paradox is not perceived as deterministic. Rather than demonstrating tested causal moderation, our data illustrate that the lived interplay between AI-driven resources and demands is contextualized by key perceived buffering conditions: nurses' proactive coping strategies, their rigorous professional boundary demarcation to preserve clinical agency, and the provision of a cohesive organizational support architecture.

The Double-Edged Sword Effect: Deconstructing the Gain and Drain Pathways

The Gain Path

Consistent with the motivational pathway of the JD-R model, our findings validate that AI serves as a potent perceived job

XSL•FO
RenderX

resource, initiating empowerment across 3 progressive dimensions: operational workflow, clinical cognition, and professional capital. First, regarding operational execution, our findings corroborate historical health information technology literature that emphasizes the physical and temporal benefits of automation [25]. We demonstrate a substantial workflow efficiency leap, where disaggregated technology classes yield distinct operational resources: generative AI tools streamline documentation automation and literature retrieval, whereas embedded clinical AI systems enhance shift scheduling and team synergy. These observations converge with recent qualitative scholarship, highlighting AI's capacity to alleviate administrative burdens and optimize workflow management [6,21,26]. Beyond routine automation, AI provides critical clinical cognitive empowerment. Specifically, embedded clinical AI systems (eg, early warning scores) act as safety sentinels, whereas generative AI tools function as on-demand cognitive tutors that bridge knowledge gaps and scaffold clinical reasoning. This positions AI as an active adjunct to human cognition [27-29], directly boosting nurses' self-efficacy [30,31] and rendering complex tasks manageable [32]. Ultimately, this operational and cognitive synergy fosters professional capital appreciation. By delegating mundane tasks to algorithms, nurses restore emotional resources, acquire cross-boundary digital skills, and experience a "core value regression"—redirecting finite attentional capacity back to the irreplaceable, humanistic core of patient care.

The Drain Path

Despite the optimistic narrative surrounding AI empowerment, our study uncovers a perceived health impairment process driven by novel digital demands. Consistent with Chuang et al [33], who demonstrated the dual impact of AI on work and life well-being under the JD-R model, our finding reinforces the model's suitability for unpacking the paradoxical effects of AI in clinical nursing contexts. A critical finding in this study is the manifestation of an AI productivity discrepancy rooted in cognitive impediment, which manifests through distinct technological mechanisms. While AI theoretically accelerates task completion, this temporal advantage is frequently neutralized by hidden cognitive loads. For generative AI tools, cognitive load stems from the threat of algorithmic hallucinations and omissions, forcing nurses into tedious prompt engineering and multisource verification; for embedded clinical AI systems, cognitive strain is driven by hypersensitive false alerts and alert fatigue [34]. As our data suggests, cognitive workload has not disappeared; it has merely shifted from primary content creation to secondary content verification. This verification pressure is further compounded by the psychological risk of automation bias (uncritical overreliance on AI outputs) versus the cognitive friction of overriding false alerts.

Furthermore, an important finding is that within the highly competitive environment of contemporary health care, the widespread accessibility of generative AI tools has inadvertently driven performance inflation. As AI lowers the technical barriers for producing polished administrative or quality-improvement work, nursing management implicitly raises baseline expectations. While prior literature often conceptualizes technological stress as individual operational difficulty, our findings uncover a novel sociotechnical phenomenon—digital involution—wherein frontline nurses are locked into a cycle of competitive perfectionism and invisible workload expansion (eg, dual-entry validation and prompt debugging). When combined with the interpretive labor required to bridge the gap between AI misinformation and patient understanding amid ambiguous legal responsibility frameworks, these demands constitute a substantial source of technostress. If left unaddressed, this invisible surge in task intensity, coupled with professional alienation and the fear of occupational displacement, threatens the long-term sustainability of the nursing workforce. These findings suggest that without organizational interventions to decouple AI efficiency from ever-increasing performance quotas, the double-edged sword of AI may inadvertently accelerate burnout rather than alleviate it.

Perceived Buffering Conditions: Navigating Demands and Resources

Another significant finding is that our qualitative findings illustrate how nurses perceive AI as shifting between a resource and a demand depending on contextual buffering conditions at individual, professional, and organizational levels. At the micro level, nurses exercise agency through proactive coping and strategic interaction proficiency [34]—tailoring prompt engineering for generative LLMs while maintaining vigilant verification for embedded clinical alerts. By selectively delegating routine tasks to algorithms while retaining high-stakes clinical judgments, practitioners optimize operational gains while mitigating cognitive risks. This resilience is anchored in professional boundary demarcation, where AI is positioned strictly as an auxiliary tool rather than an autonomous decision-maker. Such role clarity, combined with an emphasis on humanistic care, serves as a cognitive buffer against occupational replacement anxiety and technological alienation.

However, individual resilience alone is insufficient to sustain this balance. Organizational support architectures serve as the critical fulcrum. To prevent a cycle in which efficiency gains from technology inadvertently trigger escalating performance quotas, health care institutions must provide systemic resources. These include institutionally sanctioned, vertically integrated medical AI models (to eliminate the privacy and hallucination risks of generic public chatbots) and structured capacity-building programs focused on algorithmic logic, prompt calibration, and error avoidance. Ultimately, the synergy between individual professional agency and institutional scaffolding helps explain whether AI is perceived as an occupational hazard or a sustainable clinical resource.

Implications for Nursing Management and Practice

The findings of this study offer critical and actionable insights for health care administrators seeking to sustainably integrate AI into clinical nursing practice without exacerbating burnout.

First, to mitigate the cognitive impediments and legal anxiety associated with algorithmic hallucinations and false alerts, nursing administrators should actively collaborate with interprofessional hospital leadership to co-design and implement nursing-specific AI governance guidelines within the broader

XSL•FO
RenderX

organizational framework [35]. For instance, nurse leaders can establish clear operational protocols that delineate permissible AI boundaries—such as defining mandatory "human-in-the-loop" verification steps for AI-predictive risk alerts and establishing explicit guidelines that mandate clinical nurse sign-off before AI-generated documentation or care plans are finalized in patient records.

Second, health care institutions must establish a comprehensive, multidimensional AI life cycle governance framework encompassing people, processes, technology, and operations [36]. Rather than assuming that adopting institutionally sanctioned, vertically integrated medical AI models automatically eliminates privacy or reliability risks, hospital management must recognize that technological infrastructure requires robust operational oversight. Centrally funded and domain-specific clinical AI tools (eg, algorithms trained on validated guidelines and secure electronic health record databases) should be deployed alongside strict operational protocols: enforcing mandatory data deidentification, embedding clear digital audit trails, defining legal liability boundaries, and establishing explicit escalation procedures when algorithmic outputs conflict with human clinical judgment. Crucially, organizational policies must safeguard against automation bias and model drift by institutionalizing mandatory human oversight and conducting routine audits of joint human–AI system performance under both optimal and misleading AI outputs.

Third, the paradigm of capacity building support must evolve. Nursing education and continuous professional development programs should move beyond rudimentary software tutorials to cultivate comprehensive AI literacy. This involves equipping nurses with advanced prompt engineering techniques for generative AI tools, algorithmic awareness for embedded clinical AI systems, and standardized critical verification protocols.

Finally, to dismantle the escalating cycle of performance demands driven by technology, management must fundamentally recalibrate performance evaluation metrics. It is imperative that leadership ensures the temporal resources liberated by AI are channeled into a core value regression. Specifically, time saved should be redirected back to direct patient care and empathetic engagement rather than being absorbed by artificially inflated administrative workloads or escalating performance targets.

Strengths, Limitations, and Future Research

A key strength of this study lies in its robust theoretical anchorage within the JD-R model, providing theoretically grounded qualitative insights into the lived experiences of frontline nurses navigating AI. However, several limitations warrant acknowledgment. First, as a cross-sectional qualitative content analysis, this study captures participants' perceived experiences and subjective reflections; it cannot establish statistical causality, measure objective clinical performance/safety outcomes, or formally test quantitative mediation or moderation pathways. Second, because data were collected exclusively within mainland China, specific culturally embedded psychosocial dynamics may lack statistical generalizability and could manifest differently across diverse global health care ecosystems. Third, the reliance on digital recruitment strategies and voluntary participation may introduce self-selection bias toward hyper-users or tech-savvy individuals. Consequently, future research should prioritize the development and validation of psychometric instruments to quantitatively test these newly identified AI-driven JD-R pathways within larger, representative populations. Furthermore, longitudinal designs and cross-cultural comparative studies are strongly recommended to continuously track the enduring impacts of algorithmic integration on occupational well-being, evaluate joint human–AI system performance under both optimal and misleading AI conditions, and elucidate how varying organizational support architectures globally buffer these complex technological demands.

Conclusions

This qualitative study provides a nuanced theoretical unpacking of the double-edged sword effect of AI in clinical nursing through the lens of the JD-R model. While AI acts as an empowering job resource by facilitating perceived efficiency gains and clinical cognitive empowerment, it simultaneously introduces notable digital job demands. These demands manifest primarily as cognitive impediments stemming from verification overload, relational attrition challenging the empathetic core of nursing, and the psychological exhaustion induced by digital involution. Crucially, our findings illustrate that the trajectory of AI integration is not technologically deterministic. Instead, the lived interplay between AI-driven resources and demands is contextualized by key perceived buffering conditions: nurses' proactive coping strategies, rigorous professional boundary demarcation, and the provision of a robust organizational support architecture. To prevent AI from devolving into an occupational hazard, health care administrators must transcend the mere deployment of algorithms. By establishing comprehensive life cycle AI governance, providing advanced AI literacy training, and safeguarding the irreplaceable humanistic essence of the profession, organizations can sustainably harness AI to empower the nursing workforce and elevate patient care in the digital era.

Acknowledgments

The authors disclose that DeepSeek (DeepSeek AI, Hangzhou, China) was the only generative AI tool used, and it was used solely for language editing and readability improvement during the preparation of this manuscript. Specifically, DeepSeek was used to refine language expression, improve readability, and polish the English writing. No AI-generated content was incorporated into the final manuscript. All AI-assisted outputs were critically reviewed, fact-checked, and substantially revised by the authors. The authors take full responsibility for the accuracy, originality, integrity, and all content in the manuscript, including all references and citations.

XSL•FO
RenderX

Funding

This study was funded by the National High Level Hospital Clinical Research Funding (BJ-2024-198) and the 2024 Research Project of the Chinese Nursing Association (ZHKYQ202416).

Conflicts of Interest

None declared.

Multimedia Appendix 1
The Job Demands‐Resources model.
[DOC File, 119 KB - nursing_v9i1e100916_app1.doc]

Multimedia Appendix 2
Interview outline based on the Job Demands-Resources model.
[DOC File, 26 KB - nursing_v9i1e100916_app2.doc]

Multimedia Appendix 3
Participant characteristics (n=34).
[DOC File, 130 KB - nursing_v9i1e100916_app3.doc]

Multimedia Appendix 4
Participant AI exposure matrix.
[DOC File, 207 KB - nursing_v9i1e100916_app4.doc]

Multimedia Appendix 5
Thematic coding tree based on the Job Demands-Resources model.
[DOC File, 320 KB - nursing_v9i1e100916_app5.doc]

References

1. Ronquillo CE, Peltonen LM, Pruinelli L, et al. Artificial intelligence in nursing: priorities and opportunities from an international invitational think-tank of the Nursing and Artificial Intelligence Leadership Collaborative. J Adv Nurs 2021 Sep;77(9):3707-3717. [doi: 10.1111/jan.14855] [Medline: 34003504]
2. Rony MKK, Parvin MR, Ferdousi S. Advancing nursing practice with artificial intelligence: enhancing preparedness for the future. Nurs Open 2024 Jan;11(1). [doi: 10.1002/nop2.2070] [Medline: 38268252]
3. Zhang GW, Gong M, Li HJ, Wang S, Gong DX. The “Trinity” smart hospital construction policy promotes the development of hospitals and health management in China. Front Public Health 2023;11:1219407. [doi: 10.3389/fpubh.2023.1219407] [Medline: 37546298]
4. Edelson DP, Churpek MM, Carey KA, et al. Early warning scores with and without artificial intelligence. JAMA Netw Open 2024 Oct 1;7(10):e2438986. [doi: 10.1001/jamanetworkopen.2024.38986] [Medline: 39405061]
5. Hassanein S, El Arab RA, Abdrbo A, et al. Artificial intelligence in nursing: an integrative review of clinical and operational impacts. Front Digit Health 2025;7:1552372. [doi: 10.3389/fdgth.2025.1552372] [Medline: 40124108]
6. Kang R, Xuan Z, Tong L, Wang Y, Jin S, Xiao Q. Nurse researchers’ experiences and perceptions of generative AI: qualitative semistructured interview study. J Med Internet Res 2025 Aug 25;27:e65523. [doi: 10.2196/65523] [Medline: 40853413]
7. Lansdowne M, Sperling J, Lytle KS, Allen D, Merrill T, Cary MP Jr. Nurses’ perspectives on AI in clinical practice: implications for training and workforce development. Nurs Adm Q 2026;50(2):95-100. [doi: 10.1097/NAQ.0000000000000710] [Medline: 40971699]
8. Liu X, Chen Y, Guan W, et al. The mediation of AI trust on AI uncertainties and AI competence among nurses: a cross-sectional study. J Adv Nurs 2026 Jun;82(6):6063-6071. [doi: 10.1111/jan.70250] [Medline: 40995865]
9. Scott AJS, Zhao Q, Pan JF, Brown BC, Dowding D. Nurses’ experiences using AI in clinical practice: systematic review. JMIR Nurs 2026 Jun 25;9:e91238. [doi: 10.2196/91238] [Medline: 42347840]
10. Kapoor V, Skellern C, McGrath P, Brosnan S, Gofer Y, Thorburn L. Improving clinical efficiency using artificial intelligence scribe for documentation (ICE-AID study): a retrospective-prospective cohort study at a quaternary children’s hospital and health service. J Paediatr Child Health 2026 May 27. [doi: 10.1111/jpc.70419] [Medline: 42200293]
11. Joo JY, Liu MF, Ho MH. Nurses’ perceptions of artificial intelligence adoption in healthcare: a qualitative systematic review. Nurse Educ Pract 2025 Oct;88:104542. [doi: 10.1016/j.nepr.2025.104542] [Medline: 40925173]

XSL•FO
RenderX

12. Shen M, Shen Y, Liu F, Jin J. Prompts, privacy, and personalized learning: integrating AI into nursing education-a qualitative study. BMC Nurs 2025 Apr 29;24(1):470. [doi: 10.1186/s12912-025-03115-8] [Medline: 40301862]
13. Wong A, Otles E, Donnelly JP, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. JAMA Intern Med 2021 Aug 1;181(8):1065-1070. [doi: 10.1001/jamainternmed.2021.2626] [Medline: 34152373]
14. Jia J, Ning X, Liu W. The consequences and theoretical explanation of workplace AI on employees: a systematic literature review. J Digit Manag 2025;1:14. [doi: 10.1007/s44362-025-00016-3]
15. Yunis M, Jamali D, Umar M. The illusion of AI mastery: “skill inflation” in the age of accessible AI tools. J Innov Knowl 2026 Nov;18:101084. [doi: 10.1016/j.jik.2026.101084]
16. Bakker AB, Demerouti E. The Job Demands‐Resources model: state of the art. J Manag Psychol 2007 Apr 3;22(3):309-328. [doi: 10.1108/02683940710733115]
17. Li S, Meng B, Wang Q. The double-edged sword effect of relational crafting on job well-being. Front Psychol 2022;13:713737. [doi: 10.3389/fpsyg.2022.713737] [Medline: 35250699]
18. Steffensen DS Jr, McAllister CP, Perrewé PL, Wang G, Brooks CD. “You’ve got mail”: a daily investigation of email demands on job tension and work-family conflict. J Bus Psychol 2022;37(2):325-338. [doi: 10.1007/s10869-021-09748-1] [Medline: 33867662]
19. Catan-Inan F. Artificial intelligence literacy and readiness in future health care professionals: a cross-sectional study. Croat Med J 2026 Mar 20;67(1):4-13. [doi: 10.3325/cmj.2026.67.4] [Medline: 41866897]
20. O’Connor S, Vercell A, Wong D, et al. The application and use of artificial intelligence in cancer nursing: a systematic review. Eur J Oncol Nurs 2024 Feb;68:102510. [doi: 10.1016/j.ejon.2024.102510] [Medline: 38310664]
21. Zhi Z, Zhao J, Li Q, et al. Evolving perceptions and attitudes to adopting generative AI in professional settings: multicenter longitudinal qualitative study of senior Chinese hospital leaders. J Med Internet Res 2025 Jun 27;27:e75531. [doi: 10.2196/75531] [Medline: 40577803]
22. Shah SV. Accuracy, consistency, and hallucination of large language models when analyzing unstructured clinical notes in electronic medical records. JAMA Netw Open 2024 Aug 1;7(8):e2425953. [doi: 10.1001/jamanetworkopen.2024.25953] [Medline: 39136951]
23. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
24. O’Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
25. Chaudhry B, Wang J, Wu S, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med 2006 May 16;144(10):742-752. [doi: 10.7326/0003-4819-144-10-200605160-00125] [Medline: 16702590]
26. Noy S, Zhang W. Experimental evidence on the productivity effects of generative artificial intelligence. Science 2023 Jul 14;381(6654):187-192. [doi: 10.1126/science.adh2586] [Medline: 37440646]
27. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019 Jan;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
28. Shortliffe EH, Sepúlveda MJ. Clinical decision support in the era of artificial intelligence. JAMA 2018 Dec 4;320(21):2199-2200. [doi: 10.1001/jama.2018.17163] [Medline: 30398550]
29. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on the domains of nursing: scoping review. JMIR Nurs 2020 Dec 17;3(1):e23939. [doi: 10.2196/23939] [Medline: 34406963]
30. Atalla ADG, El-Gawad Mousa MA, Hashish EAA, Elseesy NAM, Abd El Kader Mohamed AI, Sobhi Mohamed SM. Embracing artificial intelligence in nursing: exploring the relationship between artificial intelligence-related attitudes, creative self-efficacy, and clinical reasoning competency among nurses. BMC Nurs 2025 Jun 20;24(1):661. [doi: 10.1186/s12912-025-03306-3] [Medline: 40542340]
31. Berdida DJE, Grande RAN, Serag RM, Abd Elmaksoud DMF. Nursing students’ artificial intelligence (AI) literacy, AI self-efficacy and AI self-competency: a cross-sectional design and structural equation model analysis. Nurse Educ Pract 2026 Jan;90:104673. [doi: 10.1016/j.nepr.2025.104673] [Medline: 41371196]
32. Lin SY, Mahoney MR, Sinsky CA. Ten ways artificial intelligence will transform primary care. J Gen Intern Med 2019 Aug;34(8):1626-1630. [doi: 10.1007/s11606-019-05035-1] [Medline: 31090027]
33. Chuang YT, Chiang HL, Lin AP. Insights from the job demands–resources model: AI’s dual impact on employees’ work and life well-being. Int J Inf Manage 2025 Aug;83:102887. [doi: 10.1016/j.ijinfomgt.2025.102887]
34. Chen F, Li Y, Chen Y, et al. Strategies for the analysis and elimination of hallucinations in artificial intelligence generated medical knowledge. J Evid Based Med 2025 Sep;18(3):e70075. [doi: 10.1111/jebm.70075] [Medline: 40983876]
35. Morey DA, Rayo MF, Woods DD. Empirically derived evaluation requirements for responsible deployments of AI in safety-critical settings. NPJ Digit Med 2025 Jun 18;8(1):374. [doi: 10.1038/s41746-025-01784-y] [Medline: 40533485]
36. Kim JY, Hasan A, Kueper J, et al. Establishing organizational AI governance in healthcare: a case study in Canada. NPJ Digit Med 2025;8(1):522. [doi: 10.1038/s41746-025-01909-3] [Medline: 40817280]

XSL•FO
RenderX

Abbreviations

DQCA: directed qualitative content analysis
JD-R: Job Demands-Resources
LLM: large language model
SRQR: Standards for Reporting Qualitative Research

Edited by E Borycki; submitted 10.May.2026; peer-reviewed by PB Dunlap, W Dong; revised version received 04.Aug.2026; accepted 07.Sep.2026; published 02.Oct.2026.

Please cite as:
Zhang X, Fang J, Chen S, Luo J
The Double-Edged Sword Effect of AI Application Among Clinical Nurses Based on the Job Demands-Resources Model: Qualitative Study
JMIR Nursing 2026;9:e100916
URL: https://nursing.jmir.org/2026/1/e100916
doi:10.2196/100916

© Xiaoyan ZHANG, Jiaxin Fang, Sihan Chen, Jiayin Luo. Originally published in JMIR Nursing (https://nursing.jmir.org), 2.Oct.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Performance of Large Language Models in the Japanese Public Health Nurse National Examination: Comparative Cross-Sectional Study

Yutaro Takahashi*, MHSc; Ryota Kumakura*, MHSc; Rie Okamoto, PhD; Shizuko Omote, PhD

Faculty of Health Sciences, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kodatsuno 5-11-80, Kanazawa, Ishikawa, Japan

*these authors contributed equally

Corresponding Author:
Yutaro Takahashi, MHSc
Faculty of Health Sciences, Institute of Medical, Pharmaceutical and Health Sciences, Kanazawa University, Kodatsuno 5-11-80, Kanazawa, Ishikawa, Japan

Abstract

Background: Large language models (LLMs) have shown promising results on Japanese national medical and nursing examinations. However, no study has evaluated LLM performance on the Japanese Public Health Nurse National Examination, which requires specialized knowledge in community health and public health nursing practice.

Objective: This study aimed to compare the performance of multiple LLMs on the Japanese Public Health Nurse National Examination and evaluate their potential utility in public health nursing education.

Methods: Three LLMs were evaluated: GPT-4o, Claude Opus 4, and Gemini 2.5 Pro. All 110 questions from the 111th Public Health Nurse National Examination were administered using standardized prompts. Questions were classified by format (text vs figure or calculation), content (general vs situational), and selection type (single vs multiple choice). Accuracy rates and 95% CIs were calculated, with statistical comparisons performed using chi-square tests.

Results: All LLMs exceeded the passing criterion (60%). The accuracy rates were as follows: 85.5% (94/110) for GPT-4o (95% CI 77.5% - 91.5%), 91.8% (101/110) for Claude Opus 4 (95% CI 85.0% - 96.2%), and 92.7% (102/110) for Gemini 2.5 Pro (95% CI 86.2% - 96.8%). No significant differences were found among the LLMs (P>.99). However, all models showed lower accuracy on multiple-choice questions than on single-choice questions, with significant intramodel differences observed for GPT-4o (10/16, 62.5% vs 82/92, 89.1%; P=.01) and Claude Opus 4 (12/16, 75% vs 87/92, 94.6%; P=.03).

Conclusions: LLMs demonstrated high performance on a public health nursing examination but showed limitations in complex reasoning requiring multiple-choice selection. These findings suggest the potential for LLM use as educational support tools while highlighting the need for cautious implementation in specialized nursing education.

(JMIR Nursing 2026;9:e82842) doi:10.2196/82842

KEYWORDS

large language models; public health nursing; licensure; nursing; artificial intelligence; AI; education

Introduction

With advances in artificial intelligence (AI), large language models (LLMs) have gained attention in various fields. High-performance LLMs have been developed, including GPT-3.5, GPT-4, Anthropic's Claude, and Google's Gemini [1-3], which have demonstrated their ability to generate contextually appropriate responses to complex questions. In Japan, the capabilities of these models have been evaluated using questions from the national medical and nursing examinations. Liu et al [4] compared GPT-4o, GPT-4, Claude Opus 3, and Gemini 1.5 Pro using Japan's national medical examinations and found that GPT-4o achieved the highest accuracy rate (89.2%) and intermodel performance depending on the subject area and question format. Takagi et al [5] also found that GPT-4 significantly outperformed GPT-3.5 in national medical examinations (79.9% vs 50.8%), revealing significant performance gaps between model generations. Taira et al [6] reported that LLM accuracy declined for questions about pharmacology, social welfare, and related legal regulations on national nursing examinations, suggesting limitations in responding to questions requiring specialized knowledge and institutional understanding. These findings demonstrate LLM utility and associated challenges in Japanese language processing and professional examinations.

However, no study has compared the performance of multiple LLMs on the Japanese Public Health Nurse National Examination, which assesses the knowledge and skills necessary

XSL•FO
RenderX

for public health nursing practice. The Public Health Nurse National Examination comprises content addressing community-based health issues, multidisciplinary collaboration, and public health nursing activities, requiring not only medical knowledge but also an understanding of Japan's community health systems [7]. Compared to clinical medicine, public health nursing presents unique challenges for LLMs. While medical examinations primarily assess biomedical knowledge and clinical reasoning, public health nursing requires the integration of social determinants of health, public health policy knowledge, community-level interventions, and understanding of local health systems. These multifaceted aspects demand complex reasoning that simultaneously considers multiple factors, which may pose greater challenges for current LLMs. Therefore, this study aimed to compare and evaluate the performance of GPT-4o (Open AI), Claude Opus 4 (Anthropic), and Gemini 2.5 Pro (Google AI) on questions from the Japanese Public Health Nurse National Examination to clarify the extent to which AI can respond to the specialized knowledge and skills necessary for public health nursing. We believe our work has significance for examining the potential for AI use in future public health nursing education.

Methods

Study Design

This comparative cross-sectional study evaluated the performance of 3 LLMs on the Japanese Public Health Nurse National Examination. We used a census sampling approach, analyzing all questions from the examination to compare accuracy rates across different LLMs and question types.

Tested LLMs

Three representative LLMs were selected: OpenAI's GPT-4o [1], Anthropic's Claude Opus 4 [2], and Google's Gemini 2.5 Pro [3].

Japanese Public Health Nurse National Examination

The Japanese Public Health Nurse National Examination is conducted based on the Act on Public Health Nurses, Midwives, and Nurses to assess the knowledge and skills necessary for public health nurses [8]. It includes 110 questions. The questions are classified as general and situational. General questions are worth 1 point each, whereas situational questions are worth 2 points each. The passing criterion is 60% correct answers for general and situational questions combined, although the percentage may be adjusted when inappropriate questions are excluded. The questions are in multiple-choice format, and the participants must select one or more correct answers from 4 to 5 options. The pass rate is 90% to 95%.

Study Population and Question Selection

This study used a census sampling approach, including all 110 questions from the 111th Public Health Nurse National Examination administered in February 2025 [9]. No sampling was conducted as the entire population of examination questions was analyzed. For questions containing figures and tables, figures were processed as image data, whereas the question text was handled as text. No questions were designated as inappropriate by the Ministry of Health, Labour and Welfare; the passing criterion was 60% [9]. The examination questions were classified according to format (text questions and figure or calculation questions), content (general and situational questions), and selection type (single- and multiple-choice questions).

Prompt Engineering

Because prompt engineering significantly affects the generated output, the question input formats were standardized. On the basis of previous research [4-6], 6 prompts corresponding to different types of questions were created (Table 1).

XSL•FO
RenderX

Table . Standardized prompts for different question types in the Japanese Public Health Nurse National Examination. Template placeholders ("<" and ">") indicate where specific content was inserted for each question type.

Prompt type	Prompt template
Prompt 1: general questions	"Japanese Public Health Nurse National Examination questions are presented. Please answer the following question in brief by selecting an option. Select one option unless otherwise specified. Question: <Question content> 1. <Option 1> 2. <Option 2> 3. <Option 3> 4. <Option 4> 5. <Option 5> (if applicable)"
Prompt 2: situational questions	"Japanese Public Health Nurse National Examination questions are presented. Based on the following situational setting, please answer the question in brief by selecting an option. Select one option unless otherwise specified. Situation: <Situational content> Question: <Question content> 1. <Option 1> 2. <Option 2> 3. <Option 3> 4. <Option 4> 5. <Option 5> (if applicable)"
Prompt 3: image questions	"Japanese Public Health Nurse National Examination questions are presented. Please review the following image and answer the question in brief by selecting an option. Select one option unless otherwise specified. Question: <Question content> 1. <Option 1> 2. <Option 2> 3. <Option 3> 4. <Option 4> 5. <Option 5> (if applicable)"
Prompt 4: situational questions with images	"Japanese Public Health Nurse National Examination questions are presented. Based on the following situational setting, please review the image and answer the question in brief by selecting an option. Select one option unless otherwise specified. Situation: <Situational content> Question: <Question content> 1. <Option 1> 2. <Option 2> 3. <Option 3> 4. <Option 4> 5. <Option 5> (if applicable)"
Prompt 5: calculation questions	"Japanese Public Health Nurse National Examination questions are presented. Please read the following question, review the mark sheet format in the image, and answer the question in brief by selecting an option. Provide your answer as a numerical value in brief. Question: <Question content>"
Prompt 6: situational calculation questions	"Japanese Public Health Nurse National Examination questions are presented. Please read the following situational setting and question, review the mark sheet format in the image, and answer the calculation question. Provide your answer as a numerical value in brief. Situation: <Situational content> Question: <Question content>"

Data Collection Procedures

Questions were input from June 25 to 26, 2025. The question text and images were directly inserted into each LLM's chat window. Each question was input in a new independent chat window to avoid potential influence from previous responses. All questions were administered in Japanese using standardized prompts.

