JMIR Nursing

Virtualizing care from hospital to community: Mobile health, telehealth, and digital patient care
Volume 8 (2025)   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

Original Papers

Review

Research Letter

Original Paper

Exploring Educators' Perceptions and Experiences of Online Teaching to Foster Caring Profession Students' Development of Virtual Caring Skills: Sequential Explanatory Mixed Methods Study

Lorelli Nowell[1], PhD; Sonja Johnston[2], PhD; Sara Dolan[1], PhD; Michele Jacobsen[2], PhD; Diane L Lorenzetti[3], PhD; Elizabeth Oddone Paolucci[3], PhD

[1]Faculty of Nursing, University of Calgary, Calgary, AB, Canada
[2]Werklund School of Education, University of Calgary, Calgary, AB, Canada
[3]Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Corresponding Author:
Lorelli Nowell, PhD
Faculty of Nursing
University of Calgary
2500 University Dr NW
Calgary, AB, T2N 1N4
Canada
Phone: 1 4036209822
Email: lnowell@ucalgary.ca

Abstract

Background: Professionals in caring disciplines have been pivotal in advancing virtual care, which leverages remote technologies to deliver effective support and services from a distance. Educators in these caring professions are required to teach students the skills and competencies needed to provide high-quality and effective care. As virtual care becomes more integral, educators must equip students in these fields with both interpersonal and technological skills, bridging traditional hands-on learning with digital literacy. However, there is a gap in evidence exploring educators' perceptions and experiences of teaching caring profession students about virtual caring skills within online environments.

Objective: This study aims to better understand caring profession educators' online teaching experiences to foster student development of virtual caring skills and competencies.

Methods: We used a sequential explanatory mixed methods approach that integrated a cross-sectional survey and individual interviews with educators from caring professions to better understand caring professional educators' online teaching experiences to foster student development of virtual caring skills and competencies. The survey's primary objectives were to examine the various elements of existing e-learning opportunities, delve into educators' perspectives and encounters with these opportunities, and identify the factors that either facilitated or hindered online teaching practices to support students in developing virtual caring skills and competencies. The individual interview guides were based on survey findings and a systematic review of the evidence to gain deeper insights into educators' experiences and perspectives.

Results: A total of 82 survey participants and 8 interview participants were drawn from educators in the fields of education, medicine, nursing, and social work. Various instructional methods were used to help students develop virtual caring skills, including reflections on learning, online modules, online discussion boards, demonstrations of remote care, and consultation with clients. There was a statistically significant difference between educators' level of experience teaching online and their satisfaction with online teaching and learning technologies ($P<.001$) and between educators' faculties (departments) and their satisfaction with online teaching and learning technologies ($P=.001$). Participants identified barriers (time constraints, underdeveloped curriculum, decreased student engagement, and limited access to virtual caring equipment and technology), facilitators (clearly defined learning objectives, technology software and support, teaching support, stakeholder engagement, and flexibility), and principles of teaching virtual caring skills in online environments (connection, interaction, compassion, empathy, care, and vulnerability).

Conclusions: Our study identifies the barriers, facilitators, and principles in teaching virtual caring skills, offering practical strategies for educators in caring professions. This study contributes to the growing body of educational research on virtual caring skills by offering educator insights and suggestions for improved teaching and learning strategies in caring professions' programs.

As educational practices evolve, future research should explore how traditionally in-person educators can effectively teach virtual caring skills across diverse contexts.

(JMIR Nursing 2025;8:e64548) doi:10.2196/64548

KEYWORDS

health care education; virtual care; telehealth; online teaching; mixed methods study; student; teaching; virtual caring skills; cross-sectional survey; interview

Introduction

Background

Professionals in caring fields, including educators, physicians, nurses, and social workers, have played a crucial role in the ongoing development of virtual care where remote information technologies are used to ensure quality and effective care. The shift to virtual care has paved the way for innovative approaches to delivering care services, such as online teaching; remote health care and social services; and remote assistance for individuals, families, and communities to improve their social functioning, all from a distance. These virtual interactions demand digital literacy skills and comfort with technology, skills that traditionally may not have been intentionally integrated into formal education.

As virtual caring practices become integral to care provision, it is imperative that educators support caring profession students in acquiring the interpersonal and technological competencies necessary for providing virtual care. Traditionally, educators in caring professions relied on face-to-face lectures and seminar-style instruction with work-integrated learning placements, where students gained hands-on skills and collaborated with experienced educators and practicing health professionals in settings such as K-12 classrooms, hospitals, and counseling centers [1,2].

The shift to virtual teaching and care settings has challenged caring profession educators to incorporate alternative strategies for providing essential educational experiences to students [3-5] and placed added responsibilities on caring professionals to implement virtual care effectively in practice [2,3]. While the literature has long emphasized the need to support educators in meeting students' requirements [6,7], this need has become even more critical with the increasing prevalence of virtual care environments [8,9].

Higher education institutions have an opportunity to re-evaluate their approach to delivering online education in caring professions and identify the essential technological competencies necessary for success in today's virtual world. Given the significant transformation in education and care delivery, it is imperative that caring professionals possess the requisite skills and competencies to adapt and thrive in these new virtual environments. However, many caring profession educators face challenges when creating effective online learning experiences to prepare students for new virtual work environments, including limited bandwidth, the lack of technological devices, unfamiliarity with technological platforms, a lack of connection with students, and a lack of student engagement [10-13]. Learning new technologies can be cumbersome and frustrating [14], and technical issues can disrupt interactions that typically occur face-to-face [15-19]. These challenges underscore the necessity for a structured, evidence-based approach to developing and implementing educational technologies in online teaching and learning contexts to support virtual caring skill development [10,20-22].

The authors recently completed a systematic review from which they identified innovative online education initiatives that harnessed learning technologies for the education of caring professionals and demonstrated a growing emphasis on assisting students in cultivating effective virtual caring practices suitable for today's virtual environments [23]. The systematic review [23] highlighted a pressing need for greater emphasis on assessing and training educators to immerse students in digital technologies, thus fostering the development of both interpersonal and digital skills essential for delivering virtual care. More research is needed regarding educators' experiences and perceptions of teaching virtual caring skills.

This Study

Adding to the limited body of literature would potentially enhance the understanding of best practices in online instruction to promote the development of virtual caring skills. Therefore, we conducted this study to answer the following research questions: (1) How do caring professions' educators *describe* the online instructional methods used that support student development of virtual caring skills and competencies? (2) What are caring professions' educators' *experiences and perceptions* of online learning opportunities for helping students develop virtual caring skills and competencies? and (3) What are the *facilitators and barriers* to creating and engaging in online teaching that supports students' development of virtual caring skills and competencies?