The definition of "correct" answers was based on the official answers published by the Ministry of Health, Labour, and Welfare [9]. Only answers that clearly matched the official correct answers and followed the instructions provided in the question text were considered "correct." Ambiguous answers, evident mistakes, unclear responses, and responses with an excessive number of answer choices were considered incorrect. All responses from the LLMs were independently reviewed and scored by 2 authors (YT and RK), with any discrepancies resolved through discussion.

Data Analysis

For each LLM, the number of correct answers, accuracy rates, 95% CIs, total scores, and score rates were calculated. Accuracy rates were compared across question format (text vs figure or calculation), question content (general vs situational), and selection type (single vs multiple choice) both between LLMs (inter-LLM comparison) and within each LLM (intra-LLM comparison). Numerical input calculation questions were excluded from the selection type analysis.

Statistical comparisons were performed using chi-square tests for expected cell frequencies of ≥5 and the Fisher exact test when expected cell frequencies were <5. For multiple pairwise inter-LLM comparisons, the Bonferroni correction was applied to control for type I error. Statistical significance was set at $P \leq .05$ (2 tailed). All statistical analyses were conducted using Stata (version 18.0; StataCorp LLC).

Ethical Considerations

Ethics approval was not required because only data from a published database were analyzed.

Results

Detailed results are presented in Table 2. The accuracy rates for each LLM for all 110 questions were as follows: 85.5% (n=94) of the questions for GPT-4o (95% CI 77.5% - 91.5%), 91.8% (n=101) of the questions for Claude Opus 4 (95% CI 85.0% - 96.2%), and 92.7% (n=102) of the questions for Gemini 2.5 Pro (95% CI 86.2% - 96.8%). The corresponding scores

XSL•FO
RenderX

were 86.2% (125/145), 91.7% (133/145), and 93.1% (135/145), respectively, all of which exceeded the passing criterion (60%). In terms of question characteristics, the accuracy rates for general questions (n=75) were as follows: 84% (n=63) for GPT-4o, 92% (n=69) for Claude Opus 4, and 92% (n=69) for Gemini 2.5 Pro. The corresponding rates for situational questions (n=35) were 88.6% (n=31), 91.4% (n=32), and 94.3% (n=33), respectively.

Table . Performance of large language models (LLMs) on the 111th Public Health Nurse National Examination.

Category	GPT-4o		Claude Opus 4		Gemini 2.5 Pro		LLM comparison *P* value		
	Correct answers, n (%)	*P* value	Correct answers, n (%)	*P* value	Correct answers, n (%)	*P* value	GPT-4o vs Claude Opus 4	GPT-4o vs Gemini	Claude Opus 4 vs Gemini
Overall accuracy rate (n=110)	94 (85.5)	—[a]	101 (91.8)	—	102 (92.7)	—	>.99	>.99	>.99
Overall score (n=145)	125 (86.2)	—	133 (91.7)	—	135 (93.1)	—	>.99	>.99	>.99
By question content		.53		>.99		>.99			
General questions (n=75)	63 (84.0)		69 (92.0)		69 (92.0)		>.99	>.99	>.99
Situational questions (n=35)	31 (88.6)		32 (91.4)		33 (94.3)		>.99	>.99	>.99
By question format		.69		>.99		.60			
Text questions (n=98)	84 (85.7)		90 (91.8)		90 (91.8)		>.99	>.99	>.99
Figure or calculation questions (n=12)	10 (83.3)		11 (91.7)		12 (100.0)		>.99	>.99	>.99
By selection type		.04		.03		.09			
Single-choice questions (n=92)	82 (89.1)		87 (94.6)		87 (94.6)		>.99	>.99	>.99
Multiple-choice questions (n=16)	10 (62.5)		12 (75.0)		13 (81.3)		>.99	>.99	>.99

[a]Not applicable.

In terms of question format, the accuracy rates for the text questions (n=98) were as follows: 85.7% (n=84) for GPT-4o, 91.8% (n=90) for Claude Opus 4, and 91.8% (n=90) for Gemini 2.5 Pro. The corresponding rates for figure or calculation questions (n=12) were 83.3% (n=10), 91.7% (n=11), and 100% (n=12), respectively. The accuracy rates for single-choice questions (n=92) were as follows: 89.1% (n=82) for GPT-4o, 94.6% (n=87) for Claude Opus 4, and 94.6% (n=87) for Gemini 2.5 Pro. However, for the multiple-choice questions (n=16), all LLMs showed decreased accuracy, and the corresponding rates were 62.5% (n=10), 75% (n=12), and 81.3% (n=13), respectively.

Statistical comparisons among the LLMs showed no significant differences. Intra-LLM comparisons revealed significant differences between single- and multiple-choice questions for GPT-4o (*P*=.01) and Claude Opus 4 (*P*=.03), with the accuracy rates for multiple-choice questions being significantly lower.

Discussion

Principal Findings

This study compared and evaluated the performances of multiple LLMs on the Japanese Public Health Nurse National Examination. All the LLMs significantly exceeded the passing criterion of 60%. These results indicate that LLMs have acquired considerable specialized knowledge, with the performance of GPT-4o (94/110, 85.5%) being comparable to that of LLMs in previous medical examinations [4] and superior to that of older-generation models [6].

While all LLMs showed high overall accuracy rates, a clear performance decline was observed for multiple-choice questions. This contrasts with the high accuracy rates for single-choice questions, with significant differences found for GPT-4o and Gemini 2.5 Pro. This phenomenon demonstrates the current limitations of LLMs in complex reasoning, which requires the

XSL•FO
RenderX

simultaneous evaluation of multiple concepts. As public health nursing practice requires comprehensive judgment considering multiple factors such as regional characteristics, residents' needs, social resources, and multidisciplinary collaboration, the performance decline on multiple-choice questions indicates that LLMs have limitations in the complex decision-making faced in actual public health nursing practice.

The accuracy rates of the LLMs evaluated in this study (94/110, 85.5% to 102/110, 92.7%) and their scores (125/145, 86.2% to 135/145, 93.1%) substantially exceeded the passing standard (60%). The pass rate for the 111th Public Health Nurse National Examination was 94% [9]; however, this represents the proportion of examinees who exceeded the passing standard, and the overall mean score for all examinees is not publicly available. Therefore, a direct comparison between the overall academic performance of examinees and LLM performance is challenging.

In the field of medical education, multiple studies comparing LLM and student performance have been reported. A study comparing final-year emergency medicine students with AI models [10] demonstrated that students achieved a 79.4% accuracy rate, outperforming ChatGPT (72.5%) and Gemini (54.4%). The superiority of students was particularly pronounced in image-based questions, highlighting current limitations in AI models' visual information processing capabilities. Additionally, a study using 1070 medical imaging questions [11] found that GPT-4 correctly answered 67.8% of the questions it attempted, whereas the students' passing mean was 63%. However, the student majority vote achieved a 94.5% accuracy rate, substantially surpassing the AI. This demonstrates that even when individual students' abilities may be equal to or slightly inferior to those of AI, collective student judgment significantly exceeds AI performance.

In our study, while LLM performance on figure or calculation questions was high (10/12, 83.3% to 12/12, 100%), the small number of questions (n=12) necessitates larger-scale validation. More importantly, the learning processes of LLMs and humans are fundamentally different. LLMs learn patterns from large volumes of text data, whereas public health nurses acquire decision-making capabilities by integrating practical experience with theoretical knowledge. Furthermore, human public health nurses possess essential practical competencies that are not measurable through written examinations, including ethical judgment, empathy, and interpersonal communication skills. As these previous studies [10,11] demonstrate, while AI shows potential as a supplementary educational tool, it cannot replace human capabilities, particularly in areas requiring visual interpretation, clinical reasoning, and collective judgment. Therefore, LLMs should be appropriately positioned as educational and learning support tools rather than as replacements for human public health nurses.

The results of this study have important implications from the perspective of competency development in public health nursing education. The "Practical Competencies Required of Public Health Nurses and Achievement Goals and Levels at Graduation" document by the Japanese Ministry of Health, Labour and Welfare [12] classifies public health nurse competencies into five domains: (1) ability to clarify community health issues and develop plans; (2) ability to provide continuous support and collaborative organizational activities for individuals, families, groups, and organizations to enhance community health promotion capacity and evaluate these activities; (3) community health crisis management capacity; (4) ability to develop projects, policies, social resources, and systems to enhance community health levels; and (5) professional autonomy and continuous quality improvement capacity. LLMs may be particularly effective in providing learning support during information gathering and assessment stages within domain 1. This competency includes information collection for clarifying community health issues, community diagnosis, and prioritization of health issues, where LLMs are expected to play a supplementary role in confirming foundational knowledge and organizing information.

For domain 1, the achievement level at graduation is set at either level 1 ("able to implement independently with minimal guidance") or level 2 ("able to implement under supervision [from supervising public health nurses or faculty]") [12]. As revealed in this study, LLMs demonstrated a performance decline in multiple-choice questions and have limitations in complex judgment tasks. Therefore, when using LLMs as educational tools, it is crucial to cultivate students' ability to critically evaluate LLM outputs and maintain practical judgment skills based on community characteristics to reach these achievement levels.

Several ethical considerations must be addressed when using LLMs in public health nursing education. The Japanese Ministry of Education, Culture, Sports, Science and Technology guidelines [13] indicate that directly using generative AI outputs does not deepen students' own learning, that differences in generative AI types (paid vs free versions) may create disparities in student outcomes leading to unfairness, and that confidential and personal information may be unintentionally leaked or disclosed. A survey of Japanese medical students [14] found that while 41.9% had experience using ChatGPT, only 10.2% had used it for medical assignments and 47% held negative views about its use for medical reports. Many students felt that, considering the time required to verify AI responses, independent learning would be more efficient, highlighting the essential need to cultivate critical evaluation skills for LLM outputs. A narrative review on chatbot integration in nursing education [15] also emphasizes the importance of ethical considerations, indicating the urgent need to establish ethical frameworks for AI use across nursing education.

The results of this study revealed that while LLMs demonstrated high accuracy rates on the Public Health Nurse National Examination, performance declined on multiple-choice questions. This finding has important implications for using LLMs as learning support tools in public health nursing education. Given the demonstrated limitations of LLMs in complex judgment requiring simultaneous consideration of multiple factors, LLMs are suitable for supplementary roles such as confirming foundational knowledge and gathering information, whereas faculty instruction remains crucial for learning scenarios requiring complex judgment. The Ministry of Education, Culture, Sports, Science and Technology has

XSL•FO
RenderX

issued guidelines [13] on the educational use of generative AI at universities and colleges of technology, and public health nurse training institutions are also called upon to develop guidelines that clearly specify appropriate use scenarios and limitations for LLMs. A nationwide survey on information and communications technology (ICT) use among public health nurses in local governments [16] found that 82.8% responded that they did not know the procedures for promoting ICT use, indicating challenges in adapting to digital technology even among practicing public health nurses. This suggests the importance of providing systematic digital literacy education from the public health nurse training stage.

A detailed examination is needed regarding curriculum integration and faculty training. For example, development of specific implementation strategies is required, including in which courses and how LLMs should be introduced, how to design a phased introduction process, and how faculty should learn appropriate LLM use methods. In particular, establishing organizational training systems for enhancing faculty AI literacy and developing assessment methods premised on LLM use are important future challenges. To accumulate knowledge regarding these strategies, pilot program implementation and evaluation will be necessary.

As practical implications, there is potential for the use of LLMs as continuing education and self-directed learning support tools for practicing public health nurses. A concept analysis of LLMs in nursing education [17] positions LLMs as transformative tools that provide accessible and personalized learning support and promote cognitive and skill development. In public health nursing education as well, use is anticipated for responding to new public health issues and during information gathering stages in community diagnosis. Additionally, there is potential for the use of LLMs as an auxiliary tool in situations requiring rapid information organization, such as during disasters or emerging infectious disease outbreaks. However, a survey on ICT use among public health nurses in local governments [16] found that 89.1% of municipalities expressed concerns about individuals who have difficulty adapting to digital technology, necessitating careful introduction that considers the essence of interpersonal support in public health nursing work. Furthermore, while 55.9% in the same survey actively promoted ICT use, only 26.7% perceived progress as smooth, indicating challenges in digital literacy education for practicing public health nurses and establishing organizational support systems. As noted in the aforementioned concept analysis [17], careful attention must be paid to LLM limitations and ethical implications, ensuring that LLM integration aligns with the values and goals of nursing education. Therefore, when using LLMs in practical settings, it is essential to critically evaluate LLM outputs and integrate them with community characteristics and practical knowledge, considering the limitations in complex judgment revealed in this study.

The findings of this study occupy an important position within the broader context of AI use in health profession education. A narrative review on chatbots in nursing education [15] demonstrated a surge in related research from 2021 to 2023 (with 2023 accounting for 70% of publications), indicating growing scholarly interest in this field. Together with LLM evaluation studies on medical licensing examinations [4,5], it is becoming increasingly clear that LLMs demonstrate high performance across medical licensing examinations generally. However, the performance decline on multiple-choice questions demonstrated in this study indicates the existence of LLM limitations in complex judgment unique to public health nursing, such as integration of social determinants of health and planning of community-level interventions. This is consistent with findings from previous studies in medical education [10,11] showing that collective student judgment far exceeds that of AI, supporting the appropriate positioning of LLMs as educational and learning support tools rather than replacements for human health professionals. Additionally, a survey of medical students [14] showed that 47% held negative views about LLM use for medical assignments, indicating the recognition of the need for verification. A concept analysis in nursing education [17] also pointed out the need for careful consideration of LLM limitations and ethical implications. Furthermore, the aforementioned review on chatbot integration in nursing education [15] emphasizes the importance of ethical considerations and the urgency of original research while acknowledging it as a promising field. Collectively, these findings suggest that across health profession education generally, while LLMs have potential as useful auxiliary tools, cultivating the ability to understand their limitations and use them critically is a common challenge across all types of health profession education.

Limitations and Future Directions

This study has the following strengths. First, this is the first study to evaluate the performance of multiple LLMs on the Public Health Nurse National Examination. While LLM evaluations have been conducted on medical and nursing licensing examinations, this study represents the first systematic evaluation in the public health nursing field. Second, reproducibility was ensured through the use of standardized prompts. Six prompts were created according to question types, achieving consistent evaluation. Third, detailed analysis by question format revealed the important finding of performance decline on multiple-choice questions. This discovery demonstrates LLM limitations in complex reasoning requiring the simultaneous consideration of multiple factors, with important implications for future educational implementation.

This study has several limitations. First, as a cross-sectional evaluation of a single year, temporal changes in LLM performance and reproducibility across different examination years could not be evaluated in this study. Second, accuracy rates alone do not clarify the quality of reasoning processes or correlation with actual public health nursing practice competencies. Third, while LLM versions and settings may influence results, this study was limited to evaluation using specific versions (GPT-4o, Claude Opus 4, and Gemini 2.5 Pro). Fourth, results may vary depending on prompt expression methods, and there is no guarantee that the standardized prompts used in this study are optimal. Fifth, because the overall mean score for examinees is not publicly available, direct performance comparison between LLMs and human public health nurse examinees is difficult. Furthermore, this study verified LLM performance in the educational evaluation context of a national

XSL•FO
RenderX

examination and did not evaluate their utility or safety as decision support tools in actual public health nursing practice. Future research should include continued evaluation over multiple years, qualitative analysis of reasoning processes, validation of utility in practical settings, comparative studies with human public health nurses, and examination of applicability to decision support in actual practice.

Conclusions

The LLMs evaluated demonstrated high performance on the Public Health Nurse National Examination; however, they also had limitations in solving problems requiring complex judgment. These findings provide important foundational data showing the possibilities and challenges of AI use in public health nursing. On the basis of these results, LLMs should be cautiously used as supplementary tools in public health nursing education.

Funding

This work was supported by a Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (grant 25K14186).

Data Availability

The data supporting the findings of this study are available from the corresponding author (YT) upon request.

Authors' Contributions

YT contributed to conceptualization, methodology, investigation, funding acquisition, supervision, and manuscript review and editing. RK contributed to methodology, investigation, data curation, formal analysis, and writing of the original draft. RO and SO contributed to manuscript review and editing.

Conflicts of Interest

None declared.

References

1. Hello GPT-4o. OpenAI. 2024. URL: https://openai.com/index/hello-gpt-4o/ [accessed 2025-08-19]
2. Introducing Claude 4. Anthropic. 2025. URL: https://www.anthropic.com/news/claude-4 [accessed 2025-08-19]
3. Gemini 2.5 Pro. Google Cloud. 2025. URL: https://cloud.google.com/vertex-ai/generative-ai/docs/models/gemini/2-5-pro [accessed 2025-08-19]
4. Liu M, Okuhara T, Dai Z, et al. Evaluating the effectiveness of advanced large language models in medical knowledge: a comparative study using Japanese national medical examination. Int J Med Inform 2025 Jan;193:105673. [doi: 10.1016/j.ijmedinf.2024.105673] [Medline: 39471700]
5. Takagi S, Watari T, Erabi A, Sakaguchi K. Performance of GPT-3.5 and GPT-4 on the Japanese Medical Licensing Examination: comparison study. JMIR Med Educ 2023 Jun 29;9:e48002. [doi: 10.2196/48002] [Medline: 37384388]
6. Taira K, Itaya T, Hanada A. Performance of the large language model ChatGPT on the National Nurse Examinations in Japan: evaluation study. JMIR Nurs 2023 Jun 27;6:e47305. [doi: 10.2196/47305] [Medline: 37368470]
7. About the "National Examination Question Standards for Public Health Nurses, Midwives, and Nurses, 2023 Edition". Ministry of Health, Labor and Welfare Japan. 2022. URL: https://www.mhlw.go.jp/stf/shingi2/0000159020_00001.html [accessed 2025-08-19]
8. Act on Public Health Nurses, Midwives, and Nurses. Japanese Law Translation. 2009. URL: https://www.japaneselawtranslation.go.jp/ja/laws/view/3993 [accessed 2025-08-19]
9. Announcement of results for the 111th National Public Health Nurse Examination, the 108th National Midwife Examination, and the 114th National Nurse Examination. Ministry of Health, Labour and Welfare Japan. 2025. URL: https://www.mhlw.go.jp/general/sikaku/successlist/2025/siken03_04_05/about.html [accessed 2025-08-19]
10. Al-Thani SN, Anjum S, Bhutta ZA, et al. Comparative performance of ChatGPT, Gemini, and final-year emergency medicine clerkship students in answering multiple-choice questions: implications for the use of AI in medical education. Int J Emerg Med 2025 Aug 7;18(1):146. [doi: 10.1186/s12245-025-00949-6] [Medline: 40775272]
11. Roos J, Martin R, Kaczmarczyk R. Evaluating Bard Gemini Pro and GPT-4 Vision against student performance in medical visual question answering: comparative case study. JMIR Form Res 2024 Dec 17;8:e57592. [doi: 10.2196/57592] [Medline: 39714199]
12. Guidelines for the management of nursing schools revised. Ministry of Health, Labour and Welfare Japan. 2023. URL: https://www.mhlw.go.jp/kango_kyouiku/news/4.html [accessed 2025-08-19]
13. Regarding the educational aspects of generative AI at universities and technical colleges. Ministry of Education, Culture, Sports, Science and Technology Japan. 2023. URL: https://www.mext.go.jp/b_menu/houdou/2023/mext_01260.html [accessed 2025-12-18]

XSL•FO
RenderX

14. Amano I, Obi-Nagata K, Ninomiya A, Fujiwara Y, Koibuchi N. Use and evaluation of generative artificial intelligence by medical students in Japan. JMA J 2025 Jul 2;8(3):730-735. [doi: 10.31662/jmaj.2024-0375] [Medline: 40786463]
15. Zhang F, Liu X, Wu W, Zhu S. Evolution of chatbots in nursing education: narrative review. JMIR Med Educ 2024 Jun 13;10:e54987. [doi: 10.2196/54987] [Medline: 38889074]
16. Akatsuka E, Taguchi A, Yoshida C, Miyagawa S, Sugiyama D. Current status and challenges in the utilization of information and communication technology and digitalization in public health nursing activities: a national survey of supervisory public health nurses in local governments. Nihon Koshu Eisei Zasshi 2025 Oct 9;72(9):606-615. [doi: 10.11236/jph.24-095] [Medline: 40484688]
17. Harrington J, Booth RG, Jackson KT. Large language models in nursing education: concept analysis. JMIR Nurs 2025 Aug 22;8:e77948. [doi: 10.2196/77948] [Medline: 40845300]

Abbreviations

AI: artificial intelligence
ICT: information and communications technology
LLM: large language model

Edited by E Borycki; submitted 22.Aug.2025; peer-reviewed by B Dovdon, C DeBlieck; revised version received 22.Dec.2025; accepted 03.Feb.2026; published 20.Feb.2026.

Please cite as:
Takahashi Y, Kumakura R, Okamoto R, Omote S
Performance of Large Language Models in the Japanese Public Health Nurse National Examination: Comparative Cross-Sectional Study
JMIR Nursing 2026;9:e82842
URL: https://nursing.jmir.org/2026/1/e82842
doi:10.2196/82842

© Yutaro Takahashi, Ryota Kumakura, Rie Okamoto, Shizuko Omote. Originally published in JMIR Nursing (https://nursing.jmir.org), 20.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Theoretical Exploration of Error Thresholds for Clinical AI Decision Support in Nursing: Exploratory Simulation Study Grounded in Human-AI Reliance Data

Hiroyuki Tajima, PhD

Faculty of Nursing, Shumei University, 1-1 Daigaku-cho, Yachiyo, Chiba, Japan

Corresponding Author:
Hiroyuki Tajima, PhD
Faculty of Nursing, Shumei University, 1-1 Daigaku-cho, Yachiyo, Chiba, Japan

Abstract

Background: Clinical AI decision support is being introduced into nursing practice; however, existing large language models (LLMs) demonstrate only moderate accuracy on complex clinical tasks, raising questions about the level of accuracy required for safe clinical use across varying levels of clinician experience and task complexity.

Objective: The aim of the study is to develop an empirically calibrated simulation model of human-AI reliance and error in nursing decision-making and estimate the AI accuracy required to achieve specified error rate targets.

Methods: A linear reliance model with coefficients for AI accuracy (A), clinician experience (E), and task complexity (C) was calibrated using weighted least squares against 9 empirical data points from 3 independent randomized experiments on AI-assisted decision-making (N=3502). Predicted error was computed as reliance×(1−*A*) across a 27-cell factorial design. Study-level bootstrap (2000 iterations) quantified calibration uncertainty. To contextualize the simulation's operating range, the accuracy of contemporary general-purpose LLMs on complex clinical tasks was drawn from published benchmarks.

Results: Calibration placed β_A at 0.201 (bootstrap 95% CI 0.023 - 0.234; $P(\beta_A>0)>.99$). For the novice×high-complexity combination, the minimum AI accuracy values required to achieve error rates <10% and<20% were 0.89 and 0.78, respectively (bootstrap 95% CIs 0.88 - 0.90 and 0.75 - 0.79). At the moderate accuracy levels currently reported for general-purpose LLMs on complex clinical tasks (approximately 0.5 - 0.7), the model predicts error rates of roughly 26% to 41% in this high-risk condition.

Conclusions: In this model, keeping predicted error rates below a stringent target (<10%) for high-complexity nursing decision support by novice clinicians requires AI accuracy of at least approximately 0.89, a level that current general-purpose LLMs may not reliably reach on complex clinical tasks. Because the model is calibrated on nonnursing reliance data, these thresholds are illustrative model outputs, not nursing-derived empirical standards. The $A_{threshold}$ framework provides a decision-theoretic tool for evaluating the minimum AI accuracy requirement by user-and-task profile. Behavioral validation in nursing contexts remains an essential next step. Because the framework is independent of any specific model, it remains applicable as AI systems improve.

(JMIR Nursing 2026;9:e99590) doi:10.2196/99590

KEYWORDS

clinical decision support; automation bias; large language models; nursing informatics; Monte Carlo simulation; error threshold

Introduction

The integration of AI, particularly large language models (LLMs), into nursing education and clinical practice has increased substantially in recent years [1-4]. Although these systems present considerable opportunities to augment clinical reasoning and decision-making, they simultaneously introduce novel risks associated with human reliance on AI-generated outputs [5]. A critical concern is how accurate current AI systems must be for safe use in clinical tasks, particularly because clinical decision support systems can improve care but also introduce safety risks when their outputs are relied upon uncritically [6-8].

Research in human factors and automation science has long demonstrated that users often develop inappropriate levels of trust in automated systems, resulting in overreliance and consequent errors [9,10]. Parasuraman and Riley [9] identified "misuse"—overreliance on automation that produces failures in monitoring or detecting system errors, as a principal failure mode of human-automation interaction. Lee and See [10] subsequently articulated a framework of appropriate reliance in which trust should be calibrated to a system's demonstrated capabilities. In health care specifically, systematic reviews have documented automation bias as a recurring safety concern in computerized clinical decision support, identifying task complexity, cognitive load, and verification demands as

XSL•FO
RenderX

principal mediators of overreliance [11-13]. Despite the established relevance of these frameworks, their quantitative application to health care AI, and to nursing in particular, remains largely underexplored.

Recent empirical studies have begun to characterize human-AI reliance relationships in AI-assisted decision-making. Yin et al [14] conducted 3 preregistered randomized experiments (total n=3793) and reported that participants relied more on AI systems, as the systems' stated and observed accuracy increased, with agreement fractions rising from approximately 0.75 at 60% accuracy to 0.82 at 95% accuracy in their first experiment. Lu and Yin [15] reported a similar monotonic trend in a subsequent study (n=466, 13,980 trials). In clinical settings, Kücking et al [16] demonstrated that AI recommendation correctness exerted a strong bidirectional influence on the diagnostic accuracy of 223 physicians and nurses assessing wound maceration; correct AI recommendations increased the odds of correct decisions approximately 10-fold (odds ratio [OR] 10.0; P<.001), whereas incorrect recommendations significantly reduced accuracy. These findings indicate that automation complacency concerns extend to contemporary AI-based clinical decision support systems and motivate a theoretically grounded, quantitative investigation of the conditions under which AI-assisted clinical errors are most likely to occur.

However, a critical gap remains between the available empirical literature and clinical use decisions. Current studies characterize human-AI reliance in qualitative terms or across a limited range of accuracy levels; yet, they do not provide clinicians or administrators with a decision-theoretic tool for determining whether an AI system, given its measured accuracy, is appropriate for a specific clinical task and user group. Such a tool is increasingly needed because, although LLMs achieve high scores on standardized medical knowledge benchmarks [17], their accuracy on more complex, open-ended clinical decision-making tasks remains only moderate [18,19], placing current systems within an operating range where the clinical safety implications are neither self-evident nor reliably inferable from the raw accuracy value alone.

To address this gap, this study developed a linear simulation model of human-AI reliance in nursing decision-making, empirically calibrated against 9 data points derived from 3 independent randomized experiments on AI-assisted decision-making (total N=3502). Error rates were computed as the product of predicted reliance and AI failure probability and evaluated across a 27-cell factorial design spanning 3 AI accuracy levels, 3 clinician experience levels, and 3 task complexity levels. This work has two primary contributions: (1) empirical calibration of the reliance model from 3 independent large-scale datasets, which avoids the circularity inherent in purely theoretical parameter selection; and (2) introduction of $A_{threshold}$, a decision-theoretic parameter that quantifies the minimum AI accuracy required to achieve a specified clinical error rate target for a specific user-and-task profile.

Clinical AI decision support is being introduced into nursing practice without quantitative guidance on the minimum AI accuracy required for safe use. The $A_{threshold}$ framework presented here is expected to benefit nursing educators, clinical informaticists, hospital administrators evaluating AI deployment, and regulators developing accuracy standards for clinical AI decision support systems. This work is submitted to the JMIR Nursing theme issue on Artificial Intelligence (AI) in Nursing, addressing the in-scope topics of LLM use in settings where nurses provide care, comparisons of AI algorithm effectiveness in supporting decision-making among nurses, and strengths and limitations of AI technologies in nursing practice.

Methods

Study Design

This study used a theoretical Monte Carlo simulation calibrated against published human-AI reliance data. The published human data were used to calibrate the model's accuracy-related parameters, while all other parameters not determined by calibration were specified a priori from theoretical considerations and tested through sensitivity analysis. To ensure that the simulation's operating range was ecologically plausible, the accuracy of contemporary general-purpose LLMs on complex clinical tasks was characterized using published benchmarks [18,19] rather than a separate in-house evaluation.

Ethical Considerations

This study did not involve human participants or the use of patient data; it comprised a computational simulation calibrated against previously published, publicly available aggregate data. Accordingly, institutional review board approval was not required.

Theoretical Framework and Model Formulation

Reliance Model

Human reliance on AI was modeled as a linear function of AI accuracy, clinician experience, and task complexity:

$$\text{Reliance}=\beta+\beta_A\cdot A+\beta_E\cdot E+\beta_C\cdot C+\beta_{AE}\cdot(A\times E)+\beta_{AC}\cdot(A\times C)+\varepsilon_1 \quad (1)$$

where A denotes AI accuracy on [0, 1]; E signifies clinical experience normalized by dividing the years of experience by 10 (so that 0, 2, and 10 years map to 0.0, 0.2, and 1.0, respectively); C represents task complexity coded as 0.0 (low), 0.5 (moderate), and 1.0 (high); and $\varepsilon_1 \sim N(0, \sigma=0.05)$. To preserve its probabilistic interpretation, reliance was clipped to the interval [0, 1].

The linear form was chosen over more elaborate specifications (eg, a logistic trust mediator with a quadratic accuracy penalty) for 3 reasons. First, it is parsimonious, as it contains the minimum parameters needed to capture additive effects of accuracy, experience, and complexity with their 2-way interactions. Second, it can be directly calibrated against published reliance data using standard regression methods, whereas more elaborate specifications add extra parameters that cannot be reliably estimated from the available data. Third, empirical studies of human-AI reliance consistently reported approximately linear relationships between AI accuracy and reliance across the 0.5 - 1.0 range [14,15], which empirically justifies the linear specification.

XSL•FO
RenderX

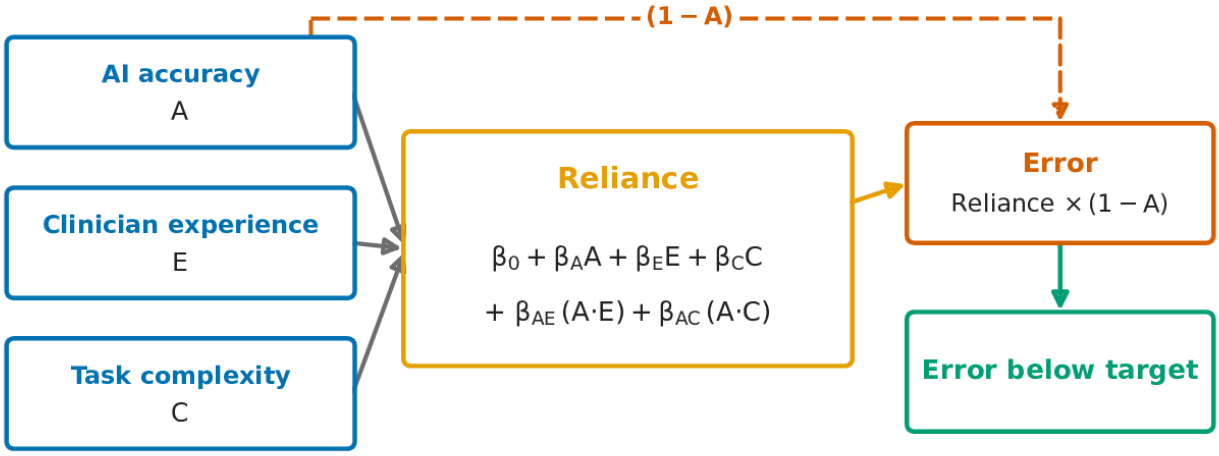
Error Model

Error was modeled as the product of behavioral reliance and the AI failure rate with additive Gaussian noise:

$$\text{Error} = \text{reliance} \times (1 - A) + \varepsilon_2 \quad (2)$$

where $\varepsilon_2 \sim N(0, \sigma = 0.02)$, with error subsequently clipped to [0, 1]. This specification assumes that an erroneous AI recommendation leads to a clinical error only to the extent that the clinician relies on it. Figure 1 presents the conceptual model of human-AI interaction in nursing decision-making.

Figure 1. Conceptual model of AI-assisted clinical decision-making. Inputs (AI accuracy A, clinician experience E, and task complexity C) feed the reliance equation with additive main effects and 2-way interactions ($A \times E$ and $A \times C$). The accuracy-related coefficients β and β_A were empirically calibrated against 9 data points drawn from 3 independent randomized experiments on AI-assisted decision-making (total N=3502). The remaining coefficients were fixed from prior theoretical considerations. Error is computed as reliance×(1−*A*), so AI accuracy enters the error term both indirectly through reliance and directly through the (1−*A*) factor. The predicted error is then compared against a specified target error rate. β_0 and β_A were calibrated from 9 empirical reliance points [14,15]; β_E, β_C, β_{AE}, and β_{AC} were fixed from the literature.



Model Calibration

The *A*-related coefficients of the reliance model (β and β_A) were calibrated using weighted least squares against 9 empirical data points derived from 3 independent randomized experiments on AI-assisted decision-making. Studies met 3 prespecified criteria: experimental manipulation of AI accuracy at ≥2 levels in the 0.5 - 1.0 range, reported behavioral measure of reliance (agreement fraction, switch fraction, or equivalent), and total study n≥100.

In total, 3 studies met all 3 criteria. In all 3 experiments, laypeople recruited via Amazon Mechanical Turk used a machine-learning classifier to predict the outcomes of speed-dating events (whether a given participant would want to meet their date again), and reliance was operationalized as the agreement (or switch) fraction between the participant's final decision and the model's prediction. Lu and Yin [15] (experiment 2; n=466 crowdworkers; 13,980 trials) manipulated the designed AI accuracy (50% and 80%) and measured the final agreement rate. Individual-level data were obtained from authors' publicly available dataset [20] and reanalyzed to obtain worker-level reliance means at each accuracy level (*A*=0.50: n=252; mean 0.642, SD 0.139; *A*=0.80: n=214; mean 0.664, SD 0.134). Yin et al [14] (experiment 1, phase 1; n=1994; stated accuracy ∈ {60%, 70%, 90%, 95%}; agreement fraction with observed accuracy of 80% and experiment 3, phase 2; n=1042; observed accuracy ∈ {55%, 80%, 100%}; agreement fraction averaged across 2 stated accuracy levels) contributed 4 and 3 data points, respectively. Values were extracted from the published figures (Figures 3a and 5a) of Yin et al [14] by visual inspection. The 2 Yin et al [14] experiments used here (experiment 1, phase 1 and experiment 3, phase 2) together comprised 3036 of the 3793 participants reported across that study's 3 experiments; adding the 466 crowdworkers of Lu and Yin [15] yields the total calibration sample of 3502. Experiments and phases that did not manipulate AI accuracy across at least 2 levels in the 0.5 - 1.0 range were not used for calibration.

All calibration targets correspond to the clinician-experience value *E*=0 (laypeople crowdworkers) and the task complexity value *C*=0.5 (moderate complexity; speed-dating outcome prediction requires the integration of multifeature personal profiles). Calibration was performed by minimizing the sum of the squared residuals weighted by the square root of each study's sample size using the Nelder-Mead simplex algorithm in SciPy (version 1.13; NumFOCUS). The other coefficients (β_E, β_C, β_{AE}, and β_{AC}) were set from prior theoretical considerations and are described in the Parameter Specification section.

Calibration yielded β=0.471 (baseline reliance at *A*=0, *E*=0, *C*=0.5 with literature-fixed complexity contribution) and β_A=0.201 (reliance increase per unit of AI accuracy). The calibrated model achieved a root-mean-square error (RMSE) of 0.054 across the 9 empirical points.

To evaluate the precision of the calibrated coefficients given the limited number of independent data sources, study-level bootstrap resampling was run with 2000 iterations. For each iteration, the 3 studies were sampled with replacement, and β and β_A were recalibrated on the resulting target set. Leave-one-study-out cross-validation was also conducted by

XSL•FO
RenderX

excluding each study in turn and recalibrating the remaining 2. The bootstrap distribution and leave-one-out results are presented in the Results section.

Parameter Specification

The remaining 4 coefficients were specified a priori based on theoretical considerations and were not estimated from the calibration data. The coefficient for clinical experience in the reliance equation (β_E=−0.20) reflects two complementary mechanisms: (1) expertise development cultivates reliance on internalized clinical knowledge and pattern recognition rather than on external decision aids [21,22] and (2) empirical studies in clinical AI decision support have shown that longer work experience is directly associated with reduced susceptibility to incorrect AI recommendations (OR 1.89 for work experience in a recent study of 223 health care professionals [16]). The coefficient for task complexity (β_C=+0.20) operationalizes evidence that increased cognitive load is associated with reliance on decision aids [9]. The 2 interaction coefficients (β_{AE}=+0.10 and β_{AC}=+0.10) represent, respectively, expertise-based attenuation and complexity-based amplification of accuracy effects. These values are theoretically motivated rather than empirically estimated. The specific magnitudes (|β|=0.10 - 0.20) were chosen as conservative nominal values, deliberately set at or below the magnitude of the calibrated accuracy slope ($|\beta_A|$≈0.20) so that these unvalidated assumptions would not dominate the empirically calibrated reliance-accuracy relationship; they are not presented as precise estimates. Because 9 aggregate calibration points cannot identify these 4 coefficients empirically, alternative specifications were not selected on the basis of statistical fit; instead, the influence of these values was bounded through the sensitivity analyses described below, which span a plausible range of alternative magnitudes.

Simulation Conditions

In a Monte Carlo simulation, an outcome of interest is estimated by repeatedly drawing random values from specified probability distributions and averaging the results across many simulated trials; here, this approach propagated the small random variation in the modeled reliance and error terms into stable estimates of the mean predicted error rate for each condition. A full factorial design was used: 3 levels of AI accuracy (0.5, 0.7, and 0.9)×3 levels of clinical experience (0, 2, and 10 years)×3 levels of task complexity (low, moderate, and high)=27 conditions. Each condition was simulated with 10,000 Monte Carlo trials (total n=270,000 simulated decision trials). The pseudorandom number generator was initialized with seed 42 using the numpy default_rng interface for full reproducibility.

$A_{threshold}$ Analysis

The minimum AI accuracy required to achieve a specified clinical error rate target (denoted as $A_{threshold}$) was computed analytically for each combination of clinical experience and task complexity. Given the linear reliance specification, the predicted mean error is a quadratic function of A:

$$\text{Error}(A|E, C)=[(\beta+\beta_E\cdot E+\beta_C\cdot C)+(\beta_A+\beta_{AE}\cdot E+\beta_{AC}\cdot C)\cdot A]\times(1-A) \quad (3)$$

For a specific target error threshold T, $A_{threshold}$ was computed as the smallest $A \in [0, 1]$ at which error($A|E, C$)≤T, obtained through a dense grid search over 1000 values of A. In total, 4 target error thresholds were evaluated, namely, 5%, 10%, 15%, and 20%. These values were not adopted from a single validated benchmark; rather, they constitute a graded series of progressively stricter error rate targets intended to span the range of error tolerances plausibly relevant to clinical decision support, from comparatively lenient targets appropriate to low-stakes supportive use to stringent targets appropriate to higher-stakes settings. Consistent with the exploratory aim of this study, the resulting $A_{threshold}$ values should be interpreted as normative, hypothesis-generating estimates rather than as empirically validated error rate targets, and the specific threshold relevant to any given deployment will depend on the clinical context.

Sensitivity Analysis

To assess model robustness, each literature-fixed coefficient (β_E, β_C, β_{AE}, and β_{AC}) was systematically perturbed by ±20% from its baseline value. The calibrated coefficients (β and β_A) were not perturbed because they were empirically determined rather than theoretically assumed. For each variant, the entire 27-cell simulation was rerun, and the resulting cell-level error rates were compared with baseline using Spearman rank correlation and maximum absolute error difference.

Joint Human-AI Decision Model

The primary error specification (error=reliance×(1−A)) attributes a clinical error solely to reliance on an incorrect AI recommendation and therefore represents the limiting case, in which the clinician contributes no independent accuracy. Human-AI teams do not always outperform either party alone and can even perform worse on decision-making tasks [23,24]—for example, a randomized trial found that giving physicians access to an LLM did not significantly improve their diagnostic reasoning over conventional resources [25]. In addition, users frequently overrely on AI by following incorrect recommendations [13], and appropriate reliance requires accepting correct advice while rejecting incorrect advice [26]. For these reasons, a joint human-AI decision model was additionally analyzed that represents both outcomes of an AI-assisted decision explicitly: when the clinician follows the AI, an error occurs if the AI is incorrect; and when the clinician overrides the AI, an error occurs if the clinician's own independent judgment is incorrect. Writing R for reliance (the probability of following the AI), A for AI accuracy, and p_c for the clinician's independent accuracy, the expected team error is $\text{error}_{joint}=R\times(1-A)+(1-R)\times(1-p_c)$. The primary model is recovered exactly as the special case $p_c\to1$ (a clinician who is always correct when acting independently); this was verified to reproduce the reported $A_{threshold}$ values (0.894 and 0.779 for the novice×high-complexity cell at the <10% and <20% targets). An asymmetric variant additionally distinguished passive deference to an incorrect recommendation from active correction of one by scaling the corrective term by a factor ρ ∈ [0, 1], where ρ=1 recovers the symmetric model, and smaller ρ represents clinicians who more often catch and overturn a wrong

XSL•FO
RenderX

recommendation. $A_{threshold}$ was recomputed under this model across clinician-accuracy levels $p_c \in \{0.6, 0.7, 0.8, 0.9, 1.0\}$.

Extended Sensitivity Analysis

To probe robustness beyond the ±20% perturbations, 3 additional analyses were performed. First, each literature-fixed coefficient was varied over a wide one-at-a-time range (−100%, −50%, +50%, and +100% of its baseline magnitude, together with a full sign reversal), and the 27-cell simulation was rerun for each variant. Second, a global Monte Carlo analysis simultaneously sampled all 4 literature-fixed coefficients from prior distributions across 50,000 draws under 2 priors: a theory-signed prior that preserved the assumed direction of each coefficient while allowing its magnitude to vary widely, and a fully sign-agnostic prior that additionally permitted every assumed sign to reverse. The resulting distribution of $A_{threshold}$ values for the highest-risk (novice×high-complexity) and lowest-deference (experienced×low-complexity) cells was summarized by its median, 5th- to 95th-percentile interval, and the probability of exceeding an accuracy of 0.70. Third, to test whether the conclusions were an artifact of the linear reliance form, 2 bounded alternative structures—a logistic (sigmoid) reliance function and a concave saturating function—were fit to the same 9 calibration points, and $A_{threshold}$ was recomputed from each. These analyses are reported in the Results section, summarized in Multimedia Appendix 1, and provided in full in the repository.

Results

Model Calibration

Weighted least-squares calibration of the reliance model against the 9 empirical data points yielded $\beta = 0.471$ and $\beta_A = 0.201$. The calibrated model achieved an RMSE of 0.054 across the calibration targets (unweighted $R^2 = 0.34$, reflecting between-study heterogeneity in baseline reliance that the model does not attempt to capture). Multimedia Appendix 2 overlays the calibrated model predictions on the 9 empirical data points, stratified by the source study.

Study-level bootstrap resampling (2000 iterations) yielded a 95% CI for β_A of 0.023 - 0.234, which reflects the small number of independent calibration sources and the between-study heterogeneity in baseline reliance. Despite this width, the sign of β_A was stable across all 2000 bootstrap iterations ($P(\beta_A > 0) > .99$), confirming that the monotonic positive relationship between AI accuracy and reliance is empirically robust. The leave-one-study-out cross-validation produced recalibrated β_A values of 0.094 (excluding Lu and Yin [15]), 0.246 (excluding experiment 3 by Yin et al [14]), and 0.199 (excluding experiment 1 phase 1 by Yin et al [14]). This finding indicates that the calibration depends most heavily on the study by Lu and Yin [15] but does not collapse upon its removal. Per-study calibration yielded β_A estimates of 0.023, 0.037, and 0.158, all positive: the joint calibration value of 0.201 represents the precision-weighted aggregate across these heterogeneous study-level estimates.

Two features of the calibration result warrant emphasis. First, the calibrated slope ($\beta_A = 0.201$) indicates that each 0.10-unit increase in AI accuracy corresponds to approximately a 0.02-unit increase in the predicted mean reliance. This is considerably weaker than the reliance-accuracy relationships assumed in some theoretically motivated frameworks and reflects the fact that human reliance on AI systems is sensitive to, but not dominated by, the accuracy of the system. Second, the 3 studies contributing calibration data have different baseline reliance levels (Lu and Yin [15]: mean≈0.65; experiment 1 by Yin et al [14]: mean≈0.78; and experiment 3 by Yin et al [14]: mean≈0.82), a heterogeneity that the linear model averages across rather than resolves. Residuals were accordingly approximately symmetric around 0, but the RMSE was markedly larger than the SEs of the study means, an outcome we interpret as reflecting unmodeled study-level factors rather than misspecification of the functional form in the accuracy dimension.

Predicted Reliance and Error Rates

Table 1 presents the predicted error rates across all 27 simulation conditions. The underlying predicted reliance values ranged from 0.421 (experienced clinicians on low-complexity tasks with $A = 0.5$) to 0.938 (novice clinicians on high-complexity tasks with $A = 0.9$), reflecting the combined influence of accuracy, experience, and complexity on the tendency to accept AI recommendations. Notably, the calibrated slope on A is sufficient to produce an approximately 0.08-unit increase in mean reliance between $A = 0.5$ and $A = 0.9$, but insufficient to generate an interior maximum of predicted error rate as a function of A.

XSL•FO
RenderX

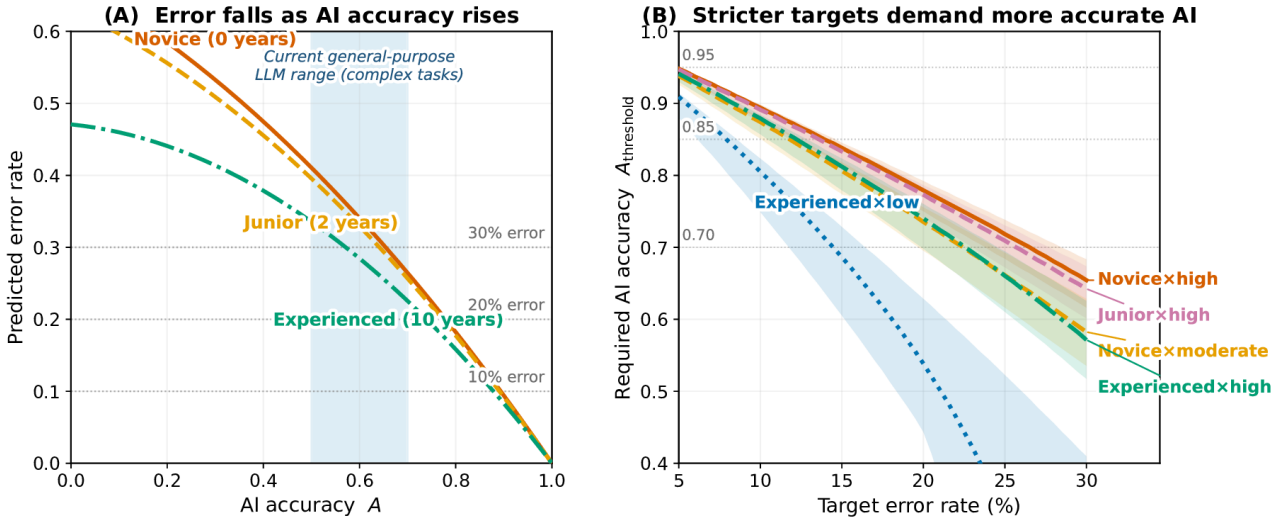
Table . Predicted error rates by simulation condition[a].

AI accuracy	Experience (years)	Low complexity, mean (SD)	Moderate complexity, mean (SD)	High complexity, mean (SD)
0.5	0	0.29 (0.03)	0.35 (0.03)	0.41 (0.03)
0.5	2	0.27 (0.03)	0.33 (0.03)	0.40 (0.03)
0.5	10	0.21 (0.03)	0.27 (0.03)	0.34 (0.03)
0.7	0	0.18 (0.03)	0.22 (0.03)	0.26 (0.03)
0.7	2	0.18 (0.03)	0.22 (0.03)	0.26 (0.03)
0.7	10	0.14 (0.03)	0.19 (0.03)	0.23 (0.03)
0.9	0	0.07 (0.02)	0.08 (0.02)	0.09 (0.02)
0.9	2	0.06 (0.02)	0.08 (0.02)	0.09 (0.02)
0.9	10	0.05 (0.02)	0.07 (0.02)	0.08 (0.02)

[a]Values are mean (SD) of simulated error rates across 10,000 Monte Carlo trials per cell (seed=42).

The predicted error rates exhibited a clean 3-tier structure when grouped by AI accuracy level. At A=0.9, all 9 cells fell below a 10% error rate (range 0.054 - 0.094). At A=0.7, the predicted error rates ranged from 0.145 to 0.265, spanning the 10% - 20% and 20% - 30% ranges, with all 3 novice cells exceeding 18% predicted error. At A=0.5, the predicted error rates ranged from 0.211 to 0.411, with all 6 moderate- and high-complexity cells exceeding 27% predicted error. For the cell most representative of the clinically high-risk scenario (novice clinicians performing high-complexity tasks), the predicted error rates were 0.411 at A=0.5, 0.265 at A=0.7, and 0.094 at A=0.9 (Figure 2A).

Figure 2. Predicted error rates and accuracy thresholds for AI decision support in nursing. (A) Predicted mean error rate as a function of AI accuracy A at high task complexity (C=1.0) for 3 experience levels. The horizontal dotted lines mark 10%, 20%, and 30% target error rates. The shaded vertical band marks a reference range for the accuracy of contemporary general-purpose LLMs on complex clinical tasks (approximately 0.5 - 0.7), based on published benchmarks [18,19]. (B) $A_{threshold}$ curves plotting the required AI accuracy against the target error-rate threshold for 5 representative experience-by-complexity combinations. Solid and dashed lines show point estimates from the calibrated model. The shaded bands represent study-level bootstrap 95% CIs (2000 iterations). Horizontal reference lines mark a reference LLM accuracy level (0.70), plausible near-term frontier LLM accuracy (0.85), and demanding target accuracy (0.95). Note that $A_{threshold}$ uncertainty is narrow at low target error rates (left side) and wider at higher target error rates (right side), reflecting the mathematical structure of error=reliance×(1−A). LLM: large language model.



The A=0.7 simulation condition is closest to the accuracy levels currently reported for general-purpose LLMs on complex clinical tasks [18,19], in which the predicted error rates for novice clinicians range from 0.184 (low-complexity tasks) to 0.265 (high-complexity tasks). At high task complexity, the condition most relevant to ambiguous clinical presentations, the predicted error rate of 0.265 implies that approximately 1 in 4 AI-supported clinical decisions would be incorrect when the system is used by a novice clinician at this accuracy level.

$A_{threshold}$ Analysis

Table 2 presents $A_{threshold}$ values, the minimum AI accuracy required to achieve a specified error rate, for each of the 9 experience-by-complexity combinations at 4 target error thresholds (5%, 10%, 15%, and 20%). For novice clinicians performing high-complexity tasks, the minimum AI accuracy values required to achieve error rates <10%, <15%, and <20% were 0.894, 0.838, and 0.779, respectively. Corresponding values for experienced clinicians performing low-complexity tasks were 0.806, 0.686, and 0.539, respectively. These

XSL•FO
RenderX

differences highlight the interaction between clinician experience, task complexity, and the minimum accuracy requirement for clinical use. Because the reliance model is calibrated on nonnursing behavioral data, these numerical thresholds should be read as illustrative outputs of the model rather than as nursing-derived empirical standards; their value lies in the relative ordering across user-and-task profiles rather than in the specific accuracy figures.

Table . $A_{threshold}$ values—minimum AI accuracy required to achieve specified error rate targets[a].

Experience and complexity	Error<5%	Error<10%	Error<15%	Error<20%
0 years				
Low	0.925	0.845	0.759	0.670
Moderate	0.939	0.874	0.807	0.736
High	0.949	0.894	0.838	0.779
2 years				
Low	0.922	0.839	0.749	0.653
Moderate	0.937	0.870	0.800	0.726
High	0.947	0.891	0.834	0.773
10 years				
Low	0.909	0.806	0.686	0.539
Moderate	0.929	0.851	0.766	0.670
High	0.942	0.879	0.812	0.740

[a]$A_{threshold}$ is the smallest AI accuracy value $A \in [0, 1]$ at which the calibrated model's predicted mean error rate falls below the specified error rate target, computed analytically from the closed-form quadratic expression derived from the linear reliance specification. Values are rounded to 3 decimal places and were computed from the full-precision calibrated coefficients; exact coefficient values are available in the repository.

Bootstrap propagation of calibration uncertainty into $A_{threshold}$ values revealed an important asymmetry: although β_A has a wide bootstrap 95% CI, the $A_{threshold}$ values are tightly bounded. For the novice×high-complexity combination, the bootstrap 95% CIs for the $A_{threshold}$ to achieve error rates of 10%, 15%, and 20% were 0.877 - 0.898, 0.814 - 0.845, and 0.750 - 0.792, respectively. This tight propagation reflects the mathematical structure of error=reliance×(1−*A*): at low target error rates, the (1−*A*) factor must be small, which constrains *A* to lie close to 1 regardless of the exact value of β_A within its CI. Correspondingly, across the accuracy range currently reported for general-purpose LLMs on complex clinical tasks [18,19], the predicted error rate for novice clinicians on high-complexity tasks ranges from approximately 0.26 at the upper end of that range to approximately 0.41 at the lower end. The conclusion that AI systems operating in this accuracy range yield error rates well above stringent targets for this user-and-task combination is robust across the bootstrap distribution of the calibrated coefficients.

Figure 2B presents the $A_{threshold}$ function graphically, plotting the required AI accuracy against the target error-rate threshold for 5 representative experience-by-complexity combinations. Horizontal reference lines mark a reference LLM accuracy level (0.70), plausible near-term frontier LLM accuracy (0.85), and demanding target accuracy (0.95). At a reference accuracy level of 0.70, a total of 4 of the 9 cells achieve predicted error rates <20%: the 3 low-complexity cells (ranging from 14.5% for experienced clinicians to 18.4% for novices) and the experienced×moderate-complexity cell (18.5%). The remaining 5 cells—comprising novice and 2-year-experience clinicians on moderate- or high-complexity tasks, plus experienced clinicians on high-complexity tasks—have predicted error rates ranging from 21.6% to 26.5%, consistently exceeding a 20% error rate.

Sensitivity Analysis

When each of the 4 literature-fixed coefficients (β_E, β_C, β_{AE}, and β_{AC}) was varied by ±20% from its baseline value, the overall error ranking of the 27 cells was nearly preserved in every perturbation. Across all 8 perturbations (4 coefficients×2 directions), the Spearman rank correlation between the perturbed and baseline cell-level error rankings ranged from 0.991 to 1.000, with a mean of 0.998 (SD 0.003). The maximum absolute change in any cell's predicted error rate across all perturbations was 0.020, occurring for the β_E+20% perturbation in the experienced×moderate-complexity×*A*=0.5 cell (baseline 0.273; perturbed 0.253). These results indicate that the principal conclusions of the simulation—the tiered error-rate structure of the 27 cells and the $A_{threshold}$ values reported in Table 2—are robust to moderate perturbation of the literature-fixed coefficients.

Joint Decision Model and Extended Robustness

Under the joint human-AI decision model, the principal conclusion for the highest-risk condition was unchanged. For novice clinicians performing high-complexity tasks, the minimum AI accuracy required to keep predicted team error below 10% moved only from 0.894 at $p_c \to 1$ to 0.917 at p_c=0.6, and the threshold for the<20% target moved from 0.779 to 0.818 across the same range; because reliance is high in this cell, the override path contributes little, and clinician accuracy has

XSL•FO
RenderX

limited leverage. A qualitatively different pattern emerged in the lowest-deference cell (experienced clinicians on low-complexity tasks), where team error acquired an irreducible floor equal to $(1-R(A=1))\times(1-p_c)$. This floor grew from 0% at $p_c\to1$ to 8.6% at p_c=0.8 and 17.1% at p_c=0.6, so that the<10% target became unreachable by any AI accuracy once p_c fell to 0.7 or below. This structural feature—that improving AI accuracy alone cannot drive team error below the clinician's own error contribution—indicates that a single accuracy number is insufficient and that the human contribution to the joint decision must be represented. Under the asymmetric variant, allowing clinicians to correct a wrong recommendation (ρ<1) lowered the required accuracy substantially (for the novice×high-complexity cell at the <20% target, from 0.779 at ρ=1 to 0.673 at ρ=0.7 and 0.516 at ρ=0.5), indicating that the primary model is conservative with respect to correction.