Methods

Design

We adopted a sequential explanatory mixed methods study design [24] to gather, analyze, and integrate quantitative and qualitative data. We used a cross-sectional survey and conducted individual interviews to gain insights into the online teaching experiences of educators in caring professions in supporting students to develop virtual caring skills and competencies. The integration of the 2 research phases became apparent when the design of the interview guide was informed by the survey findings, enabling us to delve deeper into the results obtained from the survey. Furthermore, integration occurred as we used the qualitative findings to better understand the quantitative findings, ultimately forming interpretations from the integrated findings.

Sample and Participants

Voluntary participation was sought from educators in caring professions, including education, medicine, nursing, and social work (including those cross appointed to arts and veterinary medicine) across a midsized research-intensive institution in western Canada. Any self-reported educators from the abovementioned faculties were included in the study. No completed surveys or interviews were excluded.

Data Collection

We crafted a survey using established methods as outlined by Rattray and Jones [25]. The survey's primary objectives were to examine the various elements of existing e-learning opportunities, delve into educators' perspectives and encounters with these opportunities, and identify the factors that either facilitated or hindered online teaching practices to support students in developing virtual caring skills and competencies. The survey encompassed a combination of Likert scale, closed-ended, and open-ended questions, covering demographics, experiences, instructional methods, satisfaction levels, technology use, effectiveness, and readiness. To ensure the survey's validity, both in terms of face and content, we conducted a pilot study with a sample of 10 educators who did not participate in the study. Their suggested edits were incorporated into the survey before its dissemination.

To distribute the survey securely, we used an online platform, Qualtrics (Qualtrics International Inc). Our recruitment efforts spanned various channels such as email, Twitter (subsequently rebranded as X), Instagram (Meta Platforms), and Facebook (Meta Platforms), mirroring the methods used in prior studies [26,27]. Completion of the survey was considered as an indication of informed consent. In addition, we invited all survey participants to share their email addresses if they were interested in participating in a follow-up interview.

To gain deeper insights into educators' experiences and perspectives, we developed a semistructured interview guide based on the findings from a systematic review [23] and the responses received in the survey. We reached out to all survey participants who provided their email addresses and conducted interviews lasting between 30 and 60 minutes via the Zoom (Zoom Communications) platform. Before each interview, we confirmed oral consent, and the sessions were audio-recorded and transcribed verbatim.

Data Analysis

The closed-ended survey responses were obtained from Qualtrics and subsequently imported into the SPSS (version 28; IBM Corp) statistical software package for analysis. Descriptive statistics were calculated to summarize the characteristics of the study sample, including factors such as age, gender, faculty affiliation, length of time in current position, and previous experience with online teaching and learning technologies. Variations in data distribution were summarized and visually presented through tables and graphical representations, following the guidelines outlined by Polit and Beck [28]. In addition, 1-way ANOVA and Kruskal-Wallis *H* tests were conducted to analyze differences in satisfaction and likelihood to use online teaching and learning technologies in the future to support students in developing virtual caring skills. These analyses were conducted as deemed appropriate, following the recommendations of Polit and Beck [28]. To enhance readability and facilitate subsequent post hoc analyses, participant-reported ages were collapsed into 4 categories: ≤39, 40-49, 60-59, and ≥60 years. Team members with experience in statistical analysis met and contributed to ensure the accuracy of these findings.

For the analysis of open-ended survey responses and interview transcripts, each was assigned a unique identifier and imported into NVivo (version 14; Lumivero) to manage qualitative data. Our qualitative data analysis followed a thematic approach using an inductive process, aligning with the methods proposed by Braun and Clarke [29] and Nowell et al [30]. To gain a comprehensive understanding of the data, 2 researchers (LN and SJ) independently reviewed the entire qualitative dataset. Consensus coding was completed as both researchers coded the same transcripts and compared results on a one-to-one basis. Each researcher assigned sections of text to relevant codes, and the coding was then merged and discussed. Regular monthly meetings were held to establish and ensure a shared understanding of initial codes.

Larger team meetings, involving all authors, were conducted to collectively scrutinize and further refine emerging patterns in the qualitative data, ultimately confirming the identified themes and subthemes. Throughout the analysis process, written memos and meeting minutes were maintained to document our approach and decisions. Adhering to research and reporting standards, we followed the Standards for Reporting Qualitative Research outlined by O'Brien et al [31] when reporting this study.

Data Integration

Integration occurred at 2 points in this study. First, the quantitative findings were used to inform the qualitative interview guide. Following an independent analysis of all qualitative and quantitative data, the data were integrated using a joint display as an analysis tool. During this analysis, qualitative data were used to explain and corroborate quantitative findings [32]. Quantitative findings were compared to qualitative themes to examine similarities and differences. Through this methodology, we were able to develop interpretations regarding educators' perceptions and experiences.

Ethical Considerations

We obtained approval from our local Conjoint Health Research Ethics Board (REB22-0748) to carry out this study. Educators were offered the opportunity to join the study voluntarily, with the assurance that their involvement in the survey would remain anonymous and would not affect their university employment status or career advancement. Completion and submission of the online surveys implied consent. Before participating in interviews, all respondents gave informed verbal consent. Interviews were administered by a graduate student who had no prior supervisory relationship with the participants. To protect participant anonymity, distinct identifiers were assigned to each participant, and the data were aggregated accordingly. No compensation was provided to participants for participating in this study.

Rigor

We used several techniques to ensure the rigor of our study. Regular team meetings provided opportunities for debriefing, introspection, and deliberate questioning of our interpretations, as suggested by Morse [33]. We maintained a comprehensive audit trail that included codebooks, meeting minutes, and shared files to document all study-related decisions, following the guidelines proposed by Carnevale [34]. While 2 researchers were responsible for coding all qualitative data, the broader research team assessed and deliberated on decisions related to themes and subthemes. We revisited the raw survey and interview data to further validate our findings and ensure that they authentically represented the voices of the educator participants.

Results

Participant Demographics

A total of 82 educators started the survey, and 72 (88%) completed the entire survey. The 10 (12%) participants who did not complete the entire survey completed up to the final 5 survey items. We included all responses provided by participants in our final analysis as they yielded valuable insights and contributed to our overall study findings. Of the 82 survey participants, 19 (23%) agreed to be contacted for a follow-up interview of which 8 (10%) responded and completed an interview. Table 1 provides participant demographics for the survey and interviews.

Table 1. Participant demographics.

Demographic and demographic subcategory	Survey (n=82), n (%)	Interview (n=8), n (%)
Age (y)		
<39	11 (13)	0 (0)
40-49	25 (30)	1 (13)
50-59	29 (35)	3 (38)
>60	16 (20)	4 (50)
No response	1 (1)	0 (0)
Gender		
Men	18 (22)	2 (25)
Women	58 (71)	6 (75)
Gender diverse[a]	6 (7)	0 (0)
Faculty		
Education	21 (26)	3 (38)
Medicine	34 (41)	3 (38)
Nursing	16 (20)	2 (25)
Social work	7 (9)	0 (0)
Other (joint appointments)	4 (5)	0 (0)
Experience[b]		
Beginner	35 (43)	3 (38)
Intermediate	24 (29)	2 (25)
Expert	23 (28)	3 (38)

[a]Gender diverse included gender fluid, nonbinary, queer, and individuals who prefer not to disclose. Some categories were collapsed due to the need to maintain anonymity, particularly with small numbers in particular subcategories.

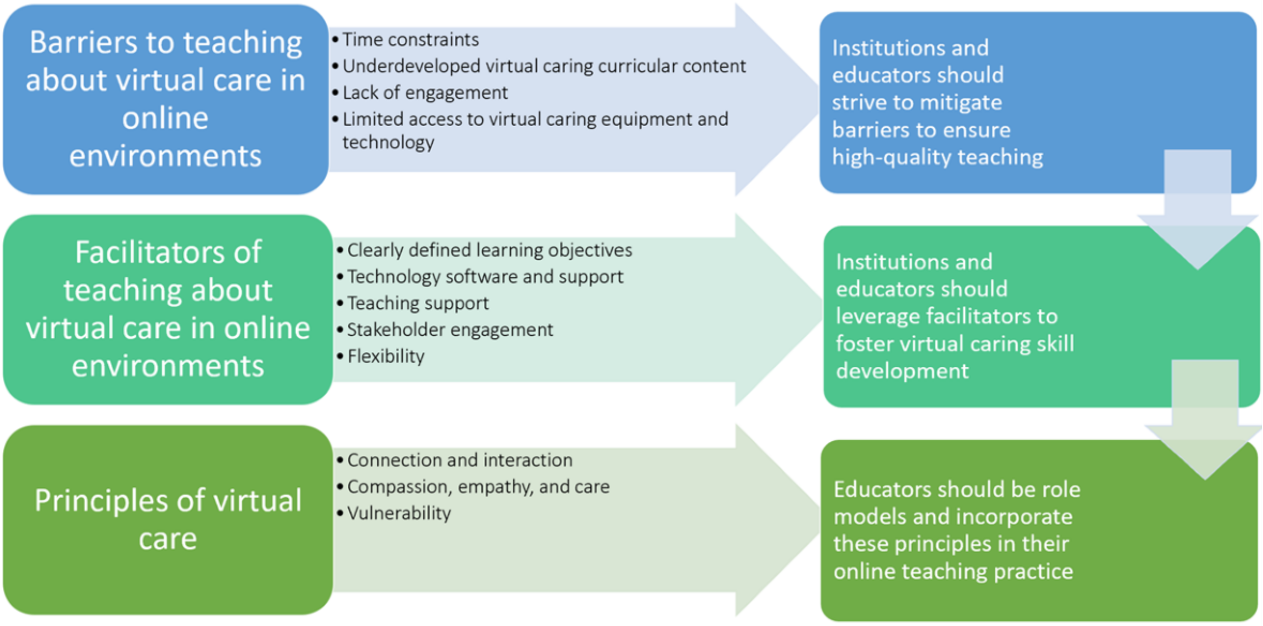
[b]Beginner=taught <4 online courses; intermediate=taught 5-7 online courses; expert=taught ≥8 online courses.

Quantitative Results

Overview

Educator survey respondents (n=82) indicated that a variety of online instructional methods were used to help students develop virtual caring skills in a *select all that apply* survey item (Figure 1). The most frequently reported online instructional methods included using reflections on learning (50/82, 61%), online modules (35/82, 43%), and online discussion boards (49/82, 60%). Educators reported using demonstrations of remote care (23/82, 28%) and consultation with clients (21/82, 26%). Respondents that used the option of *other* (7/82, 9%) described using verbal check-ins, synchronous meetings, simulations, social media, and flipped classrooms. Some respondents indicated that they have not used any online instructional methods to develop virtual caring skills (17/82, 21%).

Figure 1. Overview of themes, subthemes, and implications.



Satisfaction With Online Teaching and Learning Strategies

Survey participants (n=80) reported their level of satisfaction with online teaching and learning strategies, with 71 (89%) participants indicating that they were either satisfied or somewhat satisfied with the approaches used in their classrooms. However, a notable proportion, approximately 11% (9/80) of the participants, reported dissatisfaction.

Likelihood of Using Online Teaching and Learning Technologies

Among educators who responded to the question (n=70) about the likelihood of using online teaching and learning technologies to support students in developing virtual caring skills in the future, 53 (76%) indicated that they were very likely or somewhat likely to engage in this modality. Conversely, 17 (24%) educators responded that they were not likely to use online teaching and learning for the development of virtual caring skills in the future.

We conducted 1-way ANOVA tests to explore potential differences in satisfaction and likelihood to use technology scores among groups based on gender, age, faculty, years of experience in current position, or experience with online teaching and learning technologies. Table 2 summarizes the ANOVA test results.

In the survey, an expert was defined as an educator who had designed and taught ≥8 classes. There was a statistically significant difference between educators' level of experience teaching online and their satisfaction with online teaching and learning technologies ($F_{2,77}$=11.465; P<.001), with a large effect size (η^2=0.23) [28]. A Bonferroni post hoc analysis demonstrated that educators with expert experience with teaching using technology reported significantly higher satisfaction (mean 2.82, SD 0.39) compared to those at beginner (mean 2.06, SD 0.69) or intermediate levels (mean 2.21, SD 0.59). No statistically significant difference was found between those at beginner and intermediate levels. A statistically significant difference was found between educators' faculties (departments) and their satisfaction with online teaching and learning technologies ($F_{4,75}$=5.119; P=.001), with a large effect size (η^2=0.21). A Bonferroni post hoc analysis found that educators from the faculty of education (mean 2.75, SD 0.44) rated their satisfaction with online teaching and learning technologies significantly higher than faculty from medicine (mean 2.12, SD 0.64) or nursing (mean 2.07, SD 0.70). There were no statistically significant differences found between social work and the remaining faculties. Notably, all other comparisons via 1-way ANOVA tests yielded no statistically significant results.