The extended sensitivity analyses reinforced the ±20% results. Under wide one-at-a-time perturbation (up to ±100% and sign reversal of each literature-fixed coefficient), the novice×high-complexity threshold was mathematically invariant to β_E and β_A (both multiply the experience term, which is 0 in that cell) and varied only with the complexity-related coefficients; even under a full sign reversal of β_C, the required accuracy for the <10% target remained between 0.805 and 0.900, and the rank ordering of the 27 cells was preserved (Spearman ρ≥0.77 across all perturbations, and ρ>0.95 for most). In the global Monte Carlo analysis (50,000 draws per prior), the theory-signed prior yielded a novice×high-complexity $A_{threshold}$ with median 0.894 (5th-95th percentile 0.863 - 0.900) for the <10% target and 0.780 (5th-95th percentile 0.709 - 0.800) for the <20% target, with the required accuracy exceeding 0.70 in 100% and 97.4% of draws, respectively. Even under the fully sign-agnostic prior, in which every assumed coefficient sign was permitted to reverse, the <10% threshold for this cell retained a median of 0.844 and exceeded 0.70 in 84.4% of the sampled parameter space; values in the low tail arose only under sign combinations contrary to the theoretical expectations. Finally, both bounded alternative reliance structures fit the 9 calibration points essentially as well as the linear form (RMSE 0.0542 for the logistic and concave-saturating forms vs 0.054 for the linear form) and produced near-identical thresholds for the novice×high-complexity cell (0.900 and 0.794 - 0.795 for the <10% and <20% targets, respectively), indicating that the conclusions are not an artifact of the linear specification. Multimedia Appendix 1 summarizes the joint-model and global-sensitivity results.

Discussion

Principal Findings

In this study, an empirically calibrated simulation model was developed to quantify the AI accuracy required to achieve specified error rate targets in nursing decision-making across combinations of clinician experience and task complexity. Three principal findings warrant emphasis. We emphasize at the outset that the contribution of this work is a reusable, model-agnostic theoretical framework rather than an evaluation of any particular AI system. Its value lies in the mapping it defines from AI accuracy, clinician experience, and task complexity to a predicted error rate and a corresponding minimum-accuracy requirement—a mapping that does not depend on which model is current or on the specific accuracy figure entered into it.

First, the moderate accuracy levels currently reported for general-purpose LLMs on complex clinical tasks [18,19] fall within an operating zone where the predicted error rates exceed stringent targets for novice clinicians performing complex tasks. At the upper end of this reported range (approximately 0.70), the simulation predicts an error rate of approximately 0.26 in the novice×high-complexity cell—rising to roughly 0.41 at the lower end of the range—substantially exceeding the stringent error targets relevant to autonomous clinical decision support in high-risk clinical scenarios.

Second, the calibrated reliance model—grounded in 9 empirical data points from 3 independent randomized experiments (total N=3502)—indicates that human reliance on AI recommendations rises monotonically with AI accuracy but at a modest rate (β_A=0.20). This slope is considerably shallower than the reliance-accuracy relationships assumed in some theoretical frameworks and reflects the cumulative evidence that users calibrate their reliance in response to AI accuracy, but do so imperfectly and with substantial residual reliance even on low-accuracy systems. Because reliance does not rise steeply enough to offset the declining AI failure probability at high accuracy levels, the predicted clinical error rates in this model decrease monotonically as AI accuracy increases, without an interior maximum. The clinical implication is that no "dangerous middle" region exists in which error rates peak; the clinical concern, rather, is the absolute magnitude of error rates across a broad low-to-moderate accuracy range.

Third, $A_{threshold}$ values, the minimum AI accuracy required to achieve a specified error rate target, exhibit a pronounced dependence on clinician experience and task complexity. For the highest-risk combination (novice clinicians performing high-complexity tasks), achieving an error rate <10% requires AI accuracy of at least 0.89, and achieving an error rate <20% requires accuracy of at least 0.78. For experienced clinicians performing low-complexity tasks, the corresponding thresholds are 0.81 and 0.54. These differences, spanning up to 24 percentage points of the required accuracy across user-and-task profiles (at the <20% error target), indicate that a single accuracy benchmark cannot meaningfully govern clinical use decisions; the threshold depends on the user-and-task context.

Comparison With Prior Work

The predictions of the proposed model converge with findings from 2 independent lines of empirical research on human-AI interaction.

First, in the AI-assisted decision-making literature, the calibrated model is grounded directly in the reliance data from Lu and Yin [15] and Yin et al [14], representing 3 randomized experiments with a combined sample of 3502 participants. The linear form of the model aligns with the empirically observed monotonic relationship between AI accuracy and human reliance, and the absence of an interior reliance maximum in the range $A \in [0.5, 1.0]$, observed in both studies, directly shaped the structural

XSL•FO
RenderX

form of the reliance equation. This empirical basis distinguishes the present framework from earlier theoretical treatments that posited nonlinear reliance patterns without supporting behavioral data.

Second, in the clinical decision support literature, Kücking et al [16] reported that AI recommendation correctness exerted a strong bidirectional influence on the diagnostic performance of 223 physicians and nurses (OR 10.0; P<.001 for correct AI recommendations; reciprocal declines for incorrect recommendations). They also found that longer work experience was associated with higher diagnostic accuracy (OR 1.89), and formal qualifications further improved accuracy (OR 1.40). These clinical expertise effects are directionally consistent with the negative experience coefficient (β_E=−0.20) specified a priori in the present model. In a related study by the same group [27], the overall human diagnostic accuracy averaged 79.3% (up to 85% in the formally qualified subgroup); by comparison, general-purpose LLMs have been reported to achieve lower accuracy on complex clinical tasks [18,19], situating the operating range explored in the present simulation within a clinically meaningful band relative to human performance.

The introduced $A_{threshold}$ framework is best understood in relation to 3 established frameworks for human-automation interaction. Lee and See [10] proposed a process-oriented framework in which trust calibration determines appropriate reliance but did not provide quantitative guidance for translating system accuracy into use case decisions. Parasuraman and Riley [9] articulated a use/misuse/disuse/abuse taxonomy that classifies human-automation interaction failure modes but treats accuracy as an input among many rather than as a use-determining quantity. More recently, the literature on trustworthy AI [6] has emphasized explainability, fairness, and accountability as design principles, but has generally not specified accuracy thresholds for clinical use.

The present framework differs from prior approaches in 3 ways. First, it converts a single quantitative input—AI accuracy on the intended task—into a deployment decision using an empirically calibrated function. Second, it explicitly adjusts the deployment threshold based on user experience and task complexity, recognizing that no universal accuracy cutoff is appropriate across different use cases. Third, the graphical guidance provided in Figure 2B allows clinicians and administrators to read off the minimum required accuracy for a specific error tolerance and user-task profile without requiring statistical expertise. These properties make the $A_{threshold}$ framework complementary to, rather than competitive with, existing process- and design-oriented frameworks: $A_{threshold}$ provides a quantitative gating criterion that can be applied alongside trust calibration and trustworthy AI principles in deployment review processes.

Taken together, the present framework is consistent with both the quantitative reliance patterns observed in general-purpose AI-assisted decision-making and the qualitative expertise effects observed in clinical decision support. The principal gap in the empirical foundation—behavioral reliance data collected specifically in nursing populations performing clinical decision-making tasks across multiple AI accuracy levels—remains a priority for future research.

Theoretical Integration

The finding that $A_{threshold}$ varies strongly with task complexity can be interpreted through the lens of dual-process theory [28]. At low task complexity, system 2 (analytical) processing may remain engaged even when AI accuracy is moderate, because the clinician retains sufficient cognitive resources to evaluate the AI recommendation independently. At high task complexity, by contrast, system 1 (automatic) processing may predominate, as cognitive resources are preferentially allocated to the task itself rather than to critical evaluation of the AI recommendation. This asymmetry provides a theoretical rationale for the model's larger complexity-by-accuracy interaction coefficient and for the model-implied result that novice clinicians, who have not yet developed efficient system 2 strategies for clinical reasoning [21,22], face the most stringent AI accuracy requirements.

Within the framework of Lee and See [10], the model implies that appropriate reliance is unlikely to be fully achieved through user education or interface design alone when AI accuracy falls substantially below the $A_{threshold}$ for the relevant user-and-task combination. In such contexts, the residual error rate is a function of system accuracy rather than of user awareness, and the appropriate intervention is either to restrict AI use to user-and-task combinations for which the $A_{threshold}$ is met or to improve AI accuracy itself.

A further implication concerns the durability of the framework, as language models continue to improve. Because the $A_{threshold}$ framework maps any given AI accuracy onto a predicted error rate and a corresponding minimum-accuracy requirement, its utility does not depend on the accuracy level of any particular contemporary model. As model performance advances, the same framework can be reapplied to locate an improved system within the accuracy-error landscape and to re-evaluate which user-and-task combinations it can safely support. The specific operating range attributed here to current general-purpose LLMs is thus a movable reference point within a framework that is itself invariant to such change.

Clinical Implications

The following implications are exploratory and contingent on behavioral validation in nursing settings; they describe how the $A_{threshold}$ framework might inform practice rather than offering validated recommendations and should be interpreted in conjunction with the limitations. In particular, a system's meeting a given $A_{threshold}$ on a benchmark should be treated as a necessary but not a sufficient condition for nursing use, and the framework must not be interpreted as authorization to deploy a system without behavioral validation in the relevant nursing context.

First, the minimum AI accuracy required for clinical use should be specified as a function of the intended user and task, not as a single benchmark. Although an AI system operating at the moderate accuracy levels typical of current general-purpose LLMs on complex clinical tasks [18,19] may be suitable for use by experienced clinicians in low-complexity tasks, it is

XSL•FO
RenderX

predicted to produce error rates >25% when used by novice clinicians in high-complexity decision-making. Therefore, clinical use decisions should be based on the specific $A_{threshold}$ value that corresponds to the intended user population and task profile, rather than on the raw AI accuracy value alone. Moreover, the tolerable error rate is itself context-dependent—shaped by the clinical setting, the severity of potential harm, the degree of human supervision, and whether the system is used in a supportive or an autonomous capacity—so a target that is reasonable for an educational support tool may be inadequate for high-risk acute care.

Second, as AI is increasingly integrated across nursing practice and curricular reform has been called for to prepare nurses to work safely alongside these technologies [29], nursing education programs should incorporate explicit instruction on the relationship among AI accuracy, user experience, and clinical error rates. Students should understand that AI systems in the low-to-moderate accuracy range typical of current general-purpose LLMs may be inadequate for autonomous use in high-complexity clinical decisions. In these settings, AI should augment, not replace, clinical judgment. The $A_{threshold}$ framework offers a practical decision tool for this educational purpose.

Third, clinical workflows could adopt differentiated deployment protocols informed by the $A_{threshold}$ framework. When the current AI accuracy for user-and-task combinations falls below the relevant $A_{threshold}$, enhanced monitoring safeguards, such as mandatory verification steps, second-clinician review, or confidence-based routing, should be applied preferentially. These measures provide a practical intermediate solution while AI accuracy continues to increase toward levels suitable for autonomous use.

Fourth, regulatory and institutional review processes for clinical AI decision support could be informed by encouraging deployment proposals to specify the intended user population and task profile and to consider the corresponding $A_{threshold}$. Current regulatory frameworks do not typically request this level of user-and-task-specific justification; the $A_{threshold}$ framework is offered as a candidate quantitative input to such review, pending behavioral validation, rather than as a validated regulatory standard. This need is underscored by recent calls for explicit regulatory oversight of LLMs used in health care [30,31].

Limitations

Several limitations merit consideration.

First, although the accuracy-related coefficients of the reliance model were empirically calibrated against published behavioral data, the framework itself remains a theoretical construct, and its predictions require direct behavioral validation. Calibration confirms that the monotonic reliance slope is consistent with empirical observations in AI-assisted decision-making, but it does not establish that the noncalibrated coefficients (β_E, β_C, β_{AE}, and β_{AC}) accurately reflect nurses performing clinical tasks. The same caution applies to the joint decision model introduced in response to review: the clinician's independent accuracy (p_c) and the correction factor (ρ) are illustrative parameters rather than calibrated quantities, so the joint-model results should be read as a structured sensitivity analysis over a plausible range of clinician behavior—showing which conclusions are robust to, and which depend on, the human contribution to the decision—rather than as point predictions for any specific clinician accuracy. The framework should therefore be interpreted as a decision-support and hypothesis-generating tool rather than as a validated predictive model for individual clinical use.

Second, the accuracy levels used to contextualize the simulation's operating range were drawn from published benchmarks of general-purpose LLMs on complex clinical tasks rather than from a nursing-specific evaluation. Reported LLM accuracy varies substantially across models, prompting strategies, task types, and clinical domains, and the accuracy of frontier LLMs on complex clinical tasks remains an active area of research. The operating range examined here should therefore be regarded as an illustrative band informed by the current literature rather than as a definitive estimate of any specific system's clinical performance.

Third, calibration data were drawn from 3 studies that differed systematically in baseline reliance levels (Lu and Yin [15]: ~0.65; experiment 1 by Yin et al [14]: ~0.78; and experiment 3 by Yin et al [14]: ~0.82). The linear model averages across this heterogeneity, producing an RMSE (0.054) larger than the SEs of individual study means. Therefore, the calibrated slope (β_A=0.201) should be interpreted as a central tendency across heterogeneous contexts rather than a precise parameter applicable to any specific setting. Part of this heterogeneity may reflect methodological differences in how AI accuracy was manipulated (stated, observed, or designed), which the calibration pooled as proxies for the underlying *A* variable. The heterogeneity is informative and motivates future research parameterizing contextual factors beyond AI accuracy. With only 3 independent calibration studies, the study-level bootstrap has a limited sample space, so the reported CIs should be viewed as lower bounds on the true uncertainty. Consistent with this, the $A_{threshold}$ values are reported to 3 decimal places to reflect the numerical precision of the simulation, not the empirical precision of the underlying parameter, which is bounded by a wide CI whose lower limit approaches 0. The thresholds are therefore best interpreted in relative and ordinal terms—comparing the accuracy demands of different user-and-task profiles—rather than as exact accuracy targets to be applied at the reported number of significant figures.

Fourth, the calibration data were derived from studies in which crowdworkers predicted speed-dating outcomes rather than clinicians making clinical decisions. Although this is the most empirically rich reliance-by-accuracy source currently available, the clinical context gap is substantial. Clinicians may exhibit baseline reliance, accuracy sensitivities, and experience effects that differ from those of crowdworkers, and the task complexity assigned to speed-dating prediction (*C*=0.5) reflects a qualitative judgment rather than direct measurement. Even so, the directional structure of the model is consistent with recent experimental evidence from clinical populations that include

XSL•FO
RenderX

nurses. In a web-based experiment with 223 dermatologists, reliance on AI was stronger for correct than for incorrect advice and decreased with greater medical experience [32]. More directly, in a wound-maceration study, 223 physicians and nurses made decisions that were sensitive to whether the AI recommendation was correct or incorrect [16], and a related report measuring the rate of agreement with incorrect recommendations found that higher diagnostic performance and certified wound-care training were associated with lower agreement [33]. These patterns match the positive accuracy-reliance slope and the experience-based attenuation encoded in the model, supporting cautious extrapolation of its directional structure to clinical settings. Nursing-specific calibration data obtained across multiple AI accuracy levels would substantially strengthen the empirical foundation of the framework.

Fifth, the linear reliance specification was chosen for parsimony and tractability, since 9 aggregate calibration points are insufficient to identify more elaborate functional forms. Alternative specifications (logistic, power-law, and segmented-linear) might provide superior fit with richer data. The linear form should therefore be interpreted as the simplest specification consistent with the available data, not as a claim that the underlying reliance-accuracy relationship is truly linear.

Sixth, the model does not incorporate contextual factors—such as time pressure, team dynamics, organizational accountability, and environmental stressors—that are known to influence clinical decision-making. These factors likely contribute substantially to real-world reliance and to the observed between-study heterogeneity. Future investigations should explicitly parameterize the most influential contextual factors, which will require more calibration data than are currently available.

Seventh, the reliance and error expressions are clipped to [0, 1] to preserve their probabilistic interpretation. Within the reported 27-cell grid, clipping affected <1% of simulated values and did not materially alter the reported results. At extreme accuracy levels outside the reported grid, clipping becomes more frequent, and the linear reliance expression should not be extrapolated beyond its reported domain.

Conclusions

This study provides an empirically calibrated simulation framework for quantifying the AI accuracy required to achieve specified error rate targets in nursing decision-making. Calibration against 9 empirical data points derived from 3 independent randomized experiments (total N=3502) indicates that human reliance on AI rises gradually but modestly with AI accuracy. For the highest-risk combination—novice clinicians performing high-complexity tasks—the predicted clinical error rates remain above a stringent target (eg, <10%) when AI accuracy is below approximately 0.89. The moderate accuracy levels currently reported for general-purpose LLMs on complex clinical tasks [18,19] fall well below this threshold, suggesting that, on present evidence, such systems are not yet adequate for autonomous use by novice nurses on complex clinical decisions and should serve as an adjunct to, rather than a replacement for, clinical judgment. The $A_{threshold}$ framework introduced here provides a decision-theoretic tool for specifying minimum accuracy requirements as a function of clinician experience and task complexity, and it can inform AI deployment, nursing education, and regulatory policy. Direct behavioral validation in nursing populations remains an essential next step.

Acknowledgments

The author thanks Lu and Yin and Yin et al for making their experimental data and figures publicly available, which made empirical calibration possible. During the preparation of this work, the author used Claude (Anthropic) as a writing assistance tool to draft and revise sections of the manuscript in English, assist in the analysis of publicly available empirical reliance data, and generate Python code for model calibration and Monte Carlo simulation. After using this tool, the author reviewed and edited the content as needed and took full responsibility for the content of the publication.

Funding

The author declared no financial support was received for this work.

Data Availability

The data and code that support the findings of this study are openly available in the nursing-AI-error-thresholds repository on GitHub under the MIT License [34]. The complete reproducibility package includes all calibration scripts, the bootstrap procedure (2000 study-level iterations), the 27-cell factorial simulation, the joint decision model, and the extended (wide, global, and alternative-structure) sensitivity analyses, together with vector-format source files for every figure reported in this paper. Rerunning the figure-generation script from a clean Python ≥3.10 environment reproduces the calibrated coefficients (β_0=0.4708; β_A=0.2010), the bootstrap 95% CI for β_A (0.023-0.234), and all reported numerical results to at least 3 decimal places across platforms. The pseudorandom number generator was initialized with seed 42 using the NumPy default_rng interface, and the 27-cell factorial simulation used 10,000 Monte Carlo trials per cell. Analyses were conducted in Python (version 3.11) with NumPy (version 1.26) and SciPy (version 1.13). A preprint of this manuscript is available on SSRN [35].

XSL•FO
RenderX

Authors' Contributions

HT is the sole author and contributed the following roles according to the CRediT taxonomy: conceptualization, methodology, software, formal analysis, investigation, data curation, writing—original draft, writing—review and editing, and visualization. The author had full access to all data and analyses and takes full responsibility for the integrity of the work and the decision to submit the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Extended robustness analyses. (A) Minimum AI accuracy ($A_{threshold}$) required to keep predicted team error below the 10% and 20% targets under the joint human-AI decision model, plotted against the clinician's independent accuracy p_c, for the highest-risk (novice×high-complexity) and lowest-deference (experienced×low-complexity) cells. The primary model corresponds to $p_c \to 1$. In the lowest-deference cell, the team error acquires an irreducible floor that grows as p_c declines, so the 10% target becomes unreachable by any AI accuracy at $p_c \leq 0.7$. (B) Distribution of the novice×high-complexity $A_{threshold}$ from the global Monte Carlo sensitivity analysis (50,000 draws per prior) for the <10% and <20% targets, under a theory-signed prior and a fully sign-agnostic prior. The vertical line marks the reported estimate (0.894) and the dotted line the 0.70 anchor.
[PNG File, 209 KB - nursing_v9i1e99590_app1.png]

Multimedia Appendix 2

Calibration of the linear reliance model against 9 empirical data points from 3 independent studies (N=3502).
[PNG File, 176 KB - nursing_v9i1e99590_app2.png]

References

1. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019 Jan;25(1):44-56. [doi: 10.1038/s41591-018-0300-7] [Medline: 30617339]
2. Blease C, Kaptchuk TJ, Bernstein MH, Mandl KD, Halamka JD, DesRoches CM. Artificial intelligence and the future of primary care: exploratory qualitative study of UK general practitioners' views. J Med Internet Res 2019 Mar 20;21(3):e12802. [doi: 10.2196/12802] [Medline: 30892270]
3. Ronquillo CE, Peltonen LM, Pruinelli L, et al. Artificial intelligence in nursing: priorities and opportunities from an international invitational think-tank of the Nursing and Artificial Intelligence Leadership Collaborative. J Adv Nurs 2021 Sep;77(9):3707-3717. [doi: 10.1111/jan.14855] [Medline: 34003504]
4. von Gerich H, Moen H, Block LJ, et al. Artificial intelligence-based technologies in nursing: a scoping literature review of the evidence. Int J Nurs Stud 2022 Mar;127:104153. [doi: 10.1016/j.ijnurstu.2021.104153] [Medline: 35092870]
5. Cabitza F, Rasoini R, Gensini GF. Unintended consequences of machine learning in medicine. JAMA 2017 Aug 8;318(6):517-518. [doi: 10.1001/jama.2017.7797] [Medline: 28727867]
6. Amann J, Blasimme A, Vayena E, Frey D, Madai VI, Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak 2020 Nov 30;20(1):310. [doi: 10.1186/s12911-020-01332-6] [Medline: 33256715]
7. Bussone A, Stumpf S, O'Sullivan D. The role of explanations on trust and reliance in clinical decision support systems. Presented at: 2015 International Conference on Healthcare Informatics (ICHI); Oct 21-23, 2015. [doi: 10.1109/ICHI.2015.26]
8. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020;3:17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
9. Parasuraman R, Riley V. Humans and automation: use, misuse, disuse, abuse. Hum Factors 1997 Jun;39(2):230-253. [doi: 10.1518/001872097778543886]
10. Lee JD, See KA. Trust in automation: designing for appropriate reliance. Hum Factors 2004;46(1):50-80. [doi: 10.1518/hfes.46.1.50_30392] [Medline: 15151155]
11. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc 2012;19(1):121-127. [doi: 10.1136/amiajnl-2011-000089] [Medline: 21685142]
12. Lyell D, Coiera E. Automation bias and verification complexity: a systematic review. J Am Med Inform Assoc 2017 Mar 1;24(2):423-431. [doi: 10.1093/jamia/ocw105] [Medline: 27516495]
13. Buçinca Z, Malaya MB, Gajos KZ. To trust or to think: cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. Proc ACM Hum-Comput Interact 2021 Apr 13;5(CSCW1):1-21. [doi: 10.1145/3449287]
14. Yin M, Wortman Vaughan J, Wallach H. Understanding the effect of accuracy on trust in machine learning models. Presented at: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19); May 4-9, 2019. [doi: 10.1145/3290605.3300509]

XSL•FO
RenderX

15. Lu Z, Yin M. Human reliance on machine learning models when performance feedback is limited: heuristics and risks. Presented at: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21); May 8-13, 2021. [doi: 10.1145/3411764.3445562]
16. Kücking F, Busch DA, Przysucha M, et al. Impact of AI recommendation correctness on diagnostic accuracy in clinical decision-making. Int J Med Inform 2026 Mar 1;207:106223. [doi: 10.1016/j.ijmedinf.2025.106223] [Medline: 41391283]
17. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nature 2023 Aug;620(7972):172-180. [doi: 10.1038/s41586-023-06291-2] [Medline: 37438534]
18. Hager P, Jungmann F, Holland R, et al. Evaluation and mitigation of the limitations of large language models in clinical decision-making. Nat Med 2024 Sep;30(9):2613-2622. [doi: 10.1038/s41591-024-03097-1] [Medline: 38965432]
19. Eriksen AV, Möller S, Ryg J. Use of GPT-4 to diagnose complex clinical cases. NEJM AI 2024 Jan;1(1):AIp2300031. [doi: 10.1056/AIp2300031]
20. Lu Z, Yin M. Trustworthy-ML: data and analysis code for "Human Reliance on Machine Learning Models When Performance Feedback is Limited: Heuristics and Risks". GitHub. 2021. URL: https://github.com/ZhuoranLu/Trustworthy-ML [accessed 2026-07-30]
21. Benner P. From Novice to Expert: Excellence and Power in Clinical Nursing Practice: Addison-Wesley; 1984.
22. Benner P. Using the Dreyfus model of skill acquisition to describe and interpret skill acquisition and clinical judgment in nursing practice and education. Bull Sci Technol Soc 2004 Jun;24(3):188-199. [doi: 10.1177/0270467604265061]
23. Bansal G, Wu T, Zhou J, et al. Does the whole exceed its parts? the effect of AI explanations on complementary team performance. Presented at: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21); May 8-13, 2021. [doi: 10.1145/3411764.3445717]
24. Vaccaro M, Almaatouq A, Malone T. When combinations of humans and AI are useful: a systematic review and meta-analysis. Nat Hum Behav 2024 Dec;8(12):2293-2303. [doi: 10.1038/s41562-024-02024-1] [Medline: 39468277]
25. Goh E, Gallo R, Hom J, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. JAMA Netw Open 2024 Oct 1;7(10):e2440969. [doi: 10.1001/jamanetworkopen.2024.40969] [Medline: 39466245]
26. Schemmer M, Kuehl N, Benz C, Bartos A, Satzger G. Appropriate reliance on AI advice: conceptualization and the effect of explanations. Presented at: Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI '23); Mar 27-31, 2023. [doi: 10.1145/3581641.3584066]
27. Kücking F, Hübner UH, Busch D. Diagnostic accuracy differences in detecting wound maceration between humans and artificial intelligence: the role of human expertise revisited. J Am Med Inform Assoc 2025 Sep 1;32(9):1425-1433. [doi: 10.1093/jamia/ocaf116] [Medline: 40668943]
28. Kahneman D. Thinking, Fast and Slow: Farrar, Straus and Giroux; 2011.
29. Buchanan C, Howitt ML, Wilson R, Booth RG, Risling T, Bamford M. Predicted influences of artificial intelligence on nursing education: scoping review. JMIR Nurs 2021;4(1):e23933. [doi: 10.2196/23933] [Medline: 34345794]
30. Gilbert S, Harvey H, Melvin T, Vollebregt E, Wicks P. Large language model AI chatbots require approval as medical devices. Nat Med 2023 Oct;29(10):2396-2398. [doi: 10.1038/s41591-023-02412-6] [Medline: 37391665]
31. Meskó B, Topol EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. NPJ Digit Med 2023 Jul 6;6(1):120. [doi: 10.1038/s41746-023-00873-0] [Medline: 37414860]
32. Küper A, Lodde GC, Livingstone E, Schadendorf D, Krämer N. Psychological factors influencing appropriate reliance on AI-enabled clinical decision support systems: experimental web-based study among dermatologists. J Med Internet Res 2025 Apr 4;27:e58660. [doi: 10.2196/58660] [Medline: 40184614]
33. Kücking F, Hübner U, Przysucha M, et al. Automation bias in AI-decision support: results from an empirical study. Stud Health Technol Inform 2024 Aug 30;317:298-304. [doi: 10.3233/SHTI240871] [Medline: 39234734]
34. Tajima H. Nursing-AI-error-thresholds: reproducibility package (calibration, bootstrap, 27-cell simulation, joint model, and sensitivity analyses). GitHub. 2026. URL: https://github.com/aiprof202604-tech/nursing-ai-error-thresholds [accessed 2026-07-30]
35. Tajima H. Theoretical exploration of error thresholds for clinical AI decision support in nursing: an exploratory simulation study grounded in human-AI reliance data. SSRN. Preprint posted online on Apr 24, 2026. [doi: 10.2139/ssrn.6632719]

Abbreviations

LLM: large language model
OR: odds ratio
RMSE: root-mean-square error

XSL•FO
RenderX

Edited by E Borycki, J Sarvestan; submitted 27.Apr.2026; peer-reviewed by P Quindroit, A Kim; revised version received 16.Jul.2026; accepted 17.Jul.2026; published 05.Aug.2026.

Please cite as:
Tajima H
Theoretical Exploration of Error Thresholds for Clinical AI Decision Support in Nursing: Exploratory Simulation Study Grounded in Human-AI Reliance Data
JMIR Nursing 2026;9:e99590
URL: https://nursing.jmir.org/2026/1/e99590
doi:10.2196/99590

© Hiroyuki Tajima. Originally published in JMIR Nursing (https://nursing.jmir.org), 5.Aug.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Community Health Nurses' Knowledge and Perceptions of AI in Canada: National Cross-Sectional Survey

Mary Henderson Betkus[1*], BA, BSCN, MBA, MSc; Davina Banner[1*], RN, PhD; Leanne Currie[2*], RN, PhD; Piper Jackson[3*], BA, PhD; Shannon Freeman[1*], PhD

[1]School of Nursing, University of Northern British Columbia, 3333 University Way, Prince George, BC, Canada

[2]School of Nursing, University of British Columbia, Vancouver, BC, Canada

[3]Department of Computing Science, Thompson Rivers University, Kamloops, BC, Canada

[*]all authors contributed equally

Corresponding Author:
Mary Henderson Betkus, BA, BSCN, MBA, MSc
School of Nursing, University of Northern British Columbia, 3333 University Way, Prince George, BC, Canada

Abstract

Background: Artificial intelligence (AI) continues to expand into nursing and health care. Many examples of AI applications driven by machine or deep learning are in use. Examples include wearable devices or alerts for risk prediction. AI tends to be promoted by nonnurses, creating a risk that AI is not designed to best serve registered nurses. Community health nurses (CHNs) are a small but essential group. CHNs' familiarity with AI and their perceptions about its effect on their practice are unknown.