Through Levene tests, two 1-way ANOVA test pairings were found to have unequal variances: (1) faculties and likelihood of using online teaching and learning technology and (2) years of experience and likelihood to use online teaching and learning technology. The Kruskal-Wallis *H* test, a nonparametric equivalent, was used to examine those relationships. A Kruskal-Wallis *H* test demonstrated that there was a statistically significant difference in the likelihood of using online teaching and learning technologies and the different faculties (H_4=13.44; P=.009), with a mean rank likelihood of 52.0 for the faculty of social work, 43.2 for the faculty of nursing, 36.8 for the faculty of education, 34.6 for other faculties, and 27.0 for the faculty of medicine. A pairwise comparison revealed that educators from the faculty of social work had a significantly higher likelihood of using online teaching and learning technologies than the faculty of medicine when considering the Bonferroni correction for multiple tests, P=.01. This was the only significant relationship found in the pairwise comparison after applying the Bonferroni correction. Kruskal-Wallis *H* test demonstrated that there was no statistically significant difference in the likelihood of using online teaching and learning technologies and years of experience (H_5=3.956; P=.56).

Table 2. ANOVA test results.

Variable comparison	Descriptive statistics		ANOVA		
	Participants, n (%)	Mean (SD)	*F* test (*df*)	η^2	*P* value
Gender and satisfaction (n=80)			1.38 (2, 77)	0.04	.26
Men	18 (23)	2.11 (0.68)			
Women	56 (70)	2.39 (0.65)			
Gender diverse	6 (8)	2.17 (0.75)			
Age (y) and satisfaction (n=80)			0.31 (3, 74)	0.01	.82
0-39	10 (13)	2.40 (0.52)			
40-49	25 (31)	2.24 (0.66)			
50-59	29 (36)	2.28 (0.80)			
>60	14 (18)	2.43 (0.51)			
Faculty and satisfaction (n=80)			5.12 (4, 75)	0.21	.001
Education	20 (25)	2.75 (0.44)			
Medicine	34 (43)	2.13 (0.64)			
Nursing	15 (19)	2.07 (0.70)			
Social work	7 (9)	2.71 (0.49)			
Other	4 (5)	2.00 (0.82)			
Experience (y) and satisfaction (n=80)			0.99 (5, 74)	0.06	.43
<1	5 (6)	2.20 (0.84)			
1-5	28 (35)	2.39 (0.63)			
6-10	21 (26)	2.43 (0.68)			
11-15	12 (15)	2.33 (0.65)			
16-20	5 (6)	1.80 (0.84)			
>20	9 (11)	2.11 (0.60)			
Online experience and satisfaction (n=80)			11.46 (2, 77)	0.23	<.001
Beginner	34 (42)	2.06 (0.69)			
Intermediate	24 (30)	2.21 (0.59)			
Expert	22 (28)	2.82 (0.39)			
Gender and likelihood to use (n=70)			1.68 (2, 67)	0.05	.20
Men	14 (20)	2.07 (0.92)			
Women	52 (74)	2.38 (0.80)			
Gender diverse	4 (6)	1.75 (0.96)			
Age (y) and likelihood to use (n=70)			0.28 (3, 36)	0.01	.84
0-39	6 (9)	2.50 (0.55)			
40-49	22 (31)	2.18 (0.91)			
50-59	26 (37)	2.35 (0.89)			
>60	14 (20)	2.29 (0.73)			
Online experience and likelihood to use (n=70)			1.92 (2, 67)	0.05	.16
Beginner	27 (39)	2.11 (0.89)			
Intermediate	22 (31)	2.23 (0.81)			
Expert	21 (30)	2.57 (0.75)			

Qualitative Findings

Overview

Figure 1 offers a summary of 3 overarching themes and their associated 12 subthemes, which were identified when analyzing the qualitative data. It also highlights potential recommendations for supporting online teaching to enhance the development of virtual caring skills. The subsequent sections delve deeper into the exploration of these findings.

Barriers to Teaching About Virtual Care in Online Environments

Educators identified several barriers that were encountered for online teaching and learning related to the development of virtual caring skills, including time constraints, underdeveloped virtual caring curricular content, a lack of engagement, and limited access to virtual caring equipment and technology. Despite these barriers, participants often highlighted their adaptability in addressing the needs of their students and teaching contexts.

Time Constraints

Participants reported time constraints as a concern, citing challenges such as the increased duration of virtual interactivity and the need to adapt clinical experiences for online platforms. Educators discussed time constraints as a limiting factor for teaching duration and described the need for adjustments. One participant shared the following:

> *Another barrier is always time, right?...For it not to be just text heavy and kind of interactive, you need time, and you don't necessarily have that.* [P6, interview, education educator]

In some cases, the concern expressed was for students who could not be online for extended durations:

> *Clinical online, it's eight hours online. We made a decision that was too much for the students to be online.* [P1, interview, nursing educator]

In addition to teaching time constraints, educators noted that additional time was required when offering experiential learning in practice:

> *Timing was always an issue. It seemed to take longer to do virtual appointments than in person.* [P8, interview, medicine educator]

The shift to virtual teaching and learning spaces prompted educators to be mindful of time constraints and their impacts. Despite these challenges, all participants adjusted to better cater to the needs of their students and those they would be caring for in practice.

Underdeveloped Virtual Caring Curricular Content

Educators reported difficulties in identifying key content related to virtual caring. As noted by a nursing educator, virtual caring content was often missing from the curriculum due to curricular overload:

> *Actually, I would say one of the things that I feel that is missing from the clinical practice for [this context] is the education part. We do a little bit of in the clinical, but to do the total education...we have to do it that way because there's no time to include absolutely everything.* [P1, interview, nursing educator]

Virtual caring skills and competencies were often considered a specialized practice and were therefore not traditionally incorporated into more generalist-focused curricula. However, the onset of the COVID-19 pandemic made virtual care a crucial competency for many caring professions. An educator from the faculty of education noted how students tried to balance the unknown of virtual care expectations with how they may be expected to practice virtually when they graduated:

> *...but [students] don't necessarily know what they're getting into because you're asking them to look at an area that is somebody else's whole specialization, and yet we expect all teachers to know this information.* [P3, interview, education educator]

These responses demonstrated the challenges educators faced due to the lack of virtual caring curricular content, potentially negatively impacting students' ability to provide virtual care in their future practice.