Objective: The research aims to understand CHNs' awareness, knowledge, and perceptions of AI in practice and gain insights to better involve them in AI.

Methods: An online cross-sectional Canadian survey targeting CHNs was conducted from April to July 2023. Descriptive statistics summarized respondents' characteristics and perceptions of AI, followed by a chi-square test used to determine a relationship between respondents' level of AI knowledge and their AI perceptions, with odds ratio (OR) to determine the strength of association.

Results: A total of 228 CHNs participated with varying response rates per question. Most respondents were female (172/188, 91.5%), average age of 45.5 (SD 11.7) years, and an average of 13.5 (SD 10.1) years of community practice experience. Most respondents (205/228, 89.9%) felt they welcomed technology into their practice. They reported their understanding of AI technologies as "good" (95/220, 43.2%) and "not good" (125/220, 56.8%). Overall, 39.6% (80/202) of respondents felt uncomfortable with the development of AI. They agreed that AI should be part of education (143/203, 70.4%), professional development (152/202, 75.2%), and that they should be consulted (195/203, 96.1%). Many respondents had concerns related to professional accountability if they accepted a wrong AI recommendation (157/202, 77.7%) or if they dismissed a correct AI recommendation (149/202, 73.8%). Respondents with "good" AI knowledge were significantly associated with, and twice as likely to indicate nursing will be revolutionized (*P*=.007; OR 2.28, 95% CI 1.25-4.18), nursing will be more exciting (*P*=.001; OR 2.52, 95% CI 1.42-4.47), health care will be more exciting (*P*=.004; OR 2.3, 95% CI 1.30-4.06), and agreed that AI is part of nursing (*P*=.01; OR 2.1, 95% CI 1.19-3.68). Respondents with "not good" AI knowledge were significantly associated with, and more likely to feel uncomfortable with AI developments (χ^2_1=4.2, *P*=.04; OR 1.84, 95% CI 1.03-3.3).

Conclusions: CHNs reporting "good" AI knowledge had more favorable perceptions toward AI. Overall, CHNs had professional concerns about accepting or dismissing AI recommendations. Potential solutions include educational resources to ensure that CHNs have a sound basis for AI in their practice, which would promote their comfort with AI. Further research should explore how CHNs could be better involved in all aspects of AI introduced into their practice.

(JMIR Nursing 2026;9:e78560) doi:10.2196/78560

KEYWORDS

registered nurses; community health nurses; artificial intelligence; survey; machine learning

XSL•FO
RenderX

Introduction

Background

Artificial intelligence (AI) covers a broad array of AI-driven applications supported by machine learning (ML) or deep learning, which have potential utility in health care and nursing practice. Many examples of AI-driven care applications exist, from wearable devices for automated detection of signs and symptoms [1], automated assessment of outcomes to support the need for a different level of care [2,3], client-specific automated predictions of risks [4-7], and bots to answer inquiries and send reminders [8]. Despite widespread uptake and use, AI is commonly driven by nonnurses (ie, scientists, engineers, and the technology industry) [9,10] and physicians [11]. The lack of participation by registered nurses (RNs) creates a risk that AI will not be designed to best serve RNs who are expected to use AI applications and their outcomes in clinical practice [11-13]. Likewise, it is unknown whether community health nurses (CHNs) have thought about how AI applications could change their practice or how AI might be useful to inform clinical practice.

The community setting has a smaller group of RNs compared to the acute care sector [14]. CHNs are RNs who provide essential services in a variety of roles (eg, home health, public health, and primary care) within community settings (eg, clients' home and schools) [15]. Home health clients are most often older adults with multiple comorbidities [2], or individuals who have chronic [16,17] and unstable conditions [18]. Public health clients can be any age, as the focus of care is on promoting better health with service delivery to groups or individuals [15]. Within these settings, CHNs make the best care decisions based on the information that exists, as well as considering other subtleties that can affect these decisions. Regardless of the setting in the community, CHNs have increased autonomy [15,17,19], and clients have reduced nursing oversight because of time between visits [20]. This decreases the amount and frequency of the client-specific data collected. Further, CHNs focus on human connections and building trustful relationships while recognizing the strengths of individuals and communities to promote and improve their health [15]. These features support the importance of having CHNs who understand the practice area involved in AI.

Nursing research within the community sector is expanding to include a focus on the use of AI (eg, ML) as a method to improve real-time risk predictions [5,20,21] and to assist with better planning or targeting of service delivery [16]. Although involving CHNs would be key to raising the right questions for AI, as well as advising and validating results [22], few researchers are reporting this type of CHN involvement in AI development. More commonly, researchers are using existing collected data [16,21,23]. This passive involvement misses the opportunity of actively involving CHNs who are familiar with the data they collect and how it may add insight to clinical issues. However, in one example, a nonnurse researcher [24] describes using CHNs to advise and evaluate throughout an AI project, concluding that nursing input validated outcomes and facilitated acceptance of the AI algorithm into practice. Hence, nursing involvement provides a relevant perspective and knowledge that influences their informed decisions, which ensures clinical relevance and accuracy of AI and related ML [24,25]. These revelations add impetus to examine CHNs' perceptions of AI in their practice and to consider how they could be better involved.

Purpose Statement

This study aims to establish a baseline understanding of Canadian CHNs' awareness, knowledge, and perceptions of current and future effects of AI on their clinical practice. This will help to gain insights into how CHNs could be better involved in AI. Therefore, the research questions guiding this study include: (1) Are CHNs aware of the emergence of AI, including ML applications, in nursing? (2) What are CHNs' main sources of knowledge for learning about current day-to-day AI? (3) How do CHNs describe their level of knowledge of AI technologies? (4) Is there a relationship between CHNs' level of knowledge of AI technologies and their perceptions of the effects of AI on clinical practice, professional accountability, and the usefulness of AI applications? (5) What AI competencies do CHNs perceive as being needed in their community practice?

Methods

Ethical Considerations

Research approval was granted from the University of Northern British Columbia (UNBC) Research Ethics Board (REB 6009080), April 2023, to conduct a single cross-sectional open survey using SurveyMonkey licensed through UNBC. The survey landing page included an informational letter to provide study details. After reading the information on the landing page, respondents were asked to voluntarily consent electronically to participate in the study. Upon confirmation of informed consent, participants were then given access to this survey. If a participant did not consent, they received a thank-you message, and access to the survey closed automatically. All aspects of data collection, storage, and analysis were password-protected and housed on an encrypted UNBC server. The invitations advertised a random draw of 5 e-gift cards at the end of the survey period as an incentive to participate in the survey.

Instrument Design

A total of 11 research papers, which used surveys to examine attitudes and perceptions toward AI, were screened for relevance to this research study's instrument design. Two papers [26,27] were validating their General Attitudes Artificial Intelligence Scale to classify individuals with positive or negative feelings toward AI. The remaining 9 research studies targeted RNs [28], nursing students [29], radiologists [30,31], physicians [32], medical students [33], a mix of health care professionals [34,35], and consumers [36]. All except Swan [28] had their survey questions included in the publication or supplemental information. A request to preview and use Swan's survey, if applicable, was granted (BA Swan, RN, PhD, personal communication, November 30, 2022).

Swan's survey was selected due to its purposeful design for use with nursing professionals. It included similar questions to the other previewed surveys, indicating that common survey topics

XSL•FO
RenderX

were covered. Further, Swan's survey was adapted by adding questions important to this study. In the adapted survey, the first question to address computer expertise was sourced from Schepman and Rodway [27], who suggested that individuals with computer expertise would be more positive about AI. Questions 32 and 33 were added from Esmaeilzadeh [36] with slight modifications to address professional accountability. More details on the survey and its adaptation are found in Multimedia Appendix 1. Swan's survey had not been tested or piloted before deployment of the survey (BA Swan, RN, PhD, personal communication December 21, 2022).

The revised survey was reviewed for clarity by a retired community nursing manager with over 35 years of community experience, as a public health nurse in direct care and management. It was confirmed that the survey took 20 minutes to complete, and a direct question exploring how nurses should be involved with AI was suggested. Therefore, Q37 "How should nurses be involved in artificial intelligence that influences their practice?" was added.

The final version consisted of 37 content questions (referencing aspects of AI) plus a demographic section, which was used to describe respondents' representation across Canada, as well as their level of experience and current position. The survey was recreated on the survey platform. Complete wording of each survey item and types of questions are found in Multimedia Appendix 1.

Recruitment

The target population was RNs licensed in Canada who practiced in the community setting (eg, home care and public health) or RNs who had a community nursing focus (eg, researchers, educators, administrators, and clinical informatic nurses). Collectively, the term CHNs will be used. The survey was only offered in English. The size of the targeted population was unknown. Canadian workforce data reported that 32,074 direct care RNs were employed in community health in 2023 [14]; however, this does not account for the others not providing direct care (eg, researchers, educators, and administrators) included in the population of interest. Therefore, an online calculator [37] was used with the parameters of 20,000 for an unknown population, distribution at 50%, with 5% margin of error and a 95% CI, indicating a sample size of 377 was needed. The emergence of AI into clinical practice remains a new field. Therefore, the power analysis was a reference point to guide this exploratory research study.

The participants were recruited by an "invitation to participate letter," which had the live link to the survey embedded into its content. This was shared through nursing sources by monthly newsletters, email lists (eg, existing organizational and collegial connections), and informal networks (eg, colleague-to-colleague and social media). Two national organizations, Community Health Nurses of Canada and Canadian Nursing Informatics Association, canvassed their membership by broadcast messages and posts in their monthly e-newsletter. Each provincial and territorial nursing association or licensing body was contacted by email, briefly explaining the research and asking if they would circulate it to their members. Licensing bodies recommended that the researcher contact the nursing associations. One provincial licensing body agreed to send out the invitations by email to their members who identified as working in the community and had previously consented to be contacted for research purposes. The nursing associations kept the invitation in their monthly newsletters, or posts on their social media sites, or sent by broadcast message to their members until the survey closed. The survey was live from April 24 to July 30, 2023.

Data Management

On the survey closure date, the full dataset was exported from the survey platform to SPSS Statistics (version 29; IBM Corp). All computer IP addresses were removed, as well as respondents who provided consent but did not complete any survey questions. As it was expected that CHNs may complete this survey using a shared workstation, multiple responses from the same IP address were included as long as they were completed at different times, for different durations, and represented unique participant responses. The use of the same IP addresses was limited to 10 instances and met the above criteria. The geographical locations were grouped into regions to determine Canada-wide representation: Eastern (Prince Edward Island, Newfoundland & Labrador, New Brunswick, and Nova Scotia), Central (Ontario and Quebec), Western (Manitoba, Saskatchewan, Alberta, and British Columbia), and Northern (Yukon, Northwest Territories, and Nunavut). Questions that offered "other" as a choice were reviewed and recoded into the appropriate existing choices already provided; otherwise, it was left as "other." All word responses were coded for a numerical value to enable analysis (eg, Likert scale responses). Surveys that were blank (n=5) were removed. Cases with missing data (greatest in the demographic section) were kept, thus maximizing the number of responses for any given question. Therefore, the count n/N and percent are presented per question, except for multiple response questions, where n values and percent are given, because participants could respond to more than one option. Chi-square analysis was conducted to examine the relationship between respondents' reported AI knowledge (Q6) and respondents' perceptions of AI in their practice (Q10-Q20 and Q22-Q35, Q21 "other" was not included). All questions were examined for their missing or incomplete data. Variation in response rates could be due to respondents' choices not to answer or complete the survey. Therefore, to minimize the potential for response bias, all questions with a less than 15% missing data rate were kept. The core set of survey questions used to examine the research questions met this proportion of missing variables, with response rates as follows: Q6 (220/228, 96.5%), Q10-Q20 (206/228, 90.4% to 208/228, 91.2%), and Q22-Q35 (202/228, 88.6% to 207/228, 90.8%). The missing data for these questions is as follows: Q6 (3.5%), Q10-Q20 (range 8.8% to 9.6%), and for Q22-Q35 (range 9.2% to 11.4%).

Data Analysis

Both descriptive and inferential statistics were used to examine the data. Descriptive analysis summarized respondents' characteristics and their perceptions of AI in nursing. Inferential statistics examined the relationship between their reported AI

XSL•FO
RenderX

knowledge (Q6) and their perceptions of the current and future effects of AI on nursing and health care.

The chi-square test for independence was used to determine an association between CHNs reported level of knowledge of AI technologies (independent variables) and their perceptions of the effects of AI (dependent variables). The CHNs were grouped by their reported level of knowledge of AI technologies to allow for comparison. CHNs described their level of knowledge of AI technologies as "excellent," "very good," "good," "fair," or "none." They were grouped as "good" level of knowledge if they indicated "good" to "excellent" and "not good" level of knowledge if they indicated "fair" or "none." The reference category chi-square test for independence was a primarily "good" level of AI knowledge; however, a "not good" level of AI knowledge was the reference category for Q26 (comfort with AI development), Q32 (concern with AI offering wrong recommendation), and Q33 (concern with dismissing appropriate AI recommendation) to promote ease in explaining the results. All statements related to CHNs' perception or attitudes about AI were a 5-point Likert scale from strongly agree (5), agree (4), neutral (3), disagree (2), and strongly disagree (1). The responses for these questions were grouped as "agree" if the respondent indicated "agree" or "strongly agree" and grouped as "not agree" if they indicated "neutral," "disagree," or "strongly disagree." Neutral was grouped with "not agree" because it was interpreted that this group of respondents had no definitive feeling either way on the subject. As the aim of the research was to gain an understanding of how to better involve nurses in AI, it was concluded that these "neutral" respondents, along with "not agree," may need more targeted strategies to better involve them. Further, the transformed response "agree or not agree" was clarified by the sentiment being examined to ease understanding. The dependent variables were considered: comfortable or not comfortable with AI development, AI applications useful or not useful, effects of AI agree or not agree, and professional accountability concerned or not concerned. Odds ratios were calculated for chi-square tests that were significant to determine the strength of association.

Correction (ie, Yates and Bonferroni) methods for statistical testing were not used. Yates continuity correction was not used because the sample size was considered large enough (range 202 to 208) to support a Pearson chi-square [38]. It is noted that the item "nurse should be consulted" produced cells under 5 (not agree); however, this seemed a reasonable result and would not benefit from Yates correction. The Bonferroni post hoc was not used because it can be too restrictive [39]. The Bonferroni post hoc (0.05/25=0.002) is given for reference only and includes the 25 items (Q10-Q20 and Q22-Q35) examined for association.

The open-text question asking the respondents "How should registered nurses be involved in AI?" was examined for types of responses. Some examples of these responses included how CHNs could be engaged in AI technologies, for example, education, advising, or consulting. These responses were quantified with the frequencies reported.

Results

Overview

A total of 296 potential respondents opened the survey, 261 met recruitment criteria, with 233 (89.3%) providing consent. As reported, blank surveys (n=5) were not included. A total of 228 surveys were included in the analyses. The response rate fluctuated per question, with the response rate better at the start of the survey and waning by the final demographic section. The item, "community years experience," had the most nonresponses (52/228, 22.8%).

Sample Characteristics

Sample characteristics (Table 1) helped to describe the sample that responded to the survey. The respondents' average age was 45.5 (SD 11.7) years, with 58.4% (104/178) younger than 50 years. Most respondents identified as female (172/188, 91.5%). The average overall years of experience for RNs was 19.8 (SD 12.2) years, with most (161/179, 89.9%) ranging from 5 years to over 35 years of experience. For community practice, the average years of experience was 13.5 (SD 10.1) years, with many (129/176, 73.3%) ranging from 5 years to over 35 years of experience. The sample had representation from the 4 Canadian regions: Eastern (47/186, 25.3%), Central (72/186, 38.7%), Western (65/186, 34.9%), and Northern (2/186, 1.1%). The practice descriptions are multiple-response questions. The reported practice settings (Table 2) included public health (65/191, 22.3%), home care (56/191, 19.2%), community health centers (44/191, 15.1%), primary care (41/191, 14%), and case management (16/191, 5.5%). Approximately half indicated they provided direct care (108/191, 51.4%), and the majority (115/190, 60.5%) held a bachelor's degree.

XSL•FO
RenderX

Table . Sample characteristics of respondents.

Characteristic	Participants
Gender, n (%)[a]	
Male	16 (8.5)
Female	172 (91.5)
Age (years), means (SD)	45.5 (11.7)
Age (years), n (%)[b]	
25-29	13 (7.3)
30-34	26 (14.6)
35-39	29 (16.3)
40-44	17 (9.6)
45-49	19 (10.7)
50-54	30 (16.9)
55-59	21 (11.8)
60 and older	23 (12.9)
RN[c] experience (years), means (SD)	19.8 (12.2)
RN experience (years), n (%)[d]	
Less than 5 years	18 (10.1)
5-9	24 (13.4)
10-14	29 (16.2)
15-19	22 (12.3)
20-24	18 (10.1)
25-29	21 (11.7)
30-34	22 (12.3)
35 and greater	25 (14)
Community experience (years), means (SD)	13.5 (10.1)
Community experience (years), n (%)[e]	
Less than 5 years	47 (26.7)
5-9	24 (13.6)
10-14	30 (17)
15-19	26 (14.8)
20-24	13 (7.4)
25-29	19 (10.8)
30-34	13 (7.4)
35 and greater	4 (2.3)
Geographic location, n (%)[f]	
Eastern Canada	47 (25.3)
Central Canada	72 (38.7)
Western Canada	65 (34.9)
Northern Canada	2 (1.1)

[a]N=188.

[b]N=178.

[c]RN: registered nurse.

XSL•FO
RenderX

[d]N=179.
[e]N=176.
[f]N=186.

Table . Education and employment data of respondents.

Characteristic	Participants
Education level, n (%)[a]	
Diploma	27 (14.2)
Bachelor	115 (60.5)
Masters	38 (20)
Doctoral or PhD	10 (5.3)
Employment sector, n (%)[b,c]	
Public	140 (70.4)
Private	45 (22.6)
Academia	14 (7)
Practice setting, n (%)[b,d]	
Health informatics	12 (4.1)
Community health	44 (15.1)
Case management	16 (5.5)
Older adult	13 (4.5)
Home care	56 (19.2)
Hospice palliative	11 (3.8)
Primary care	41 (14)
Community mental health	9 (3.1)
Public health	65 (22.3)
College or university	18 (6.2)
Other	7 (2.4)
Current position (years), n (%)[b,d]	
Direct care	108 (51.4)
Nurse informatician	6 (2.9)
Manager or administrator	34 (16.2)
Staff education	23 (11)
Researcher	6 (2.9)
Faculty	19 (9)
Strategic planning	5 (4.3)
Other	9 (4.3)

[a]N=190.
[b]Multiple response questions, n summed in each section, may be greater than N.
[c]N=187.
[d]N=191.

Acceptance of Technology and Competent Users of Technology

The survey questions 1 and 2 were used to explore the CHNs' acceptance of technology into their practice, and how they described their computer use. Almost all participants (205/228, 89.9%) agreed or strongly agreed to welcoming technology into their practice. More than half (129/227, 56.8%) identified as competent users of the internet and standard applications, and another 36.6% (83/227) indicated they were users of specialist applications. The survey, included in Multimedia Appendix 1,

XSL•FO
RenderX

is subdivided into sections related to the headings addressing each of the research questions.

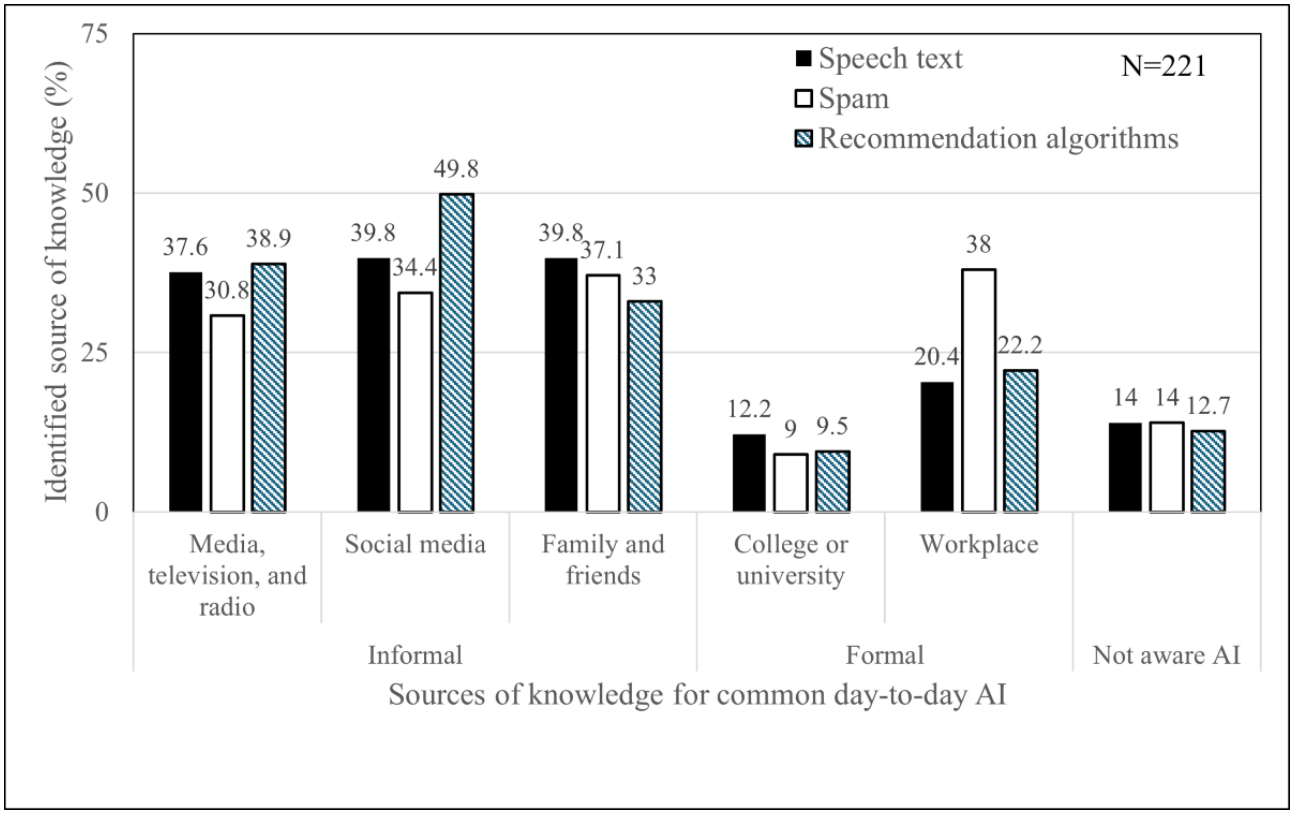
CHNs' Awareness of the Emergence of AI, Including ML Applications, in Nursing

CHNs' awareness of AI (Q8) in health care was more prevalent than their awareness of AI in nursing. The respondents were aware of AI (multiple response questions) in health care (123/220, 55.9%), but fewer were aware of AI in nursing (67/220, 30.5%). This was similar for ML and deep learning (Q9): respondents had heard of it in health care (84/220, 38.2%), and fewer had heard of it in nursing (35/220, 15.9%).

CHNs' Main Sources of Knowledge for Learning About Current Day-to-Day AI

The key sources of knowledge for learning about current day-to-day AI (Q3-Q5) varied between informal and formal methods. The respondents' major source of knowledge (multiple response questions) of common forms of day-to-day AI applications (Figure 1; ie, speech-text, spam, and recommendation algorithms) was informal resources such as media, television, or radio (range 68/221, 30.8% to 86/221, 38.9%); social media (range 76/221, 34.4% to 110/221, 49.8%); and family and friends (range 73/221, 33% to 88/221, 39.8%). Formal sources were indicated less often: colleges and universities (range 20/221, 9% to 27/221, 12.2%) and workplace (range 45/221, 20.4% to 84/221, 38%). It is worth noting that some respondents were not aware that these applications (speech-to-text, spam, and recommendation algorithms) were forms of AI (range 28/221, 12.7% to 31/221, 14%).

Figure 1. Respondents' key sources of knowledge for day-to-day artificial intelligence. AI: artificial intelligence.



CHNs' Description of Their Level of Knowledge of AI Technologies

Respondents described their understanding of the technologies used in AI (Q6) as none (42/220, 19.1%), fair (83/220, 37.7%), good (67/220, 30.5%), very good (23/220, 10.5%), and excellent (5/220, 2.3%). These results were grouped into 2 levels of AI knowledge: "good" level of knowledge included good to excellent (95/220, 43.2%), and "not good" level of knowledge used fair and none (125/220, 56.8%). Level of AI knowledge (Q6) was used in the chi-square test, as AI was more commonly known with a more balanced representation. For ML or deep learning (Q7), it was a similar trend, more nurses indicated "not good" level of knowledge (148/220, 67.3%) than "good" level of knowledge (72/220, 32.7%).

The Relationship Between CHNs' Level of Knowledge of AI Technologies and Their Perceptions of the Effects of AI on Clinical Practice, Professional Accountability, and the Usefulness of AI Applications.

Effects of AI on Clinical Practice

Questions 22 to 31 and 34 to 35 examined the respondents' perception of the effects of AI on their practice. An overview of respondents' perceptions indicated 39.6% (80/202) felt uncomfortable with the developments in AI, ML, and deep learning. Over half (133/206, 64.6%) of the respondents agreed that AI would revolutionize both health care and nursing. Few respondents agreed that the human nurse, 10.2% (21/205), or members of the interprofessional team, 12.6% (26/207), would

XSL•FO
RenderX

be replaced. Almost half of respondents felt AI would make nursing more exciting, 44.1% (89/202), and similarly, health care more exciting, 47.5% (96/202). Likewise, 44.8% (91/203) perceived AI to be part of nursing. Many respondents (143/203, 70.4%) felt that AI should be part of nursing education and included in professional development (152/202, 75.2%). Most respondents agreed they should be consulted (195/203, 96.1%) about AI, as well as having the opportunity to raise relevant nursing questions (189/202, 93.6%).

Examination with the chi-square test for independence (Table 3) was used to determine if there was a relationship between the respondents' reported AI knowledge and their perceptions of the potential effects of AI on clinical practice. For Q26, the reference category for level of AI knowledge was "not good." There was a significant relationship between respondents reporting "not good" level of AI knowledge and their perception of "feel uncomfortable" (ie, "agree" with statement) with AI developments (χ^2_1=4.2, *P*=.04; α=.05; small effect ϕ=.15). Respondents reporting "not good" AI knowledge were 1.84 (95% CI 1.03-3.3) times more likely to indicate developments in AI made them feel uncomfortable.

XSL•FO
RenderX

Table . Respondents' perceptions of current and future effects of AI[a] on clinical practice related to their level of AI knowledge.

Questions	Knowledge level	Effects		Chi-square (*df*)	Effect (ϕ)	OR[b] (95% CI)	*P* value
		Agree n (%)	Not agree n (%)				
Q22 revolutionize nursing				7.3 (1)	0.19	2.28 (1.25-4.18)	.007
	Good	66 (75)	22 (25)				
	Not good	67 (56.8)	51 (43.2)				
Q23 revolutionize health care				2.3 (1)	0.11	1.58 (0.88-2.86)	.13
	Good	62 (70.5)	26 (29.5)				
	Not good	71 (60.2)	47 (39.8)				
Q24 replace human RN[c]				0.85 (1)	0.07	1.53 (0.62-3.78)	.36
	Good	11 (12.5)	77 (87.5)				
	Not good	10 (8.5)	107 (91.5)				
Q25 replace interprofessional team member				0.2 (1)	0.03	1.18 (0.52-2.70)	.69
	Good	12 (13.6)	76 (86.4)				
	Not good	14 (11.8)	105 (88.2)				
Q26 uncomfortable with AI developments[d]				4.2 (1)	0.15	1.84 (1.03-3.3)	.04
	Not good	53 (45.7)	63 (54.3)				
	Good	27 (31.4)	59 (68.6)				
Q27 nursing will be more exciting				10.1 (1)	0.22	2.52 (1.42-4.47)	.001
	Good	49 (57)	37 (43)				
	Not good	40 (34.5)	76 (65.5)				
Q28 health care will be more exciting				8.3 (1)	0.20	2.3 (1.30-4.06)	.004
	Good	51 (59.3)	35 (40.7)				
	Not good	45 (38.8)	71 (61.2)				
Q29 AI is part of nursing practice				6.6 (1)	0.18	2.1 (1.19-3.68)	.01
	Good	48 (55.2)	39 (44.8)				
	Not good	43 (37.1)	73 (62.9)				
Q30 AI included in nursing education				0.7 (1)	0.06	1.3 (0.71-2.4)	.40
	Good	64 (73.6)	23 (26.4)				
	Not good	79 (68.1)	37 (37.9)				
Q31 AI included in professional development				2.0 (1)	0.10	1.6 (0.83-3.14)	.16
	Good	69 (80.2)	17 (19.8)				
	Not good	83 (71.6)	33 (28.4)				
Q34 nurses should be consulted				0.2 (1)	−0.03	0.74 (0.18-3.05)	.68
	Good	83 (95.4)	4 (4.6)				
	Not good	112 (96.6)	4 (3.4)				
Q35 identify relevant AI nursing questions				2.0 (1)	−0.10	0.44 (0.14-1.39)	.15
	Good	78 (90.7)	8 (9.3)				
	Not good	111 (95.7)	5 (4.3)				

XSL•FO
RenderX

[a]AI: artificial intelligence.