Lack of Engagement

Educators expressed concerns regarding the lack of student engagement they encountered in online settings when teaching about virtual care. They cited interruptions, decisions about cameras being on or off, and struggles connecting with students and colleagues as factors negatively impacting student engagement levels. The online environment posed a complex challenge, with frequent connectivity issues and interruptions. One interview participant noted the following:

> *We had cats and dogs. We had children interrupting...every time a student would come in [to the Zoom call], the doorbell [chime] would ring and [their] dog would go berserk. [They'd] be constantly shutting mute on, and then have to do something about that dog. So, [they]'d disappear from the screen and then come back.* [P1, interview, nursing educator]]

The debate over whether students should have their cameras switched on or switched off during virtual learning was raised, particularly in terms of establishing a sense of presence:

> *To providing an ethic of care is the cameras on cameras off issue...the preference for students to have their cameras off makes for a very difficult teaching environment...I can't see your face; I can't see your reaction.* [P3, interview, education educator]

Beyond visible student presence in a virtual class, educators also raised a concern about how virtual caring challenged their own engagement and sensory perceptions:

> *I can't sense what's going on for them...You can't feel the energy in the room, right? You can't see body language. You can't see nonverbal communication...These are professions where we rely on all of our senses. And in a virtual environment, they're not all there.* [P6, interview, education educator]

One survey respondent identified the following:

> *I can't see faces or check in with people who might show signs of confusion the same way I can in person. You can't "read the room" online.* [P23, survey, medical educator]

Limited Access to Virtual Caring Equipment and Technology

Educators expressed concerns about the limited access to virtual caring equipment and technology, which had a detrimental impact on interactivity. For one educator in medicine, the lack of equipment was an ongoing challenge:

> *I would say that the interaction suffered. We struggled with not having enough private computer space in the hospital. We struggled with not having cameras for the learners, and microphones, and that went on for quite a while.* [P8, interview, medicine educator]

For a nursing educator, the lack of student internet access was a challenge in teaching virtual care and creating environments for students to practice their virtual caring skills:

> *And there was one student who had to do [Zoom] on [their] cell phone, and she was using her minutes on her phone. It was getting too expensive. It was so much better if [they] just didn't use [their] cell phone...there were other students, their internet would go down.* [P1, interview, nursing educator]

Teaching and providing virtual care in rural and remote areas brought attention to the privilege of internet access and resources as well as the challenges faced by clients:

> *There are limitations in some of the other countries about their access to Wi-Fi…many people do not have access to Wi-Fi at home. Therefore, the scheduling is important. I think many centers then also have interruptions of their Wi-Fi and are constantly on and off, on and off, and that creates some problems for them. And finally, there's a few centers that the reason for that happening is that they lose electricity.* [P5, interview, medicine educator]

> *Remote learning in rural areas, it all depends on bandwidth...At the beginning, I didn't realize the reason why people weren't turning their cameras on...Tech is always a barrier...whether it's bandwidth, whether it's Zoom not working, whether our own internet.* [P6, interview, education educator]

The challenges, such as lack of equipment and poor internet accessibility, directly impacted educators' ability to teach students virtual caring skills and competencies. These considerations can also be challenges in working with experts or patients outside of the virtual classroom.

Facilitators to Teaching About Virtual Care in Online Environments

Educators identified several facilitators for online teaching and learning related to virtual caring skills. These facilitators included well-defined learning objectives, supportive technology software and assistance, effective teaching support, active stakeholder engagement, and a commitment to flexibility.

Clearly Defined Learning Objectives

Educators brought up their awareness of key graduate expectations, competencies, and learning objectives in both the open-ended survey questions and interview responses. Some educators were challenged in aligning new virtual contexts with previously defined learning competencies:

> *I've had to reconsider how my own caring is conveyed and recognized in different circumstances. I've also begun to theorize about how caring is connected to key graduate learning expectations and competencies.* [P66, survey, education educator]

Other participants, like this one from medicine, asserted that the learning objectives should remain consistent despite the shift to online learning:

> *I don't think we've changed the learning objectives. I think that they remain relatively constant, it's how you achieve them. And with the remote learning, the remote learning has allowed the interaction, but it's the interaction I think that's more important than the virtual way of doing things.* [P5, interview, medicine educator]

Other educators spoke about how the processes of learning caring competencies might not change in virtual contexts, but students may struggle to see the value of acquiring virtual caring skills:

> *If they [students] don't care about something, it doesn't become part of a learning repertoire. Then what you have to then wonder if you're just covering material for the sake of covering...it's not enough that I care about the ideas, I need to get them to care about the ideas as well.* [P4, interview, education educator]

Despite the various viewpoints on how learning objectives were achieved, there was consistent support across the participants for the development and use of clear learning objectives related to virtual care. Particularly for participants who had relied previously on in-person assessments of learning objectives, there was an intentionality to focus and be explicit on what the learning objective was and how virtual care considerations were necessary.

Technology Software and Support

Participants identified that possessing knowledge and intentionally using technology and virtual caring software could enhance the development of virtual caring skills. Others identified the benefits of providing orientation and skill development sessions to familiarize individuals with the use of technology:

> *Some of the [online] programs demanded a lot of interaction...so it started off with teaching people how to do things [in the online programs]*. [P5, interview, medicine educator]

Furthermore, survey participants asserted that ongoing technical assistance was important to successfully integrate new technologies into the virtual caring curriculum. One survey participant commented on the positive advancement of technology and its influence on education. They wrote, "The technology has come so far that teaching online is often equivalent to in person" (P51, survey, social work). Some of the examples of technology use included telehealth, podcasting, video creation, Zoom, and virtual simulations.

In one example with clinical practice, learners were actively engaged with a particular client population online. Students were tasked with using technology and software to interact with the client. Experimenting with the various features of the technology provided an additional way for students to learn new ways to establish connections with clients:

> *They hear the [diagnosis], and they're like, oh, they can't do anything. But they were having fun with the little apps that turn your hair green, or give you bunny ears, and stuff like that. So, they're going through and playing with all that kind of stuff. I don't even know where half that stuff is or how they find it. But it's hilarious and it's fun to watch. So, it becomes a medium and a tool kind of thing.* [P7, interview, nursing educator]

It is important for students to gain a clear understanding of how to use virtual caring technologies efficiently and effectively to make meaningful caring connections with clients. The perspective from both survey and interview participants reinforced that having access to the tools was important, with support and familiarity requiring time and resources for tool use competency.

Teaching Support

Educators identified that various teaching supports were necessary for fostering initial self-awareness and skill development when teaching in virtual settings. Ongoing development and the exchange of best practices helped build and sustain confidence and competence in using virtual caring technology. Many educators turned to others for teaching support, including teaching and learning departments and teams, or external networks to help support their personal learning needs. Others found teaching support from within their own faculty and professional organizations:

> *My colleagues and my own field professional organization was better in terms of teaching strategies or things to do within a lesson.* [P6, interview, education educator]

Overall, educators were motivated to seek out ways to enhance their teaching practice of virtual caring skills.