[b]OR: odds ratio.

[c]RN: registered nurse.

[d]Reference category was set to ”good” for all variables with the exception of Q26 where the reference category was set to “not good.”

The remaining statements (Q22-Q25 and Q27-Q35) used the reference category “good” level of AI knowledge. There were significant relationships between “good” level of AI knowledge and the following perceptions. Respondents perceived AI would revolutionize nursing (χ^2_1=7.3, *P*=.007; α=.05; small to moderate effect ϕ=.19) and were 2.28 times more likely to agree that nursing would be revolutionized (95% CI 1.25-4.18). Respondents perceived AI would make both nursing (χ^2_1=10.1, *P*=.001, α=.05, small to moderate effect ϕ=.22) and health care (χ^2_1=8.3, *P*=.004, α=.05, small to moderate effect ϕ=.20) more exciting. Respectively, these respondents were 2.52 times more likely (95% CI 1.42-4.47) and 2.3 times more likely (95% CI 1.30-4.06) to perceive that AI will make nursing and health care more exciting. These respondents perceived that AI is part of nursing practice (χ^2_1=6.6, *P*=.01; α=.05; small to moderate effect ϕ=.18) and were 2.1 times more likely to agree that AI is part of nursing practice (95% CI 1.19-3.68).

There was no association observed between level of AI knowledge and perceived effects: for revolutionizing health care (*P*=.13) nor between level of AI knowledge and perceived effects for replacing human RN (*P*=.36) or replacing interprofessional team members (*P*=.69). There was no association between level of AI knowledge and perception that AI should be part of nursing education (*P*=.4), part of professional development (*P*=.16), nurses should be consulted (*P*=.68), or nurses should identify relevant nursing questions for AI (*P*=.15).

Professional Accountability

Two statements (Q32 and Q33) used “what if” scenarios to examine CHNs’ perceptions of AI and professional accountability. One described an AI providing the wrong recommendation, and the other described a correct recommendation that was dismissed by the nurse. Respondents expressed concern regarding their responsibility in both scenarios. The majority, 77.7% (157/202), were concerned if AI offered the wrong recommendation, and likewise, 73.8% (149/202), if an appropriate AI recommendation was dismissed. Examination with chi-square test for independence (Table 4) with “not good” as the reference category revealed no association between level of AI knowledge and perceived concern if AI provided a wrong recommendation (*P*=.06). Conversely, the chi-square test for independence suggested a significant association between a “not good” level of AI knowledge and perceived concern if a correct recommendation was dismissed (χ^2_1=3.98, *P*=.046; α=.05; small effect ϕ=.14). Respondents reporting “not good” AI knowledge were 1.9 times more likely to be concerned with dismissing an appropriate AI recommendation (95% CI 1.01-3.57).

Table . Respondents’ perceptions of concern with professional accountability related to their level of AI[a] knowledge.

Questions	Knowledge level	Concern		Chi-square (*df*)	ϕ	OR[b] (95% CI)	*P* value
		Agree n (%)	Not agree n (%)				
Q32 if AI offers wrong recommendations				3.7 (1)	0.14	1.9 (0.98-3.74)	.06
	Not good	95 (82.6)	20 (17.4)				
	Good	62 (71.3)	25 (28.7)				
Q33 if correct recommendation is dismissed				3.98 (1)	0.14	1.9 (1.01-3.57)	.046
	Not good	91 (79.1)	24 (20.9)				
	Good	58 (66.7)	29 (33.3)				

[a]AI: artificial intelligence.

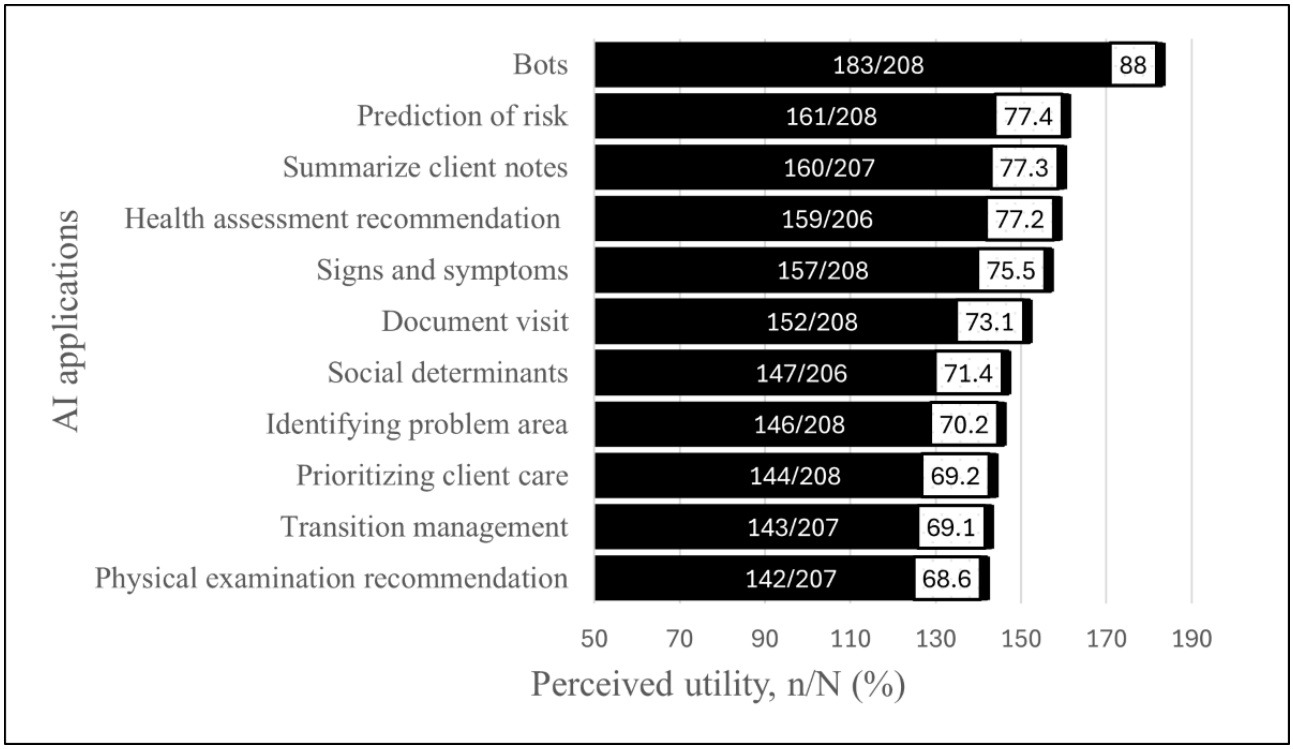
[b]OR: odds ratio.

Usefulness of AI Applications

Q10-Q20 examined the respondents’ perceptions of the utility of various AI applications. Respondents perceived that overall, each AI application would be useful (Figure 2), with agreement ranging from 68.6% (142/207) to 88% (183/208). Most respondents indicated Q15 bots (183/208, 88%), Q18 risk prediction (161/208, 77.4%), and Q20 summarizing narrative text from a client’s notes (160/207, 77.3%) would be useful. Further examination to determine if the level of AI knowledge was associated with CHNs’ perception of the utility of AI application revealed that in all but one example, there was no association between level of AI knowledge and their perception of utility (Table 5). There was a significant association between a “good” level of AI knowledge and perception of utility for Q13 transition management (χ^2_1=7.9, *P*=.005, α=.05, small to moderate effect ϕ=.2). Respondents reporting “good” AI knowledge were 2.45 times more likely to agree that transition management would be useful (95% CI 1.3-4.63).

XSL•FO
RenderX

Figure 2. Respondents' perception of the utility of AI applications. AI: artificial intelligence.



XSL•FO
RenderX

Table . Respondents' perception of the utility of AI[a] applications related to their level of AI knowledge.

Questions	Knowledge level	Utility of AI		Chi-square (*df*)	ϕ	OR[b] (95% CI)	*P* value
		Agree n (%)	Not agree n (%)				
Q10 signs and symptoms				0.27 (1)	0.034	1.18 (0.62-2.26)	.61
	Good	68 (77.3)	20 (22.7)				
	Not good	89 (74.2)	31 (25.8)				
Q11 social determinants				0.8 (1)	0.06	1.3 (0.72-2.48)	.36
	Good	65 (74.7)	22 (25.3)				
	Not good	82 (68.9)	37 (31.1)				
Q12 prioritizing client care				0.9 (1)	0.07	1.3 (0.73-2.45)	.35
	Good	64 (72.7)	24 (27.3)				
	Not good	80 (66.7)	40 (33.3)				
Q13 transition management				7.9 (1)	0.20	2.45 (1.30-4.63)	.005
	Good	70 (79.5)	18 (20.5)				
	Not good	73 (61.3)	46 (38.7)				
Q14 problem area				3.7 (1)	0.13	1.83 (0.98-3.42)	.06
	Good	68 (77.3)	20 (22.7)				
	Not good	78 (65)	42 (35)				
Q15 bots				0.03 (1)	−0.01	0.93 (0.40-2.15)	.86
	Good	77 (87.5)	11 (12.5)				
	Not good	106 (88.3)	14 (11.7)				
Q16 health assessments				0.1 (1)	−0.02	0.90 (0.47-1.74)	.76
	Good	67 (76.1)	21 (23.9)				
	Not good	92 (78)	26 (22)				
Q17 physical assessment				0.04 (1)	−0.01	0.94 (0.52-1.7)	.84
	Good	59 (67.8)	28 (32.2)				
	Not good	83 (69.2)	37 (30.8)				
Q18 prediction of risk				0.4 (1)	0.04	1.24 (0.64-2.41)	.53
	Good	70 (79.5)	18 (20.5)				
	Not good	91 (75.8)	29 (24.2)				
Q19 documentation of visit				3.2 (1)	0.13	1.8 (0.95-3.44)	.07
	Good	70 (79.5)	18 (20.5)				
	Not good	82 (68.3)	38 (31.7)				
Q20 summarize client notes				1.6 (1)	0.09	1.54 (0.78-3.05)	.21
	Good	71 (81.6)	16 (18.4)				
	Not good	89 (74.2)	31 (25.8)				

[a]AI: artificial intelligence.

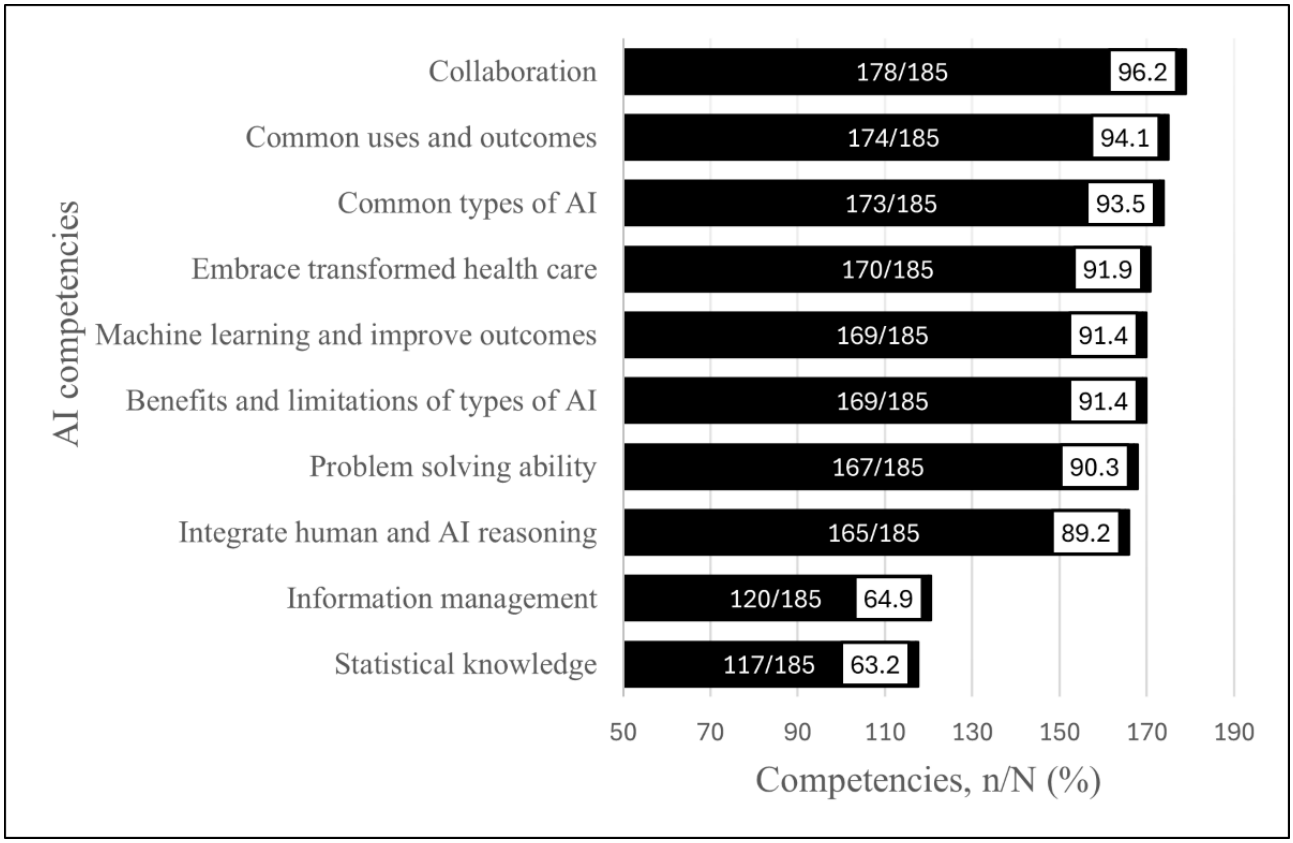
[b]OR: odds ratio.

AI Competencies CHNs Perceive as Being Needed in Their Community Practice

The survey (Q36) offered 10 competencies for respondents to indicate which were needed by CHNs (multiple-response question). The 3 competencies most identified as needed (Figure 3) were (1) communications, collaboration, and cross-functional knowledge (178/185, 96.2%); (2) knowledge of common uses and outcomes of AI (174/185, 94.1%); and (3) knowledge of common types of AI (173/185, 93.5%). The competency identified least was statistical knowledge, which also covered skills related to clinical analytics, data management, and algorithm awareness (117/185, 63.2%). Complete wording of each competency is found in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 3. Competencies identified by respondents as needed by CHNs to integrate AI into clinical practice. AI: artificial intelligence; CHN: community health nurse.



Insights on How CHNs Could Be Better Involved in AI

An open-ended question (Q37) asked how respondents thought they should be involved. It produced (70/228, 30.7%) responses, which provided insights into nurse involvement and their perspectives on related aspects of AI in their practice. Respondents expressed a need for further education (21/70, 30%) using phrases like “learn,” “knowledge acquisition,” “stay up to date,” and “education.” Most respondents (57/70, 81.4%) cited numerous roles or functions where nurses should be involved: raising relevant questions (5/70, 7.1%); advising and consulting (24/70, 34.3%); planning, development, and implementing (14/70, 20%); evaluation (12/70, 17.1%); change management (5/70, 7.1%); regulation, policy, and ethics (7/70, 10%); and all phases (13/70, 18.6%). They used terms like “key stakeholders” and “subject matter experts.” They felt nurses needed to be involved to make AI relevant. Respondents (7/70, 10%) specifically identified that direct care (front-line and end user) CHNs should be involved. Some respondents (8/70, 11.4%) referred to AI as a tool or an additional resource. Other respondents (13/70, 18.6%) acknowledged their apprehension with AI being introduced into practice. Respondents (6/70, 8.6%) referred to the need to be mindful about the human relationship with phrases like “relationships are key aspects of community health nursing” and “human connection care can not be replaced.”

Discussion

Principal Findings

The main findings indicated CHNs differ in their level of knowledge and perceptions of AI technologies in nursing and health care. Many CHNs have a limited awareness of AI emerging in health care and report even less awareness of AI emerging in nursing practice. The main sources of information for day-to-day AI applications are predominantly informal methods (eg, social media) compared to academic and workplace sources. Some CHNs are unaware that common day-to-day applications are AI-driven. Fewer CHNs describe their knowledge of AI technologies as “good.” However, the CHNs who describe their AI knowledge as “good” are twice as likely to be optimistic or have favorable perceptions of AI effects, such as revolutionizing nursing, making nursing more exciting, and agreeing that AI is part of nursing. Whereas CHNs with “not good” AI knowledge are almost twice as likely to feel uncomfortable with AI development. Regardless of the level of AI knowledge, most CHNs agree they should be involved in AI by consulting and raising nurse-relevant questions in various phases of AI development, such as implementation and ongoing evaluation. The results substantiate the need for appropriate AI education for CHNs to prepare them to participate in AI that will influence their practice.

CHNs have a limited awareness of AI emerging in nursing practice (30.5%), which aligns with results found in similar nursing research [28,40]. However, other questions in this current research are used to gain further insights into why their

XSL•FO
RenderX

understanding of various AI technologies might be limited. CHNs use informal (eg, social media, and family and friends) methods of learning about common day-to-day AI applications, with 12.7% (28/221) to 14% (31/221) of respondents being unaware that these common forms (ie, speech to text, spam, and recommendation algorithms) are driven by AI. This limited awareness could be related to relying on informal sources of knowledge. CHNs may turn to readily available sources of information because of convenience. Likewise, being aware of spam from work-related sources could be as simple as "don't see a reply, check your spam folder," while having no real understanding of the algorithms that recognize and reroute spam. This lack of understanding whether an application is driven by AI has been linked to clinical practice by another study [40] where 22% of Canadian nurses did not know if AI is used in their practice area. Similarly, Coakley et al [31] identified that approximately 40% of radiographers did not recognize work-related AI-driven applications. This raises a potential concern that CHNs may be using AI-driven applications within their practice unbeknownst to them. Lastly, over half (125/220, 56.8%) of CHNs describe their knowledge of AI technologies as "not good." This limited awareness of AI in nursing and lack of knowledge of AI technology highlights a knowledge deficit, stressing the importance of AI education for CHNs.

The composition of the survey sample strengthens the clinical value of the results. First, this Canadian sample is an experienced group of CHNs, both in years of practice as an RN and years of experience in the community sector. They describe themselves as competent and welcoming of technology. This was expected because Canada has been striving since 2000 to improve digital health connections (eg, electronic health records) within the Canadian health care system [41]. A current report [40] confirms a continual uptake in digital technology. This steady increase of new technologies into practice (eg, electronic health records and electronic assessments) emphasizes CHNs' adaptability and resiliency to new technologies in their practice, considering these decisions are made at higher levels in the organization rather than from staff who are expected to use them [42]. Second, more than half of the survey respondents provide direct care services. This means they are familiar with community practice, its clinical data, and provision of care at the client level, and have the potential to offer pragmatic insights. Third, this group of CHNs includes end users who are seldom involved in the development of AI. They are, however, important stakeholders in ensuring clinical relevance in new technology [22,24]. The various characteristics (eg, experienced, competent, and end users) of this CHN sample provide validity and relevance to the results.

A common technique for assessing the level of knowledge across surveys is asking the respondent to indicate their level of AI knowledge. Most surveys use this subjective method, finding fewer respondents rate their level of AI knowledge as "good" compared to "not good" level of knowledge) [28,32,40,43,44], aligning with the current study (level of knowledge "good" 95/220, 43.2% versus "not good" 125/220, 56.8%). None of the cited surveys uses the difference in knowledge level to compare groups and their perceptions.

The subjective evaluation of CHNs' level of AI knowledge may be underestimated or overestimated. However, professionally, CHNs self-reflect on practice and learning gaps, so they have familiarity in evaluating their competencies [45]. It seems plausible to use the self-identified AI knowledge level as a starting point to determine if there is a relationship between the level of knowledge and CHNs' perceptions of AI. The 2 groups of "good" and "not good" knowledge level of AI technologies in this study suggest that the level of AI knowledge affects some of the AI perceptions of CHNs.

The CHNs reporting "not good" level of knowledge are almost two times more likely to indicate that they are uncomfortable with the developments in AI. Intuitively, this makes sense. It can be argued that having "good" AI knowledge provides a method to evaluate the benefits or disadvantages of AI and perhaps provides some control [46]. CHNs reporting a "good" level of knowledge are more than two times more likely to feel nursing will be revolutionized, nursing and health care will become more exciting, and agree that AI is part of nursing practice. Therefore, CHNs with a "good" level of AI knowledge are more optimistic about the future effects of AI [46]. The differences between AI perceptions for CHNs with a "good" level of knowledge versus a "not good" level further stress the necessity for education and ongoing learning opportunities to decrease apprehension and promote optimism around AI [46].

Regardless of their level of knowledge, few CHNs believe that human RNs (21/205, 10.2%) or interprofessional team members (26/207, 12.6%) will be replaced by AI. This sentiment aligns with that of Swan [28]. The underlying belief that human touch is integral to nursing care, along with humans' ability to reconsider and change care when an unexpected situation arises, supports human RNs and other interprofessional members' continued importance to the care team [47-49]. CHNs' responses (Q37) defend the importance of human involvement: "relationships are a key aspect of community health nursing" and "human connection care can not be replaced." CHNs' belief that they will not be replaced does not address how they think their role within health care will change. This aspect should be examined in future research.

Professional accountability is a central feature for all regulated professionals. Several studies include some reference to the issue (eg, medical liability). This current study demonstrates a mixed outcome. There is no association between the level of AI knowledge and concern with AI providing a wrong recommendation, versus an association between the level of AI knowledge and concern with dismissing a correct recommendation. Still, the clinical importance should be addressed because most CHNs, regardless of their level of AI knowledge, have concerns about who would be held responsible for either accepting a wrong recommendation (157/202, 77.7%) or dismissing a correct recommendation (149/202, 73.8%). Other studies confirm that professionals have concerns about the use of AI in practice [32,40,43]. Further research needs to explore what CHNs feel they require to help address and remediate their concerns.

Although CHNs have a limited understanding of AI, more than two-thirds perceived the examples of AI-driven applications as

XSL•FO
RenderX

useful. It suggests that, superficially, they perceived a value in the application to their practice setting. Positive perceptions of the utility of AI applications trend across surveys [28,33,40,43]. However, it is unknown why CHNs perceive the AI applications as useful to them; that is, whether it is the function of replacing a task or the function of supporting decision-making that is important. Future research should follow the open text response "start by asking nurses what they feel could be automated," thus gaining an understanding of what makes an application useful.

A more comprehensive picture of how CHNs should be involved in AI becomes apparent through the open text responses. First, CHNs confirm their need to "learn," for "knowledge acquisition," to "stay up to date," and "education" as important prerequisites to being involved in AI. Although the quantitative sample identifies the necessity of AI to be included in nursing education and professional development, the open text responses connect AI education to facilitating CHNs' involvement (participation). Second, the open text responses share a wide range of ways CHNs can be involved. It verifies the importance of including all levels of nurses, specifically noting direct care nurses, and validates that CHNs need to be engaged during all phases of AI. Areas of involvement include: raising relevant questions, planning, development, implementing, evaluation, and monitoring to ensure AI is clinically relevant and accurate. Third, they recognize that their involvement includes regulation, workplace policies, and ethical frameworks to guide their practice because AI is a tool. Fourth, they readily admit to being apprehensive, citing concerns with loss of some of their skills (eg, assessment) along with the human connection and relationship with clients because of AI. This loss of human connection because of technology is also a common concern explored in the literature [49,50]. Further research should continue to examine how CHNs can be better involved.

Clinical Implications and Recommendations

This research reveals 2 interrelated concepts, preparation and participation. These are both essential to better involve CHNs. The first, preparation, acknowledges the importance of education and ongoing professional development. This builds the foundation that will support CHNs to become involved. AI needs to be consciously integrated both in nursing education and ongoing professional development, with attention to a standardized curriculum to ensure all nurses have a basic understanding of AI. Specific areas of concentration should address professional accountability. This will provide CHNs with knowledge to evaluate AI outputs as part of their decision-making, as well as planning and ameliorating perceived future effects of AI. The second, participation, addresses the various aspects of involving CHNs to identify relevant questions and to contribute their nursing perspective to all phases of development and implementation of AI. Professional nursing groups and health care organizations are instrumental in ensuring that the right mix of CHNs, from end user to leadership, have participation on AI advisory committees. Although this research was initiated to examine the perceptions of CHNs about AI in clinical practice, it now raises the necessity of further research to expand on these results by conducting small group consultations to gain an in-depth understanding of how best to involve CHNs.

Strengths and Limitations

The strengths of this study include establishing baseline knowledge and perceptions of AI among Canadian CHNs. An effort was made to recruit the appropriate sample population by targeting national nursing groups as well as provincial and territorial nursing organizations. This survey identifies the need for appropriate education (preparation) and confirms that CHNs want to be involved (participation). It explores the use of the self-reported level of knowledge to determine differences between the "good" and "not good" levels of knowledge.

Several limitations exist. The original survey that was foundational to this study did not have psychometric or reliability testing done. Further testing and reporting of reliability is recommended as a future step. In this variation of the survey, only 1 experienced community nurse was used to determine face validity. A notable limitation of the research is the length of the instrument, as nonresponses increased as the survey progressed. The multiple response questions allowed for several responses, which may have blurred the interpretation. For example, under "current position," a respondent could have 2 different positions within the community, for example, direct care and educator. Self-reported knowledge is subjective; we are unable to verify what the respondents know or do not know or evaluate the expertise of their knowledge. However, CHNs who feel they have knowledge were more favorable or optimistic about AI within their practice. This study did not examine whether respondents had AI-related practical experience or whether current AI is integrated into their practice. Each survey statement or question is briefly explained or described, for example, "AI will revolutionize nursing by supporting health promotion and disease prevention, helping create personalized treatment plans, speeding up administrative tasks." Each respondent could interpret it differently depending on their understanding of how this might occur and their experience with any of the concepts in the descriptor. There were no respondents identified from Prince Edward Island, Saskatchewan, or the Northwest Territories. With a nonresponse rate (42/228, 18.4%) for this question, it could not be determined where the missing respondents were located. An online survey has challenges. The recommended number of 377 respondents to have a 5% error margin was not achieved. The true response rate is unknown because it is unknown how many CHNs received the recruitment invitations due to the method used to recruit respondents. Respondents' bias or selective reporting may have occurred because it was an online survey; only nurses who could access the survey could respond. Additionally, using the term AI in the survey title may have only interested a select group of nurses. As well, the survey was only offered in English, limiting the participation and insights from Francophone nursing colleagues. Lastly, for analysis, the chi-square test can only test for the association of the categorical variables, not causation.

Conclusions

The survey results provide insights into the proposed research questions. Only a third of CHNs are aware that AI is emerging in nursing practice. CHNs use informal sources of knowledge (eg, family and friends) to learn about day-to-day AI applications, with some unaware that these day-to-day

XSL•FO
RenderX

applications are AI-driven. This raises the concern that CHNs may be using AI in their practice without realizing that the technology they are using is AI-based. CHNs who report better AI knowledge tend to be more optimistic (ie, “more exciting”) and less uncomfortable about AI and its effects on practice. However, many CHNs have concerns with AI and their professional accountability. Many CHNs agree that AI as a topic should be included in nursing education as well as professional development. This study identifies that most CHNs want to be involved in AI, highlighting that they want to be consulted and given opportunities to raise nurse-relevant questions. An important step to better involve CHNs should address the availability of appropriate and consistent education. This will help to promote the awareness of AI in nursing and alleviate professional concerns, thus preparing CHNs to be better involved.

Acknowledgments

The authors thank all respondents for taking the time to participate in this research.

Funding

No external financial support or grants were received from any public commercial or not-for-profit entities for the research, authorship, or publication of this paper.

Data Availability

The data are presented in the main paper.