Stakeholder Engagement

Study participants identified stakeholder support and engagement as important to virtual caring skill development. One survey respondent contributed that "online teaching is forcing me to get creative...I learned to rely more on facilitating students' own motivation and initiative to seek community involvement" (P40, survey, medicine educator). Educators sought to encourage students to engage with clients in the community to help inform their virtual caring practices. Another survey participant indicated the importance of consulting various stakeholders, including students and educators, regarding their experiences with virtual caring technology by suggesting faculties should do the following:

> *[Engage in] consultation with students to understand their experiences as the end user/recipient of any technologies used for developing caring skills; [develop] a long-term vision/strategy for implementing, evaluating, and updating technology; [and link] technology use to program intent/pedagogy so that it makes sense to teachers/learners and is not just used for the sake of it.* [P5, survey, nursing educator]

Others highlighted the value of engaging with a range of stakeholders, including caregivers, clients, students, and instructors, in the virtual care setting:

> *We would invite the clients and the caregivers, or whoever was in the home to set up the screen and make sure that all of the controls were kind of off so that we could control it. And so as long as they could log in, we could get them into a breakout room. We would put the student in there with them. We would put a mentor from the [organization] in there with them. And then as instructors, we would go into each breakout room and just listen, make sure everything was okay, answer any questions, and then go to the next one and kind of wander through that way. And it worked really well.* [P7, interview, nursing educator]

Participants indicated that various stakeholders bring valuable and diverse perspectives to virtual caring experiences and harnessing these viewpoints can help facilitate more effective teaching and learning about virtual care.

Flexibility

Educators identified various ways that they chose to adjust, alter, change, or remain open to alternative ways of engaging in their practices for teaching, learning, and providing virtual care. The theme of flexibility emerged prominently in the survey responses, with a focus on being flexible with students. One survey participant emphasized the importance of "just being open and available and allowing students to set the stage for how they want to show up and learn and to be open if they are finding the online approach to learning challenging" (P79, survey, social work educator). Another perspective on flexibility was that it "allowed for more flexible scheduling and allow(ed) me to reach international students easier" (P15, survey, medicine educator). The connection to students in conducting, developing, or framing the learning space was recognized as a key element in building the flexibility to permit learning that incorporated virtual learning skills. This flexibility contributed to a more dynamic and inclusive learning environment.

Principles of Virtual Care

In our analysis, we identified principles of virtual care that reflect what educators reported as important considerations to how they approached teaching and learning virtual caring skills.

These principles include emphasis on connection and interaction; compassion, empathy, and care; and vulnerability.

Connection and Interaction

Educators identified how important connections and interactions were for teaching about and providing virtual care. This perspective was particularly present for a nursing educator who described how technical nursing skills were not as important as making personal connections with the clients, which is vital when providing virtual care:

> *[Students] felt that they were missing out on some of those skills, like IV starts because obviously we didn't do that [in a virtual environment]. But no, those are not the most important skills in nursing. It's the interaction. It's the education...nursing is not all about skills.* [P1, interview, nursing educator]

Some educators were thoughtful in their approach to providing students with purposeful opportunities to develop connections with clients:

> *I would want to be in a different room, with my camera off, observing the whole encounter...be the fly on the wall...and then be able to deliver feedback after the appointment.* [P8, interview, medicine educator]

All educators identified through the interviews that personal connections and prioritizing interactions were desired, and even necessary, before skill development in virtual environments.

Compassion, Empathy, and Care

Educators shared how emotional labor and intentional considerations are required to design learning experiences around compassion, empathy, and care, particularly in virtual contexts. One survey respondent suggested that "students of any caring profession know they need emotional bravery and an ability to handle very difficult situations with empathy and calmness even when they do not feel that way" (P34, survey, social work educator). Participants also indicated they needed this emotional bravery to successfully implement online teaching and learning technologies to support students in developing virtual caring skills. Educators acknowledged the impact and challenges associated with emotional labor and considered their role as educators in addressing issues like compassion fatigue:

> *Emotional labor and compassion fatigue...because those aspects impact the degree to which somebody wants to try something new or continue a practice that used to work, that doesn't seem to be working now.* [P3, interview, education educator]

One interview participant considered the impact of learning activities with a focus on social and emotional learning for individual well-being:

> *I also am a very active and dynamic facilitator, even online, so I use teaching strategies that I would use in the classroom and I get my students to actually get up and do things if I'm talking about a social emotional learning activity, something that's for wellbeing, because taking care of yourself is as important as you know what you're teaching, and you will impact the wellbeing of your own students or patients by the way you are as well. So, if I'm talking about just a simple social-emotional piece where it is maybe a five, four, three, two mindfulness activity, I do it with them.* [P6, interview, education educator]

Compassion, empathy, and care were viewed as important considerations in teaching, learning, and providing virtual care. These qualities could manifest authentically in a variety of ways, depending on the context of the teacher, learner, or client.

Vulnerability

The theme of care extended to include a focus on educator vulnerability and the willingness to embrace new approaches, recognizing that things might not always go as planned. However, this willingness by the educator required creating safer learning and caring spaces:

> *In caring skills and competencies, there's a level of vulnerability there that you must have. And so, when you're starting out with online courses, you need to build that caring atmosphere within your virtual online environment in a way that students feel safe.*
>
> *If you have a course, you have the time, and you utilize facilitation methods that are similar to what you are expecting them to be able to do as well, then that's helpful, right? I guess it comes back to that theory practice piece.* [P6, interview, education educator]

Another participant spoke about the need to break down barriers by creating relationships that push virtual caring efforts to meet clients' needs:

> *They [clients] put up their own barriers, to be perfectly honest. Because if you want it, you'll find a way to do it. But…If you have the goal in mind that, then all you need to do is figure out how to get there. It's a lot easier…I mean, create relationships. Ask people if they want to try something. And don't think you can't do it just because nobody's done it before...See if it works. Not everything works the first time. Well, I know that's why this is important too, right? It's like you evaluate and you figure out what works, what doesn't work.* [P7, interview, nursing educator]

There was a shared sense among participants that without the educator's sense of vulnerability and willingness to try something new and create intentional efforts toward connection through compassion and care, educational practice for virtual care would not be able to move forward.

Discussion

Principal Findings

In this sequential explanatory mixed methods study [24] we explored the experiences and perceptions of educators in caring professions as they navigated online teaching to facilitate the development of virtual caring skills and competencies among students. Educators identified both barriers and facilitators to engaging in this mode of teaching and learning as well as identified key principles underlying virtual caring.