Authors' Contributions

Conceptualization: MHB, LC, SF
Data curation: MHB, SF
Formal analysis: MHB, SF, LC, PJ
Investigation: MHB
Methodology: MHB, SF, DB, LC, PJ
Project administration: MHB
Supervision: SF, DB, LC, PJ
Validation: MHB, SF, DB, LC, PJ
Visualization: MHB, SF, DB, LC, PJ
Writing – original draft: MHB
Writing – review & editing: MHB, SF, DB, LC, PJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey instrument.
[DOCX File, 39 KB - nursing_v9i1e78560_app1.docx]

Checklist 1
CHERRIES checklist.
[DOCX File, 26 KB - nursing_v9i1e78560_app2.docx]

References

1. Bat-Erdene BO, Saver JL. Automatic acute stroke symptom detection and emergency medical systems alerting by mobile health technologies: a review. J Stroke Cerebrovasc Dis 2021 Jul;30(7):105826. [doi: 10.1016/j.jstrokecerebrovasdis.2021.105826] [Medline: 33932749]
2. Bowles KH, Ratcliffe SJ, Naylor MD, Holmes JH, Keim SK, Flores EJ. Nurse generated EHR data supports post-acute care referral decision making: development and validation of a two-step algorithm. AMIA Annu Symp Proc 2017;2017:465-474. [Medline: 29854111]
3. Kang DY, Cho KJ, Kwon O, et al. Artificial intelligence algorithm to predict the need for critical care in prehospital emergency medical services. Scand J Trauma Resusc Emerg Med 2020 Mar 4;28(1):17. [doi: 10.1186/s13049-020-0713-4] [Medline: 32131867]
4. Choi J, Choi J, Jung HT. Applying machine-learning techniques to build self-reported depression prediction models. Comput Inform Nurs 2018 Jul;36(7):317-321. [doi: 10.1097/CIN.0000000000000463] [Medline: 29985815]

XSL•FO
RenderX

5. Kang Y, McHugh MD, Chittams J, Bowles KH. Utilizing home healthcare electronic health records for telehomecare patients with heart failure: a decision tree approach to detect associations with rehospitalizations. Comput Inform Nurs 2016 Apr;34(4):175-182. [doi: 10.1097/CIN.0000000000000223] [Medline: 26848645]
6. Su D, Zhang X, He K, Chen Y. Use of machine learning approach to predict depression in the elderly in China: a longitudinal study. J Affect Disord 2021 Mar 1;282:289-298. [doi: 10.1016/j.jad.2020.12.160] [Medline: 33418381]
7. Hu M, Shu X, Yu G, Wu X, Välimäki M, Feng H. A risk prediction model based on machine learning for cognitive impairment among Chinese community-dwelling elderly people with normal cognition: development and validation study. J Med Internet Res 2021 Feb 24;23(2):e20298. [doi: 10.2196/20298] [Medline: 33625369]
8. Newman MB. Artificial intelligence and case management: embracing transformative practice. Prof Case Manag 2021;26(6):304-306. [doi: 10.1097/NCM.0000000000000533] [Medline: 34609344]
9. Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. J Med Ethics 2020 Mar;46(3):205-211. [doi: 10.1136/medethics-2019-105586] [Medline: 31748206]
10. Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. Science 2015 Jul 17;349(6245):255-260. [doi: 10.1126/science.aaa8415] [Medline: 26185243]
11. Shang Z. A concept analysis on the use of artificial intelligence in nursing. Cureus 2021 May 5;13(5):e14857. [doi: 10.7759/cureus.14857] [Medline: 34113496]
12. Barda AJ, Horvat CM, Hochheiser H. A qualitative research framework for the design of user-centered displays of explanations for machine learning model predictions in healthcare. BMC Med Inform Decis Mak 2020 Oct 8;20(1):257. [doi: 10.1186/s12911-020-01276-x] [Medline: 33032582]
13. Stead WW. Clinical implications and challenges of artificial intelligence and deep learning. JAMA 2018 Sep 18;320(11):1107-1108. [doi: 10.1001/jama.2018.11029] [Medline: 30178025]
14. Nursing in Canada, Data Tables. : Canadian Institute for Health Information; 2024 URL: https://www.cihi.ca/sites/default/files/document/nursing-in-canada-2024-meth-notes-en.pdf [accessed 2025-12-29]
15. 2019 Canadian community health nursing professional practice model and standards of practice. CHNC / IISCC. 2019. URL: https://www.chnc.ca/en/standards-of-practice [accessed 2025-12-28]
16. Bose E, Radhakrishnan K. Using unsupervised machine learning to identify subgroups among home health patients with heart failure using telehealth. Comput Inform Nurs 2018 May;36(5):242-248. [doi: 10.1097/CIN.0000000000000423] [Medline: 29494361]
17. Petit dit Dariel OJ, Raby T, Ravaut F, Rothan-Tondeur M. Developing the Serious Games potential in nursing education. Nurse Educ Today 2013 Dec;33(12):1569-1575. [doi: 10.1016/j.nedt.2012.12.014] [Medline: 23332500]
18. Zolnoori M, McDonald MV, Barrón Y, et al. Improving patient prioritization during hospital-homecare transition: protocol for a mixed methods study of a clinical decision support tool implementation. JMIR Res Protoc 2021 Jan 22;10(1):e20184. [doi: 10.2196/20184] [Medline: 33480855]
19. Sockolow PS, Bowles KH, Topaz M, et al. The time is now: informatics research opportunities in home health care. Appl Clin Inform 2021 Jan;12(1):100-106. [doi: 10.1055/s-0040-1722222] [Medline: 33598906]
20. Topaz M, Woo K, Ryvicker M, Zolnoori M, Cato K. Home healthcare clinical notes predict patient hospitalization and emergency department visits. Nurs Res 2020;69(6):448-454. [doi: 10.1097/NNR.0000000000000470] [Medline: 32852359]
21. Sullivan SS, Hewner S, Chandola V, Westra BL. Mortality risk in homebound older adults predicted from routinely collected nursing data. Nurs Res 2019;68(2):156-166. [doi: 10.1097/NNR.0000000000000328] [Medline: 30531348]
22. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. Nat Med 2019 Sep;25(9):1337-1340. [doi: 10.1038/s41591-019-0548-6] [Medline: 31427808]
23. Hobensack M, Song J, Scharp D, Bowles KH, Topaz M. Machine learning applied to electronic health record data in home healthcare: a scoping review. Int J Med Inform 2023 Feb;170:104978. [doi: 10.1016/j.ijmedinf.2022.104978] [Medline: 36592572]
24. Sinn CLJ, Jones A, McMullan JL, et al. Derivation and validation of the personal support algorithm: an evidence-based framework to inform allocation of personal support services in home and community care. BMC Health Serv Res 2017 Nov 25;17(1):775. [doi: 10.1186/s12913-017-2737-7] [Medline: 29178868]
25. Douthit BJ, Hu X, Richesson RL, et al. How artificial intelligence is transforming the future of nursing. American Nurse Journal. 2020 Sep 6. URL: https://www.myamericannurse.com/how-artificial-intelligence-is-transforming-the-future-of-nursing/ [accessed 2025-01-06]
26. Schepman A, Rodway P. Initial validation of the General Attitudes Towards Artificial Intelligence Scale. Comput Hum Behav Rep 2020;1:100014. [doi: 10.1016/j.chbr.2020.100014] [Medline: 34235291]
27. Schepman A, Rodway P. The General Attitudes Towards Artificial Intelligence Scale (GAAIS): confirmatory validation and associations with personality, corporate distrust, and general trust. Int J Hum–Comput Interact 2023 Aug 9;39(13):2724-2741. [doi: 10.1080/10447318.2022.2085400]
28. Swan BA. Assessing the knowledge and attitudes of registered nurses about artificial intelligence in nursing and health care. NEC 2021;39(3):139 [FREE Full text] [doi: 10.62116/NEC.2021.39.3.139]
29. Kwak Y, Ahn JW, Seo YH. Influence of AI ethics awareness, attitude, anxiety, and self-efficacy on nursing students' behavioral intentions. BMC Nurs 2022 Sep 30;21(1):267. [doi: 10.1186/s12912-022-01048-0]

XSL•FO
RenderX

30. Botwe BO, Antwi WK, Arkoh S, Akudjedu TN. Radiographers' perspectives on the emerging integration of artificial intelligence into diagnostic imaging: the Ghana study. J Med Radiat Sci 2021 Sep;68(3):260-268. [doi: 10.1002/jmrs.460] [Medline: 33586361]
31. Coakley S, Young R, Moore N, et al. Radiographers' knowledge, attitudes and expectations of artificial intelligence in medical imaging. Radiography (Lond) 2022 Nov;28(4):943-948. [doi: 10.1016/j.radi.2022.06.020] [Medline: 35839662]
32. Oh S, Kim JH, Choi SW, Lee HJ, Hong J, Kwon SH. Physician confidence in artificial intelligence: an online mobile survey. J Med Internet Res 2019 Mar 25;21(3):e12422. [doi: 10.2196/12422] [Medline: 30907742]
33. Mehta N, Harish V, Bilimoria K, et al. Knowledge of and attitudes on artificial intelligence in healthcare: a provincial survey study of medical students. MedEdPublish 2021;10(1):75. [doi: 10.15694/mep.2021.000075.1]
34. Abdullah R, Fakieh B. Health care employees' perceptions of the use of artificial intelligence applications: survey study. J Med Internet Res 2020 May 14;22(5):e17620. [doi: 10.2196/17620] [Medline: 32406857]
35. Castagno S, Khalifa M. Perceptions of artificial intelligence among healthcare staff: a qualitative survey study. Front Artif Intell 2020;3:578983. [doi: 10.3389/frai.2020.578983] [Medline: 33733219]
36. Esmaeilzadeh P. Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. BMC Med Inform Decis Mak 2020 Jul 22;20(1):170. [doi: 10.1186/s12911-020-01191-1] [Medline: 32698869]
37. Sample size calculator. Raosoft. 2004. URL: www.raosoft.com/samplesize.html [accessed 2025-12-18]
38. Haviland MG. Yates's correction for continuity and the analysis of 2 x 2 contingency tables. Stat Med 1990 Apr;9(4):363-367. [doi: 10.1002/sim.4780090403] [Medline: 2362976]
39. Sedgwick P. Multiple significance tests: the Bonferroni correction. BMJ 2011 Jan 25;344(jan25 4):e509-e509. [doi: 10.1136/bmj.e509]
40. 2023 Canadian survey of nurses: use of digital health technology in practice: quantitative research report. : Canada Health Infoway; 2024 URL: https://www.infoway-inforoute.ca/en/component/edocman/resources/reports/benefits-evaluation/6481-2023-canadian-survey-of-nurses-use-of-digital-health-technology-in-practice?Itemid=101 [accessed 2025-12-18]
41. Summary corporate plan 2019-2020. Canada Health Infoway. 2019. URL: https://www.infoway-inforoute.ca/en/component/edocman/3654-summary-corporate-plan-2019-2020/view-document [accessed 2025-12-28]
42. Pepito JA, Locsin R. Can nurses remain relevant in a technologically advanced future? Int J Nurs Sci 2019 Jan 10;6(1):106-110. [doi: 10.1016/j.ijnss.2018.09.013] [Medline: 31406875]
43. Scheetz J, Rothschild P, McGuinness M, et al. A survey of clinicians on the use of artificial intelligence in ophthalmology, dermatology, radiology and radiation oncology. Sci Rep 2021 Mar 4;11(1):5193. [doi: 10.1038/s41598-021-84698-5] [Medline: 33664367]
44. Wong K, Gallant F, Szumacher E. Perceptions of Canadian radiation oncologists, radiation physicists, radiation therapists and radiation trainees about the impact of artificial intelligence in radiation oncology - national survey. J Med Imaging Radiat Sci 2021 Mar;52(1):44-48. [doi: 10.1016/j.jmir.2020.11.013] [Medline: 33323332]
45. Patel KM, Metersky K. Reflective practice in nursing: a concept analysis. Int J Nurs Knowl 2022 Jul;33(3):180-187. [doi: 10.1111/2047-3095.12350] [Medline: 34626459]
46. Schiavo G, Businaro S, Zancanaro M. Comprehension, apprehension, and acceptance: understanding the influence of literacy and anxiety on acceptance of artificial Intelligence. Technol Soc 2024 Jun;77:102537. [doi: 10.1016/j.techsoc.2024.102537]
47. Liang HF, Wu KM, Weng CH, Hsieh HW. Nurses' views on the potential use of robots in the pediatric unit. J Pediatr Nurs 2019;47:e58-e64. [doi: 10.1016/j.pedn.2019.04.027] [Medline: 31076190]
48. Mohanasundari SK, Kalpana M, Madhusudhan U, et al. Can artificial intelligence replace the unique nursing role? Cureus 2023 Dec;15(12):e51150. [doi: 10.7759/cureus.51150] [Medline: 38283483]
49. Nursing & compassionate care in the age of artificial intelligence: engaging the emerging future. : Associated Medical Services (AMS); 2020 URL: https://rnao.ca/sites/rnao-ca/files/RNAO-AMS_Report-Nursing_and_Compassionate_Care_in_the_Age_of_AI_Final_For_Media_Release_10.21.2020.pdf [accessed 2025-12-18]
50. Arcega J, Autman I, De Guzman B, et al. The human touch: is modern technology decreasing the value of humanity in patient care? Crit Care Nurs Q 2020;43(3):294-302. [doi: 10.1097/CNQ.0000000000000314] [Medline: 32433070]

Abbreviations:

AI: artificial intelligence
CHN: Community Health Nurse
ML: machine learning
OR: odds ratio
RN: registered nurse
UNBC: University of Northern British Columbia

XSL•FO
RenderX

Edited by E Borycki; submitted 04.Jun.2025; peer-reviewed by A Sowan, L He; revised version received 01.Dec.2025; accepted 06.Dec.2025; published 23.Jan.2026.

Please cite as:
Henderson Betkus M, Banner D, Currie L, Jackson P, Freeman S
Community Health Nurses' Knowledge and Perceptions of AI in Canada: National Cross-Sectional Survey
JMIR Nursing 2026;9:e78560
URL: https://nursing.jmir.org/2026/1/e78560
doi:10.2196/78560

© Mary Henderson Betkus, Davina Banner, Leanne Currie, Piper Jackson, Shannon Freeman. Originally published in JMIR Nursing (https://nursing.jmir.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of Telenursing Guidelines to Improve the Quality of Services in Diabetic Wound Care in a Hospital in Thailand: Case Study

Chonlada Darayon[1], BSCN; Paralee Opasanant[2], PhD; Siriporn Sangsrijan[2], PhD; Waraporn Pattaramungkhunkat[1], BSCN; Panpimol Sukwong[3], PhD

[1]Nursing Department, Mae Chan Hospital, Mae Chan, Chiang Rai, Thailand

[2]School of Nursing, University of Phayao, 19 Moo 2, Tambon Maeka, Amphur Muang, Phayao, Thailand

[3]School of Nursing, Northern College, Tak, Tak, Thailand

Corresponding Author:
Paralee Opasanant, PhD
School of Nursing, University of Phayao, 19 Moo 2, Tambon Maeka, Amphur Muang, Phayao, Thailand

Abstract

Background: The majority of patients with diabetic wounds living in Mae Chan District, Chiang Rai Province face challenges such as a shortage of nurses, limited access to health care, and insufficient resources. Strategies such as specialist networks, patient monitoring, and online care platforms are crucial to improving diabetic wound management in the community.

Objective: This study aims to develop telenursing guidelines for caring for patients with diabetic wounds and foot ulcers, and to investigate the effects of telenursing on wound healing among patients.

Methods: Participatory action research was conducted in three cycles: (1) assessing the current situation and feasibility of telenursing; (2) evaluating telenursing guidelines for wound healing; and (3) examining the effects of telenursing on wound healing, amputation rates, and patient satisfaction.

Results: The mean diabetic wound severity scores decreased after receiving telenursing care at weeks 2, 4, 6, and 8 (P<.001). No patients were found to have foot or leg amputations. The patients in the group who received telenursing care showed that their wounds healed in an average of 8.6 (SD 4.3) weeks. The satisfaction score for telenursing care was 4.7 out of 5 (SD 0.2).

Conclusions: Telenursing guidelines were developed to enhance access to wound care, reduce amputation rates, and promote wound healing, resulting in a significant reduction in wound severity and the absence of amputations. The study further demonstrated that telenursing not only expedited healing times but also reduced health care costs and improved patient satisfaction.

(JMIR Nursing 2026;9:e74228) doi:10.2196/74228

KEYWORDS

telenursing; diabetic ulcers; wound care; nursing guideline; quality of service

Introduction

Diabetes is one of the serious chronic diseases strongly associated with health complications, including macrovascular disease, microvascular disease, peripheral vascular complications, and neuropathy, leading to serious health problems such as coronary heart disease, stroke, kidney disease, and retinal abnormalities. In addition, patients with diabetes are at risk of having foot ulcers, which are the most common complications causing disability, amputation, or premature death. This is one of the most critical public health problems. Diabetic foot ulcers are among the most common complications in persons with diabetes who have problems with uncontrolled blood sugar [1]. The incidence shows that patients with diabetes develop foot ulcers at 19% to 34% and have recurrence symptoms within 3 to 5 years at 65%; moreover, foot or leg amputation is found to be shared among patients with diabetes compared to the general population at 50% to 70%, and they have a chance of premature death within 5 years [2]. Foot ulcers among persons with diabetes take longer to recover than those of normal people, taking an average of 11 to 14 weeks. This could affect the lifestyles of the patients because it causes physical, mental, and social limitations. According to a study of health expenses for caring for patients with foot ulcers in a tertiary university hospital in Singapore between 2013 and 2017, the average cost of wound care alone was US $3368 per patient per year; the cost of minor amputation below the ankle was US $10,468 per patient per year; and the cost of major amputation above or below the knee was US $30,131 per patient per year [3].

Foot ulcers [4,5] can be classified into the following types of wounds. First, infected ulcers are where the wound is infected

XSL•FO
RenderX

and becomes inflamed. The area around the wound is red and spreading from the original site. Symptoms of infected wounds are redness, swelling, pain, and pus. Second, neuropathic ulcers are the most common complication of uncontrolled diabetes. The wound is usually round in shape, with a deep hole in the middle part and thickened skin on the edges. It is found at the base of the big toe and pinky toe. Neuropathic ulcers develop from smaller wounds with symptoms of redness, swelling, numbness, or inflammation around the affected area. Third, ischemic ulcers are wounds that develop when the arteries do not deliver enough blood flow to a specific area, causing local ischemia in the skin and the underlying tissues. Initially, these ulcers manifest as blue discoloration, necrosis, or gangrene. They are painful wounds often found on the tips of the toes, especially the big and pinky toes. The characteristics of the wounds are pale, yellow, and cold, with a weak or absent pulse. They often start at the fingertips and pose an increased risk of amputation if left untreated. Diabetic ulcers are complex and chronic wounds that cannot follow the normal wound-healing process. When the wound is infected, dead tissue (sloughing or shedding) and purulent exudate from inflammation of the wound. There is no new tissue formation, and it does not heal within 4 weeks. Diabetic wounds heal slowly, and sometimes they do not heal. If the wounds are left without proper care, they will cause an infection that will spread to the bone layer (osteomyelitis), requiring organ removal (amputation). If a severe infection enters the bloodstream (septicemia), it can cause death. The essential goals in caring for patients with diabetic ulcers are to prevent disability from the wound spreading to the point where it is necessary to amputate the involved organ, to prevent new wounds from forming after treatment, and to increase the quality of life of the patients.

Diabetic wound care is integral to medical practice to heal wounds better and faster. To promote wound healing faster, (1) clean the wound to prevent infection and promote the wound healing process, (2) keep the wound elevated to reduce fluid accumulation and promote blood circulation, (3) provide adequate knowledge regarding nutrition to promote the wound healing process, (4) have enough rest because resting will reduce the metabolic process inside unnecessary cells and tissues and can carry enough oxygen and nutrients to the wounds, and (5) promote personal hygiene [6]. Diabetic wound care should be supervised by a nurse with clinical expertise who can collect, evaluate, analyze, and identify the needs and problems of patients. Expert nurses should be able to manage options for care and use medical supplies and equipment correctly and appropriately, obtain communication skills to coordinate with the treatment team and encourage family members to participate in care, identify health services and resources, protect patients' rights and privacy, and preserve patient safety and prevent harm during care [7-9]

Despite a community point prevalence of approximately 1.5% to 5% for diabetic foot ulcers, the lifetime incidence risk reaches 19% to 34% [1]. This elevation is frequently linked to suboptimal foot self-care practices; however, there are not enough nurses to meet the increasing demand for care and other issues in Mae Chan District, Chiang Rai Province, such as a lack of health equipment; resources; and access to health services, communication channels for health care advice, and continuity of care rehabilitation. Therefore, strategies for caring for diabetic wounds are necessary, such as creating a network of channels for giving advice from specialists to nurses, caregivers, and patients in the community, monitoring and following up with patients, and providing quality of service through online channels. Various health care advice channels could increase access to services, reduce waiting time, and facilitate patients in a new way of life (New Normal). Moreover, the use of digital technology systems to provide health services for continuous care and referral to primary care units with the best practice guidelines, The Nursing and Midwifery Council has announced telenursing strategies to provide health care advice and to find solutions related to health within the knowledge framework of the nursing and midwifery profession [10].

Telenursing uses information and communication technologies to develop all fields of nursing care, education, and research at a distance [11]. In Thailand, the Nursing and Midwifery Council provides advice and solutions related to health within the knowledge framework by using this digital technology system to provide services and support all forms of nursing care, including continuous care, quality of care, equal access to health, and allowing people to manage their health through online services [10]. Telenursing care increases channel access to health services, covers health care in remote areas, especially in areas with travel restrictions, and maintains professional standards and quality. Patients also do not need to commute to the hospital, which helps reduce the cost of hospital admission and other related expenses while still allowing monitoring and following up on patients' conditions through online channels.

The Mae Chan Hospital is located in the Chiang Rai Province. It is a middle-level hospital that can host up to 120 beds and support referrals from community hospitals in the joint service network of 4 districts: Mae Chan District, Mae Fah Luang District, Doi Luang District, and Chiang Saen District. The Mae Chan Hospital is responsible for caring for patients referred from community hospitals when the capacity of those hospitals in the 4 districts reaches the limit. The hospital helps increase access to treatment for service users in the area, reduces referrals to regional hospitals (advanced-level hospital), and supports primary service units in each district. The Wound and Ostomy Care Clinic, Mae Chan Hospital, provides services for patients with chronic wounds, bedsores, general wounds, penetrating wounds, drains, and ostomy care. Some nurses specialize in diabetic wounds, one specialist in ostomy care, and one in general care for patients. There were a number of people who received services from 2020 to 2022, including a total of 2142, 1571, and 1675 patients, representing 3725, 4931, and 6239 visits, respectively. There were 114, 70, and 67 patients with diabetic ulcers, accounting for 405, 281, and 291 visits, respectively. The average cost of commuting to the hospital for patients to receive wound care is US $10 per time. The average number of days to receive wound care services is 14 to 45 days per person. In general, the travel distance for patients to visit the wound care and ostomy clinic is 1 to 40 km. Due to travel distance and expenses involved in wound care at the hospital, some of the patients are unable to receive treatment regularly

XSL•FO
RenderX

and continuously, so many patients experience wound spreading and recurrence of wound infection, leading to foot or leg amputation in some cases.

According to literature reviews, telenursing is important in assisting patients to access standard care for chronic wounds. Telenursing care helps reduce complications and the number of patients admitted to the hospital, which encourages participation among patients, caregivers, and nurses in self-care practices, and increases patient satisfaction because it reduces travel time, traveling costs, and other related expenses of going to the hospital. Therefore, the research team was interested in studying the development of telenursing guidelines for caring for patients with diabetic wounds and foot ulcers, and the effects of telenursing on wound healing in patients to provide quality service and prevent the spread and recurrence of wound infections and foot and leg amputations in patients.

Methods

Study Design

This was participatory action research [12], which divided the study into 3 cycles.

Cycle 1

The current situations, problems, and needs of patients with diabetic wounds and the feasibility of using telenursing guidelines in the context of the Mae Chan Hospital, Chiang Rai Province, and community networks in the research area were studied. The researcher used lessons learned from the study to conclude the issues and needs of developing and drafting telenursing guidelines for diabetic wound care. Informants and participants in the design of diabetic wound care guidelines and telenursing system were purposively selected, consisting of (1) patients with diabetic wounds and/or their caregivers who received care in the Wound and Ostomy Care Clinic, totaling 10 persons; (2) the multidisciplinary team including a surgeon, a pharmacist, a nutritionist, and 9 registered nurses at subdistrict health-promoting hospital in the Mae Chan Hospital networks, totaling 12 people; and (3) a computer technical officer.

The interprofessional team plays a critical role in providing comprehensive care for patients with chronic diabetic wounds. The surgeon is central to diagnosing the underlying causes of chronic wounds, determining their etiology, and formulating an integrated treatment plan. In collaboration with the health care team, the surgeon ensures the effective implementation of telenursing guidelines, addressing the needs of patients who face barriers to in-person health care access. The pharmacist works closely with the physician and nursing staff to select appropriate topical treatments that promote wound healing and minimize the risk of infection. In addition to this, the pharmacist is responsible for educating both patients and their families on the proper use of medications and wound care techniques. The nutritionist plays a vital role by recommending dietary strategies that support wound healing, promote tissue regeneration, and address underlying risk factors such as poor circulation or diabetes. Their tailored nutritional guidance is designed to strengthen the immune system and accelerate recovery. Finally, the computer technical officer ensures the seamless operation of the telemedicine system, facilitating smooth communication via platforms like the "Wound Care Mae Chan" LINE app. They provide technical support, enabling real-time updates and enhancing the overall telehealth experience for both health care providers and patients.

Cycle 2

Telenursing guidelines toward diabetic wound healing were studied. Participants recruited in the study were patients with diabetic wounds and professional nurses from subdistrict health-promoting hospitals under the Mae Chan Hospital network, totaling 20 persons. A sample size calculation was done using G*Power software version 3.1.9.4 (Fual et al), with test power set at 0.70, confidence value at 0.05, and influence value (effect size) at 0.30. The number of samples obtained from the calculation was 13. An additional 50% was added to the sample size, so the total number of participants in this cycle was 20.

Cycle 3

The effects of telenursing on diabetic wound healing were studied by evaluating the wounds, foot or leg amputation, and satisfaction with using telenursing guidelines. A quasi-experimental one-group pretest-posttest design was employed in this stage of the study.

Research Instruments

The tools used for data collection are given below:

1. General information on characteristics was collected, including gender, age, education level, occupation, history of smoking, history of diabetes and duration, and types of diabetic wounds.
2. Bates-Jensen Wound Assessment Tool (BWAT): the BWAT consists of 13 wound characteristics: size, visible depth, wound edges, undermining and tunneling processes, necrotic tissue type and amount, exudate type and amount, surrounding skin discoloration, peripheral tissue edema, peripheral tissue induration, granulation tissue, and epithelialization [13]. The scores range from 13 to 65, and content validity and concurrent validity are both equal to 0.91. The assessment tool was used in a tryout among 10 patients with similar characteristics to find reliability using the Cronbach α coefficient formula, and the reliability value was 0.88. The cutoff scores of BWAT are the following: minimal scores are 13 to 20, mild severity scores are 21 to 30, moderate severity scores are 31 to 40, and critical severity scores are 41 to 65.
3. The questionnaire on satisfaction with the telenursing guidelines consists of 7 items rated on a 5-point Likert scale measuring satisfaction, with response options ranking from 5 (very satisfied), 4 (satisfied), 3 (neutral), 2 (dissatisfied), and 1 (very dissatisfied).

The tools used to conduct the study included the following:

1. Guidelines for diabetic wound care: the researcher developed the guidelines based on the existing literature review by specifying keywords to search for papers and related documents, including diabetic wounds, chronic wounds, guidelines for wound care, and wound care. The

XSL•FO
RenderX

literature review search period was set from 2014 to 2023. Information was also obtained from cycle 1, which involved studying the current situations, problems, and needs of patients with diabetic wounds and the feasibility of using telenursing guidelines. The researcher presented a draft of telenursing guidelines to 3 experts in the fields of surgery, preventive medicine, and wound and ostomy care. These experts reviewed the guidelines to ensure they were consistent with the research objectives. Content validity was performed by measuring the Index of Item Objective Congruence, which was equal to 0.80.

2. LINE app "Wound Care Mae Chan"
3. Telemedicine Service System (Telemedicine; ITELE) of Mae Chan Hospital

Ethical Considerations

The Human Research Ethics Committee at Chiang Rai Provincial Public Health Office approved this research study under approval 37/2023, with certification granted on April 1, 2023. The researcher adhered strictly to ethical principles throughout the study. Before participation, the researcher clearly communicated the objectives of the study, the research process, the expected duration, and the anticipated benefits to ensure a full understanding; subsequently, written informed consent was obtained from all participants before any data collection commenced. Informants and participants were informed that communication regarding wound care would be conducted via the LINE application, specifically designated for this research group. They were also informed of their right to withdraw from the study at any time, without any impact on their current treatment or benefits received from the hospital.