Quantitative and qualitative data were integrated following individual analysis. The most common online instructional methods used to teach virtual caring skills were reflection, online modules, and online discussion boards. Only 26% (21/80) of the participants indicated that they provided experiential learning via consultation with clients on the quantitative survey. In qualitative interviews, participants discussed barriers to this educational modality, such as lack of time, indicating that providing virtual caring experiences could be less efficient than providing in-person clinical learning. Furthermore, 21% (17/80) of the educators indicated that they had not used online technology to teach virtual caring skills. This was reflected in the qualitative data when participants discussed the challenges of fitting more content into an already crowded curriculum. As virtual environments increase in the caring professions, it is important that virtual caring curriculum becomes a more permanent fixture within program curricula [35], rather than treated as a specialty consideration that can be included if time permits. This highlights the attention for program-level considerations for technological literacy and use development. It is not enough for educators to be able to use the technology effectively and use tools in one course; instead, there is a need to identify opportunities across a program to support the learning and development of digital literacy and technology-use competencies.

Educators had varying levels of satisfaction with their online teaching and learning strategies to enhance virtual caring skills. Less than half of the participants (34/80, 43%) indicated that they were satisfied with their online teaching and learning strategies, with other educators indicating that they were either somewhat satisfied or not satisfied. Through the qualitative survey and interview data, educators expressed frustration regarding the lack of engagement or connection with their students, which created difficult teaching environments. Educators also expressed concern regarding students' access to technology devices and reliable internet. Bolster et al [35] expanded this idea when discussing that clinical patients that might have limited access to virtual caring technologies or may lack digital literacy. In this study, the challenges discussed by educators may have influenced their overall satisfaction with their ability to execute effective teaching and learning strategies. In the survey qualitative responses, those that were "satisfied" (34/80, 43%) often cited reasons such as the smooth functioning of technology and active student engagement. Educator and student interactions with technology appear to be influential to educators' satisfaction with the teaching experience. Leaders from across the United States emphasized the importance of optimizing the logistics of technology when they met for a symposium titled Crossing the Virtual Chasm: Rethinking Curriculum, Competency, and Culture in the Virtual Care Era [35]. They reported that the need to optimize logistics, including providing equitable technology access and user software training, was one of the levers that can improve virtual care education [35].

Although educators' likelihood to use online teaching and learning technology was mixed in quantitative surveys, there was notable support to develop learning objectives to enhance virtual caring skills. Educators discussed facilitators that could enhance the teaching and learning of virtual caring skills in interviews. Survey respondents who identified as very likely to use online teaching and learning technologies (37/70, 53%) indicated via qualitative responses that teaching support through professional development, ongoing technology assistance, and student engagement was essential to support students in developing virtual caring skills. Addressing challenges that arise while teaching and learning virtual caring skills in an online environment can be beneficial to student outcomes and educators' satisfaction and increase their likelihood to use such technologies. Although higher education institutions are working to keep up with evolving technologies, specialized attention will be required in the virtual caring education context [35].

Surveys and interviews were undertaken with educators across caring professions, including education, medicine, nursing, and social work, within a research-intensive educational institution in western Canada. Quantitative analysis revealed interesting insights into educators' satisfaction with online teaching and learning strategies and their likelihood to use online teaching and learning technologies. Overall, educators were somewhat satisfied with the online teaching and learning strategies they were using in their classrooms. Furthermore, they felt that they were somewhat likely to use online teaching and learning technology to support student learning of virtual caring skills. Through inferential analysis, we found that educators with experience designing and teaching ≥8 classes (considered expert level) had statistically greater satisfaction with the teaching and learning techniques they used in online learning environments. This finding indicates that educators could benefit from more experience in online teaching. This is congruent with the findings reported by Rhode et al [7], indicating that educators with more experience teaching in online environments had more positive attitudes toward online teaching and learning.

In addition, we found that educators from the faculty of education reported significantly higher satisfaction levels in teaching virtual care in an online modality compared to their counterparts in medicine or nursing. This may be largely due to the longer history that education faculties may have had in providing instruction in an online environment. This finding highlights the importance of offering additional support and professional development to educators in traditionally in-person programs, enabling them to effectively meet the needs of an increasingly online student population. In an integrative review, Cutri and Mena [36] discuss the cultural and structural challenges of traditionally in-person educators transitioning to online teaching and learning, including the workload required and readiness to transition to the online environment. Considering these challenges, academic institutions should consider implementing robust professional development programs to better support faculty engaging in online teaching and learning, ensuring optimal support for students learning virtual caring skills.

Educators identified several barriers to online teaching and learning related to the development of virtual caring skills, including time constraints, underdeveloped virtual caring curricular content, lack of engagement, and limited access to virtual caring equipment and technology. Time constraints may pose a significant challenge for educators as they strive to cover comprehensive content within limited time frames. Furthermore, educators may struggle to find room for virtual caring skills within their current curriculum, recognizing that to include additional content, other content will have to be reduced or eliminated. The underdeveloped nature of virtual caring curricular content may result in teaching and learning practices that lack the depth and breadth required to adequately prepare students for the nuances of virtual care. A notable barrier to teaching virtual caring skills in online environments, seen in this study and the literature, is the struggle to maintain student engagement, as online settings often hinder active participation and interaction. Students are more likely to be engaged when they have active learning opportunities, a positive learning climate, and meaningful interaction with faculty and peers [37]. Furthermore, limited access to virtual caring equipment and technology has exacerbated the challenge of teaching online [38] and hindered caring professionals' practical application of virtual care concepts [39]. Addressing these barriers is crucial to ensuring a robust and effective virtual care education within online learning environments.

Educators in this study identified several facilitators of online teaching and learning related to virtual caring skills, such as clearly defined learning objectives, technology software and support, teaching support, stakeholder engagement, and flexibility. Clear and well-defined learning objectives play a pivotal role in ensuring quality education, providing a road map for both educators and students to navigate curriculum with clarity and purpose. Adequate technology software and support are essential facilitators, enabling seamless integration of virtual caring skills into the online environment. Teaching support, including resources, training, and guidance, enhances educators' ability to effectively convey virtual caring concepts. In a grounded theory study, Shepherd et al [40] explored medical faculty and learner experiences regarding the learning of virtual caring skills during the COVID-19 pandemic. Despite medical faculty recognizing how virtual care can benefit patients, they were reluctant to continue to teach in virtual clinics, due to barriers at the individual, institutional, and systemic levels, citing challenging technology platforms and a lack of professional development as 2 of the limitations [40]. Stakeholder engagement, involving collaboration with health care professionals, institutions, and communities, may foster a more holistic approach to virtual care education. In addition, flexibility in instructional methods and assessment allows for adaptive learning experiences, catering to diverse student needs and optimizing the acquisition of virtual caring skills in an online setting.

Educators identified connection and interaction; compassion, empathy, and care; and vulnerability as key considerations when developing online teaching and learning experiences to support students in developing virtual caring skills. Fostering meaningful connections and interactions within the virtual learning space is essential for educators to create engaging and supportive learning environments. Encouraging compassionate and empathetic attitudes is fundamental, as these qualities are at the core of effective virtual care. Our findings mirrored the assertion by Bolster et al [35] that connection in virtual care is an essential component of "webside manner," indicating the importance of rapport building through technology. Integrating opportunities for students to understand and express vulnerability is equally important, as it promotes authenticity and a deeper understanding of the human aspect of health care. By prioritizing these elements, online educational experiences can transcend physical barriers, providing a rich and holistic foundation for students to develop the interpersonal skills necessary for effective virtual caring [16,41-44].

This study is part of a larger multistudy research project intended to provide a framework for virtual caring skill development in higher education. This study explores the educator's perspectives, while another study explores the student's perspectives. The final integrated findings will inform a framework to guide educators from varied professions as they develop virtual caring curricula. By gaining educator and student perspectives, we aim to provide a comprehensive view of core principles, competencies, teaching methods, facilitators, and barriers to teaching and learning virtual caring skills.

Strengths and Limitations

Our sequential explanatory mixed methods study provided a thorough examination of caring profession educators' perceptions of virtual caring skill development within a specific educational institution. The inclusion of participants from various caring professions offered diverse perspectives, enhancing the study's comprehensiveness. By incorporating surveys and interviews, the research amalgamated quantitative and qualitative data, enabling a more profound insight into educators' experiences and perspectives in online teaching related to virtual care. However, it is essential to acknowledge the study's limitations, warranting caution in interpreting the findings. The focus on a singular institution may limit the generalizability of these findings to broader contexts. Furthermore, the participant pool from a single institution may lack diversity, potentially affecting the external validity and transferability of findings to a more varied population. Despite these constraints, this study lays the groundwork for exploring virtual caring skill development, inspiring further research, and offering potential insights for enhancing the delivery of virtual care in educational settings.

Conclusions

Educators in caring professions require specialized knowledge and skills to effectively teach and support students in developing virtual caring skills and competencies. Our study highlights the barriers, facilitators, and principles of teaching virtual caring skills online. As we contribute to the growing body of educational research on virtual caring skills, we share insights from caring profession educators. Future research should continue to explore how educators in more traditionally in-person teaching and learning can be supported to meet modern-day needs. In addition, more evidence is needed to explore effective teaching and learning strategies to teach virtual caring skills in a variety of contexts. Our findings offer practical strategies to enhance teaching and learning within educational programs for caring professions.

Conflicts of Interest

None declared.

References

1. Bogo M. Field education for clinical social work practice: best practices and contemporary challenges. Clin Soc Work J 2015 Mar 18;43(3):317-324 [FREE Full text] [doi: 10.1007/S10615-015-0526-5]
2. Leading work-integrated learning in Canada. Future Skills Center. URL: https://www.cewilcanada.ca/ [accessed 2020-12-17]
3. Dewart G, Corcoran L, Thirsk L, Petrovic K. Nursing education in a pandemic: academic challenges in response to COVID-19. Nurse Educ Today 2020 Sep;92:104471 [FREE Full text] [doi: 10.1016/j.nedt.2020.104471] [Medline: 32502723]
4. Roskvist R, Eggleton K, Goodyear-Smith F. Provision of e-learning programmes to replace undergraduate medical students' clinical general practice attachments during COVID-19 stand-down. Educ Prim Care 2020 Jul;31(4):247-254 [FREE Full text] [doi: 10.1080/14739879.2020.1772123] [Medline: 32469632]
5. Van Nuland S, Mandzuk D, Tucker Petrick K, Cooper T. COVID-19 and its effects on teacher education in Ontario: a complex adaptive systems perspective. J Educ Teach 2020 Aug 06;46(4):442-451 [FREE Full text] [doi: 10.1080/02607476.2020.1803050]
6. Dede C, Jass Ketelhut D, Whitehouse P, Breit L, McCloskey EM. A research agenda for online teacher professional development. J Teach Educ 2008 Nov 26;60(1):8-19 [FREE Full text] [doi: 10.1177/0022487108327554]
7. Rhode J, Richter S, Miller T. Designing personalized online teaching professional development through self-assessment. TechTrends 2017 Jun 21;61(5):444-451 [FREE Full text] [doi: 10.1007/S11528-017-0211-3]
8. Darling-Hammond L, Hyler ME. Preparing educators for the time of COVID … and beyond. Eur J Teach Educ 2020 Sep 04;43(4):457-465 [FREE Full text] [doi: 10.1080/02619768.2020.1816961]
9. Quezada RL, Talbot C, Quezada-Parker KB. From bricks and mortar to remote teaching: a teacher education program‘s response to COVID-19. J Educ Teach 2020 Aug 02;46(4):472-483 [FREE Full text] [doi: 10.1080/02607476.2020.1801330]
10. Cleland J, McKimm J, Fuller R, Taylor D, Janczukowicz J, Gibbs T. Adapting to the impact of COVID-19: sharing stories, sharing practice. Med Teach 2020 Jul 13;42(7):772-775 [FREE Full text] [doi: 10.1080/0142159X.2020.1757635] [Medline: 32401079]
11. Ferri F, Grifoni P, Guzzo T. Online learning and emergency remote teaching: opportunities and challenges in emergency situations. Societies 2020 Nov 13;10(4):86 [FREE Full text] [doi: 10.3390/soc10040086]
12. Kidd W, Murray J. The COVID-19 pandemic and its effects on teacher education in England: how teacher educators moved practicum learning online. Eur J Teach Educ 2020 Sep 09;43(4):542-558 [FREE Full text] [doi: 10.1080/02619768.2020.1820480]
13. Sepulveda-Escobar P, Morrison A. Online teaching placement during the COVID-19 pandemic in Chile: challenges and opportunities. Eur J Teach Educ 2020 Sep 09;43(4):587-607 [FREE Full text] [doi: 10.1080/02619768.2020.1820981]
14. Chittleborough G. Learning how to teach chemistry with technology: pre-service teachers’ experiences with integrating technology into their learning and teaching. J Sci Teach Educ 2017 Feb 21;25(4):373-393 [FREE Full text] [doi: 10.1007/S10972-014-9387-Y]
15. Cantone RE, Palmer R, Dodson LG, Biagioli FE. Insomnia telemedicine OSCE (TeleOSCE): a simulated standardized patient video-visit case for clerkship students. MedEdPORTAL 2019 Dec 27;15:10867 [FREE Full text] [doi: 10.15766/mep_2374-8265.10867] [Medline: 32051850]
16. Lister M, Vaughn J, Brennan-Cook M, Molloy M, Kuszajewski M, Shaw RJ. Telehealth and telenursing using simulation for pre-licensure USA students. Nurse Educ Pract 2018 Mar;29:59