Participants received a remuneration of US $6 for each completed data collection session. The researchers were committed to protecting the participants' rights and privacy. As such, all data were recorded using codes instead of real names, ensuring anonymity. All participant information was kept confidential and could not be traced back to any individual. The findings of the study were presented as aggregated data to protect participant identity. Furthermore, all recorded data, including images, would be destroyed 1 year after the publication of the research results.

Results

As per the study objectives, we developed telenursing guidelines for caring for patients with diabetic wounds and foot ulcers, and studied the effects of telenursing on wound healing, using participatory action research.

In cycle 1, to study the current situations, problems, and needs of patients with diabetic wounds and the feasibility of using telenursing guidelines, we recruited 10 informants who were patients with diabetic wounds and foot ulcers aged between 30 and 74 years (average age 62.6, SD 12.4 years; duration of diabetes 1-10 years; average duration of diabetes 4.9, SD 3.5 years; average wound assessment score 51.3, SD 3.3); additional informants included a surgeon, a pharmacist, a nutritionist, 9 registered nurses at the subdistrict health-promoting hospital, and a computer technical officer. The following needs and problems were found.

Some patients could not receive wound care regularly and continuously because of the distance to the hospital, traveling difficulties, expenses for transportation, food, medical care, and related costs. Findings showed that some patients live far from the hospital. It costs them a lot of money per hospital visit. In some cases, patients have no one to take them to the hospital because family members have to work, and if they have to accompany the patients to the hospital, they have to miss their work and lose income. Moreover, they must spend extra money on transportation and food while waiting for patients to receive wound care. The cost of traveling to the hospital to receive wound care services is approximately 100 to 1000 Baht (US $2.85-$28.50) per visit, which is too much for them to spend, especially for patients and their family members who live in a remote community area.

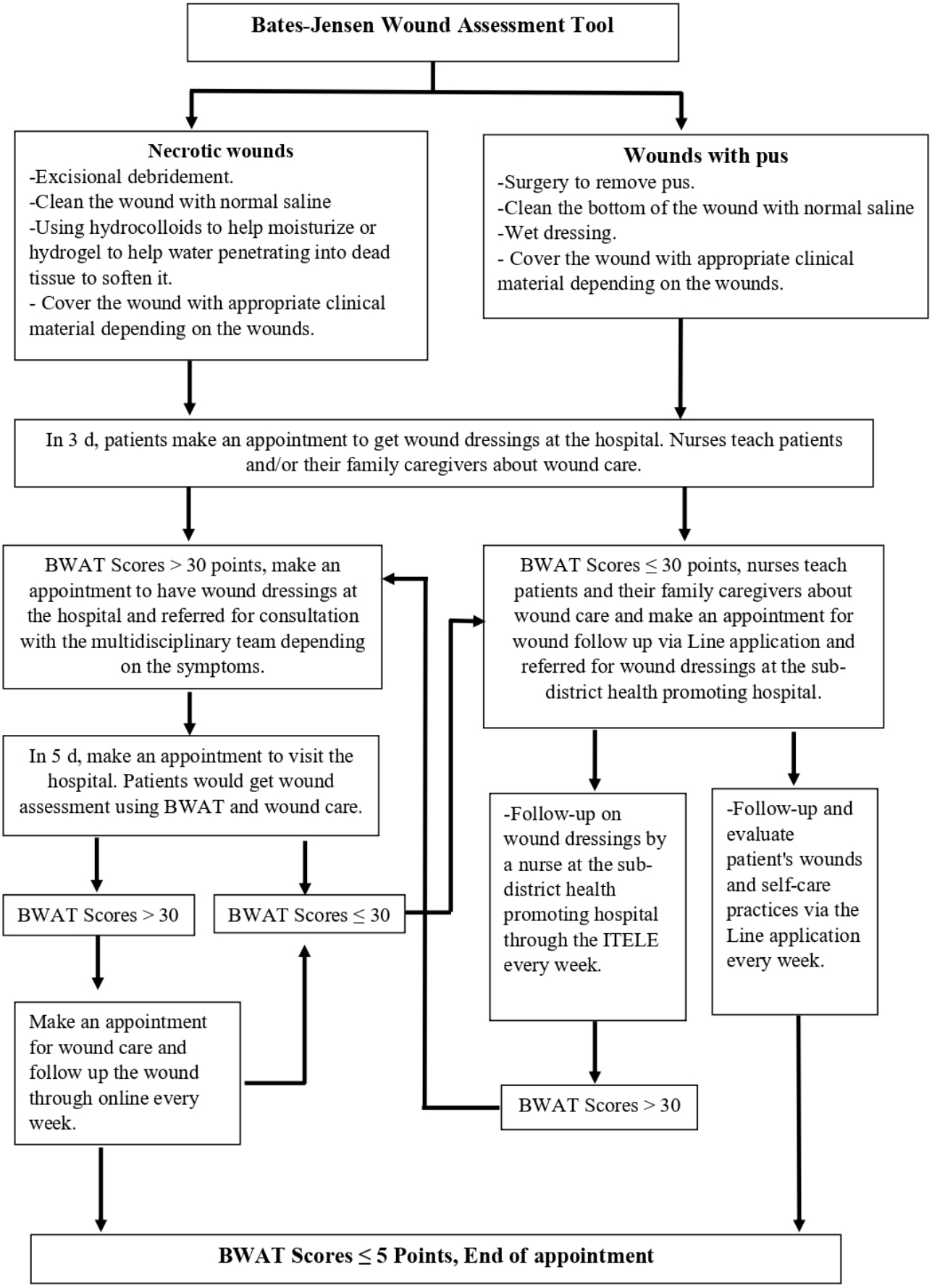
Lack of knowledge about diabetic wound care is one of the problems. Patients lack adequate knowledge of wound care, nutrition, and control of blood sugar levels. These are factors for wound healing. In some cases, patients did not follow up on the appointment, which could also affect their physical health. Because of this, many patients experienced wound spreading and recurrence of wound infection, and in some cases, may experience foot or leg amputation.

Treating severely infected wounds requires skills and expertise in advanced wound care, which exceeds the potential of community nurses working in subdistrict health-promoting hospitals, and patients did not have access to or resources for advanced wound care, such as hydrocolloid, hydro fiber, alginate, hydrogel, foam, and film dressings.

Nowadays, the majority of patients and their family caregivers have smartphones that can be used to communicate via the LINE app. Introducing telenursing for diabetic wound care among patients could help reduce barriers to receiving health care services. Based on the study of situations, needs, and problems among patients with diabetic wounds, telenursing guidelines for diabetic wound care are summarized as shown in Figure 1.

XSL•FO
RenderX

Figure 1. Telenursing guidelines for diabetic wound care. BWAT: Bates-Jensen Wound Assessment Tool; NSS: normal saline solution.



In cycle 2, the telenursing guidelines for diabetic wound healing were evaluated.

The study sample included 20 patients with diabetic wounds and foot ulcers, and half of them were female (n=10, 50%) and

XSL•FO
RenderX

aged 30 to 81 years, with an average age of 47.6 (SD 9.4) years and an average duration of diabetes of 7.6 (SD 6.4) years. Studies showed that the wounds were found most on the leg, foot, and toe (n=15, 75%). The average assessment wound severity score was 50.8 (SD 4.4), indicating that patients had wound critical severity (BWAT=41‐65 points represents critical severity), and the results are shown in Table 1.

XSL•FO
RenderX

Table . Attributes of the sample.

Items	Value
Sex, n (%)	
Male	10 (50)
Female	10 (50)
Age (y)	
30 - 49, n (%)	2 (10)
50 - 59, n (%)	3 (15)
60 - 69, n (%)	9 (45)
≥70, n (%)	6 (30)
Mean (SD)	62.6 (12.4)
Duration of diabetes (y), n (%)	
0 - 2	4 (20)
3 - 5	6 (30)
6 - 10	7 (35)
≥10	3 (15)
Wound location, n (%)	
Head and neck	1 (5)
Arm, hand, and fingers	4 (20)
Legs, feet, and toes	15 (75)
Pain, n (%)	
Pain occurs during wound dressing	7 (35)
Pain is present at all times	8 (40)
Intermittent pain	5 (25)
Wound status score	
Mean (SD)	50.8 (4.35)
Critical severity (BWAT[a]=41 - 65 points), n (%)	20 (100)
Wound characteristics, n (%)	
Size (length × width) (cm^2)	
<4	3 (15)
4-16	11 (55)
16.1-36	2 (10)
36.1-80	3 (15)
>80	1 (5)
Depth	
Obscured by necrosis	5 (25)
Full-thickness skin loss with extensive destruction, tissue necrosis, or damage to muscle, bone, or supporting structures	15 (75)
Wound edges	
Indistinct, diffuse, and none clearly visible	19 (95)
Distinct, outline clearly visible, attached, and even with wound base	1 (5)
Undermining	
<2 cm in any area	14 (70)
2 - 4 cm involving <50% wound margins	6 (30)

XSL•FO
RenderX

Items	Value
Necrotic tissue type	
Loosely adherent yellow slough	3 (15)
Adherent, soft, and black eschar	17 (85)
Necrotic tissue amount	
>50% and <75% of wound covered	14 (70)
75%-100% of wound covered	6 (30)
Exudate type	
Serosanguineous: thin, watery, and pale red or pink	1 (5)
Serous: thin, watery, and clear	7 (35)
Purulent: thin or thick, opaque, tan or yellow, and with or without odor	12 (60)
Exudate amount	
The amount of secretion is ≤25%	1 (5)
The amount of secretion is more than 25% - 75%	7 (35)
The amount of secretion is more than 75%	12 (60)
Skin color surrounding the wound	
Bright red and/or blanches to touch	16 (80)
Black or hyperpigmented	4 (20)
Peripheral tissue edema	
Pitting edema extends <4 cm around the wound	9 (45)
Crepitus and/or pitting edema extends >4 cm around the wound	11 (55)
Peripheral tissue induration	
None present	1 (5)
Induration, <2 cm around the wound	2 (10)
Induration 2 - 4 cm extending <50% around the wound	6 (30)
Induration 2 - 4 cm extending >50% around the wound	11 (55)
Granulation tissue, n (%)	
Bright, beefy red; <75% and >25% of wound filled	3 (15)
Pink and/or dull, dusky red and/or fills <25% of wound	17 (85)
Epithelialization, n (%)	
50% to <50% wound covered	3 (15)
<25% wound covered	17 (85)

[a]BWAT: Bates-Jensen Wound Assessment Tool.

The researcher prepared workshops and set up a meeting to explain the task of developing telenursing guidelines to the committee and those responsible for the continuous care, advanced wound care, and training program using the ITELE.

In cycle 3, mean wound severity scores were compared before and after receiving telenursing guidelines for diabetic wound care at week 2, week 4, week 6, and week 8 using 1-way repeated measures (ANOVA) statistical analysis; the results are shown in Table 2. The mean scores for wound severity were significantly different at the .001 level. A comparison of the wound severity scores before and after receiving telenursing nursing care guidelines for wound care at week 2, week 4, week 6, and week 8 using *t* test (1-tailed) statistical analysis is shown in Table 3.

XSL•FO
RenderX

Table . Comparison of the wound severity scores before and after receiving telenursing guidelines toward diabetic wound care using repeated measures of 1-way ANOVA statistics.

ANOVA	Sum of square	*df*	Mean square	*F* test	*df*	*P* value
Between-group	15,379.4	4	3844.8	401.7	4	<.001
Within-group	3360.2	95	35.4	—[a]	95	—
Total	18,739.6	99	—	—	99	—

[a]Not applicable.

Table . Comparison of the wound severity scores before and after receiving telenursing guidelines towards diabetic wound care at week 2, week 4, week 6, and week 8.

	Wound severity scores, mean (SD)	*t* test (*df*)	*P* value
Before and after receiving telenursing care guidelines for diabetic wound care in the second week	23.7 (5.3)	20.0 (19)	<.001
Before and after receiving telenursing care guidelines for diabetic wound care in the fourth week	29.6 (3.9)	34.3 (19)	<.001
Before and after receiving telenursing care guidelines for diabetic wound care in the sixth week	32.0 (4.0)	35.3 (19)	<.001
Before and after receiving telenursing care guidelines for diabetic wound care in the eighth week	34.0 (4.7)	36.1 (19)	<.001
After receiving telenursing care guidelines for diabetic wound care in the week 2 and week 4	5.9 (3.9)	6.8 (19)	<.001
After receiving telenursing care guidelines for diabetic wound care in the week 2 and week 6	8.3 (5.0)	7.4 (19)	<.001
After receiving telenursing care guidelines for diabetic wound care in the week 2 and week 8	10.3 (6.5)	7.1 (19)	<.001
After receiving telenursing care guidelines for diabetic wound care in the week 4 and week 6	2.4 (2.6)	4.1 (19)	<.010
After receiving telenursing care guidelines for diabetic wound care in the week 4 and week 8	4.4 (3.9)	5.1 (19)	<.001
After receiving telenursing care guidelines for diabetic wound care in the week 6 and week 8	2.0 (2.3)	3.8 (19)	<.010

According to Table 3, the wound severity scores after receiving telenursing care in the 2nd, 4th, 6th, and 8th weeks had decreased statistically (*P*<.001). The scores after receiving telenursing care in weeks 4, 6, and 8 had decreased statistically from week 2 (*P*<.001). The scores after receiving telenursing care in weeks 6 and 8 had decreased statistically from week 4 (*P*<.001), and the scores after receiving telenursing care in week 8 had decreased statistically from week 6 (*P*<.001).

Discussion

Principal Findings

The findings of this study showed that samples had a mean age of 62.6 (SD 12.4) years, an average duration of having diabetes for 4.9 (SD 3.5) years, and high levels of wound severity on average (BWAT=50.8, SD 4.3). This indicates that the majority of the sample were older people, who have a high risk for critical wound severity. First, the factors affecting the wound healing process include infection, high blood sugar levels, and nervous system disorders. This is consistent with the study by Burgess et al [14], which found that the pathophysiology related to the diabetic wound healing process includes hyperglycemia, neuropathy, microvascular complications, infection, inflammation, and immune system deficiency in chronic wounds, and the psychological impacts of diabetes mellitus. Second, the factors preventing patients from receiving wound care continuously include financial limitations, distance to the hospital, no caregivers taking them to the hospital, and so on.

XSL•FO
RenderX

From the in-depth interviews, it was found that the majority of patients live far from the hospital. Distance is one of the barriers to accessing health services. Some patients do not have anyone to care for because family members have to work. Due to this reason, some of the patients are unable to go to the hospital by themselves. If a caregiver comes to the hospital, it would cost more money for traveling expenses, which is more than the net income received each day. The average cost is US $6 to US $10 per time, which is considered a high expense. Third, the knowledge factors where the patient does not have enough knowledge regarding blood sugar control and nutrition, which are important factors to promote wound healing. The problems and needs of patients led to the development of telenursing guidelines to provide patients with access to wound care services to prevent leg or foot amputation, increase the rate of wound healing, reduce wound healing time, reduce inequality, and distribute opportunities to access advanced wound care and resources for wound healing.

A study of the development of using telenursing guidelines on diabetic wound healing found that the mean wound-based severity scores had decreased significantly after receiving telenursing guidelines for diabetic wounds in week 2, week 4, week 6, and week 8 ($P<.001$). Moreover, no patients were found to have foot or leg amputation. This can be explained by the fact that the samples in this study received baseline assessment, including illness history, duration of the wound, history of received medicine, financial status, nutrition, and wound assessment using the BWAT. Patients also received wound assessment every week from the beginning throughout the study. Having regular wound assessments is important for effective wound care. Understanding factors affecting wound healing helps nurses make appropriate decisions in wound care [15]. This allows the wound healing process to proceed normally, and it works well together by preparing the wound to be in an appropriate state according to the principles of wound bed preparation [16]. Wound management consists of three phases. First, the debridement of the wound is when dead or unhealthy tissue is removed from a wound. It stimulates wound healing to the proliferative phase, causing re-epithelialization of the wound to be smaller, which prevents or reduces the amount of infection by removing pathogenic parts such as foreign bodies. Second, the wounds are made moist by dealing with secretions, making the wound healing more effective and causing angiogenesis, the formation of granulation tissue, and the process of re-epithelialization. Third, the wound infection is controlled by cleaning the wound to reduce the number of germs and giving topical antibiotics [17,18].

Information for developing telenursing guidelines was obtained from studying the situations, problems, and needs of patients with diabetic wounds and foot ulcers, as well as from reviewing the existing literature. Based on the information, self-care behaviors are important for promoting wound healing among patients, such as controlling blood sugar levels, taking care of the wound by cleaning and getting wound care regularly and continuously, and eating proper food with good nutrition. According to this study, the patients in the group that received telenursing care could heal wounds in an average of 8.6 (SD 4.3) weeks. The patients in the group that received advanced wound care and normal wound care could heal their wounds in an average of 10.2 weeks. The patients in the group that received wound care every day (conventional wound dressings) and normal wound care could heal in an average of 11.72 weeks. The average cost of wound care for the group receiving telenursing guidelines was 11,772 Baht (US $336.34) per person, receiving advanced wound care and normal care was 14,772 Baht (US $422.06) per person, and receiving wound care and normal care every day was 22,740 Baht (US $649.71) Baht per person. In terms of the average score of satisfaction using telenursing guidelines, it was 4.7 out of 5 (SD 0.2). This is consistent with many studies stating that self-care behaviors, blood sugar level control, and continuous use of medication among patients with diabetes were better after telenursing care [19-22], and this reduces the use of resources for treating diabetic wounds, hospital admissions, and patient expenses while maintaining service quality [19,23-26].

This research is in alignment with both national and international policies that promote the incorporation of digital technology as a tool for enhancing health care services. It also functions as a conduit, bridging the gap between communities and health care systems, thereby mitigating geographical and relational disparities between the public and health care providers. Moreover, digital technology empowers individuals to stay informed about evolving health conditions, understand their rights to access high-quality, equitable, and safe health care, and fosters public awareness and collaboration in improving personal and community health outcomes [27,28].

Suggestions for Using Research Results

To continuously produce good results when implementing telenursing guidelines for improving the quality of services in diabetic wound care, those involved in the development process, including community people, patients, health care providers, nurses, health staff, and family caregivers, must be prepared in terms of knowledge in the field of diabetic wound care such as advanced wound care, the ITELE program, community nursing, self-care practices and knowledge regarding the use of remote technology in accordance with the steps of the nursing process, and health assessment, including defining problems, planning, practice, and evaluation.

There must be supervision, monitoring, and evaluation of the use of telenursing guidelines to improve nursing practices suitable to the community context and in accordance with knowledge about advanced wound care, diabetic wounds, chronic wounds, and remote technology.

Limitations

The small sample size limits the ability to generalize the findings to a broader population.

XSL•FO
RenderX

Acknowledgments

The authors would like to express their sincere gratitude to the patient with diabetes who provided useful information.

Funding

The authors acknowledge Mae Chan Hospital's financial support.

Conflicts of Interest

None declared.

References

1. Armstrong DG, Boulton AJM, Bus SA. Diabetic foot ulcers and their recurrence. N Engl J Med 2017 Jun 15;376(24):2367-2375. [doi: 10.1056/NEJMra1615439] [Medline: 28614678]
2. McDermott K, Fang M, Boulton AJM, Selvin E, Hicks CW. Etiology, epidemiology, and disparities in the burden of diabetic foot ulcers. Diabetes Care 2023 Jan 1;46(1):209-221. [doi: 10.2337/dci22-0043] [Medline: 36548709]
3. Lo ZJ, Surendra NK, Saxena A, Car J. Clinical and economic burden of diabetic foot ulcers: a 5-year longitudinal multi-ethnic cohort study from the tropics. Int Wound J 2021 Jun;18(3):375-386. [doi: 10.1111/iwj.13540] [Medline: 33497545]
4. Lim CS, Baruah M, Bahia SS. Diagnosis and management of venous leg ulcers. BMJ 2018 Aug 14;362:k3115. [doi: 10.1136/bmj.k3115] [Medline: 30108047]
5. Mayrovitz HN, Wong S, Mancuso C. Venous, arterial, and neuropathic leg ulcers with emphasis on the geriatric population. Cureus 2023 Apr;15(4):e38123. [doi: 10.7759/cureus.38123] [Medline: 37252574]
6. Seenil K, Saniloh K, Meechana J, et al. Advanced wound care. Vajirasarn Nurs J 2020;22(1):104-115 [FREE Full text]
7. Institute of Medical Research and Technology Assessment. Clinical Practice Guideline: Prevention and Management of Diabetic Foot Complications: The Agriculture Co-Operative Federation of Thailand; 2013. URL: https://diabeticfootthailand.com/wp-content/uploads/2021/08/DM-foot-2556.pdf [accessed 2026-01-03]
8. Sindhu S. Wound and Wound Care: Thailand Nursing and Midwifery Council; 2022. URL: https://uctal.thailis.or.th/bibitem?bibid=b03252322 [accessed 2026-01-03]
9. International best practice guidelines: wound management in diabetic foot ulcers. : Wounds International; 2013 URL: https://legsmatter.org/wp-content/uploads/2018/04/dfubestpracticeforweb.pdf
10. Thailand Nursing and Midwifery. Tele-nursing. : Royal Gazette; 2021 URL: https://www.tnmc.or.th/images/userfiles/files/T_0049.PDF [accessed 2026-01-03]
11. Souza-Junior VD, Mendes IAC, Mazzo A, Godoy S. Application of telenursing in nursing practice: an integrative literature review. Appl Nurs Res 2016 Feb;29:254-260. [doi: 10.1016/j.apnr.2015.05.005] [Medline: 26856523]
12. Kemmis S, McTaggart R. Participatory action research: communicative action and the public sphere. In: Denzin N, Lincoln Y, editors. Strategies of Qualitative Inquiry: SAGE Publications, Inc; 2007:271-330 URL: https://books.google.co.in/books?id=hD7W095U66cC&pg=PR5&source=gbs_selected_pages&cad=1#v=onepage&q&f=false [accessed 2026-01-03]
13. Bates-Jensen BM, McNees P. Toward an intelligent wound assessment system. Ostomy Wound Manage 1995 Aug;41(7A Suppl):80S-86S. [Medline: 7669204]
14. Burgess JL, Wyant WA, Abdo Abujamra B, Kirsner RS, Jozic I. Diabetic wound-healing science. Medicina (Kaunas) 2021 Oct 8;57(10):1072. [doi: 10.3390/medicina57101072] [Medline: 34684109]
15. Houghton P. Research 101: wound assessment tools. Wound Care Can 2018;16(1):58-64 [FREE Full text]
16. Falanga V. Wound healing and its impairment in the diabetic foot. Lancet 2005 Nov 12;366(9498):1736-1743. [doi: 10.1016/S0140-6736(05)67700-8] [Medline: 16291068]
17. Halim AS, Khoo TL, Saad AZM. Wound bed preparation from a clinical perspective. Indian J Plast Surg 2012 May;45(2):193-202. [doi: 10.4103/0970-0358.101277] [Medline: 23162216]
18. desJardins-Park HE, Gurtner GC, Wan DC, Longaker MT. From chronic wounds to scarring: the growing health care burden of under- and over-healing wounds. Adv Wound Care (New Rochelle) 2022 Sep;11(9):496-510. [doi: 10.1089/wound.2021.0039] [Medline: 34521257]
19. Ghoulami-Shilsari F, Esmaeilpour Bandboni M. Tele-nursing in chronic disease care: a systematic review. Jundishapur J Chronic Dis Care 2019;8(2):e84379. [doi: 10.5812/jjcdc.84379]
20. Moriyama M, Kazawa K, Jahan Y, et al. The effectiveness of telenursing for self-management education on cardiometabolic conditions: a pilot project on a remote Island of Ōsakikamijima, Japan. J Prim Care Community Health 2021;12:21501327211030817. [doi: 10.1177/21501327211030817] [Medline: 34238053]
21. AkbariRad M, Dehghani M, Sadeghi M, et al. The effect of telenursing on disease outcomes in people with type 2 diabetes mellitus: a narrative review. J Diabetes Res 2023;2023:4729430. [doi: 10.1155/2023/4729430] [Medline: 38098964]
22. Sriprasert K, Aungwattana S, Tamdee D. Effect of tele-nursing-based self-management program on health behaviors of persons with uncontrolled diabetes mellitus in the community [Master's Thesis]. : Chiang Mai University; 2023 URL: https://cmudc.library.cmu.ac.th/frontend/Info/item/dc:170970 [accessed 2026-01-19]

XSL•FO
RenderX

23. Nittari G, Ricci G, Savva D, et al. Telemedicine in diabetic ulcer management: a pilot study with exploration of medico-legal aspects. Nutr Metab Cardiovasc Dis 2023 Nov;33(11):2280-2286. [doi: 10.1016/j.numecd.2023.07.021] [Medline: 37591741]
24. Chen L, Cheng L, Gao W, Chen D, Wang C, Ran X. Telemedicine in chronic wound management: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Jun 25;8(6):e15574. [doi: 10.2196/15574] [Medline: 32584259]
25. Bai X, Zhang H, Jiao Y, Yuan C, Ma Y, Han L. Digital health interventions for chronic wound management: a systematic review and meta-analysis. J Med Internet Res 2024 Jul 16;26:e47904. [doi: 10.2196/47904] [Medline: 39012684]
26. Zhang X, Guo Z, Duan Y, Sun C, Luo J. Effectiveness of telemedicine on wound-related and patient-reported outcomes in patients with chronic wounds: systematic review and meta-analysis. JMIR Mhealth Uhealth 2025 Jun 10;13:e58553. [doi: 10.2196/58553] [Medline: 40493508]
27. Information and Communication Technology Center. Digital health strategy, ministry of public health (2021-2025). : Ministry of Public Health URL: https://surl.lt/tlxigx [accessed 2026-01-01]
28. Global strategy on digital health 2020-2027. World Health Organization. 2025. URL: https://iris.who.int/server/api/core/bitstreams/a64d584f-015b-471f-9d86-9ea263cfb516/content [accessed 2026-01-03]

Abbreviations

BWAT: Bates-Jensen Wound Assessment Tool

Edited by E Borycki; submitted 20.Mar.2025; peer-reviewed by A Prechaterasat, B Chaithat, SC Barros; revised version received 27.Dec.2025; accepted 28.Dec.2025; published 09.Feb.2026.

Please cite as:
Darayon C, Opasanant P, Sangsrijan S, Pattaramungkhunkat W, Sukwong P
Development of Telenursing Guidelines to Improve the Quality of Services in Diabetic Wound Care in a Hospital in Thailand: Case Study
JMIR Nursing 2026;9:e74228
URL: https://nursing.jmir.org/2026/1/e74228
doi:10.2196/74228

© Chonlada Darayon, Paralee Opasanant, Siriporn Sangsrijan, Waraporn Pattaramungkhunkat, Panpimol Sukwong. Originally published in JMIR Nursing (https://nursing.jmir.org), 9.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Retraction: Comparative Effectiveness of Health Communication Strategies in Nursing: A Mixed Methods Study of Internet, mHealth, and Social Media Versus Traditional Methods

JMIR Publications Editorial Office

JMIR Publications, 130 Queens Quay E, Suite 1100, Toronto, ON, Canada

Corresponding Author:
JMIR Publications, 130 Queens Quay E, Suite 1100, Toronto, ON, Canada

Related Article:

Retraction of: https://nursing.jmir.org/2024/1/e55744/

(JMIR Nursing 2026;9:e89640) doi:10.2196/89640

The JMIR Publications Editorial Office is retracting the article "Comparative Effectiveness of Health Communication Strategies in Nursing: A Mixed Methods Study of Internet, mHealth, and Social Media Versus Traditional Methods" by Hamarash et al [1]. This follows an investigation that identified concerns regarding the integrity of the peer-review process that occurred for this article. We regret that these issues were not identified prior to publication.

All authors did not agree with retraction.

Reference

1. Hamarash MQ, Ibrahim R, Yaas MH, Abdulghani MF, Al Mushhadany O. Comparative effectiveness of health communication strategies in nursing: a mixed methods study of internet, mHealth, and social media versus traditional methods. JMIR Nurs 2024 Nov 19;7:e55744. [doi: 10.2196/55744] [Medline: 39622705]

Submitted 15.Dec.2025; this is a non–peer-reviewed article;accepted 15.Dec.2025; published 04.Feb.2026.

Please cite as:
JMIR Publications Editorial Office
Retraction: Comparative Effectiveness of Health Communication Strategies in Nursing: A Mixed Methods Study of Internet, mHealth, and Social Media Versus Traditional Methods
JMIR Nursing 2026;9:e89640
URL: https://nursing.jmir.org/2026/1/e89640
doi:10.2196/89640

© JMIR Publications Editorial Office. Originally published in JMIR Nursing (https://nursing.jmir.org), 4.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX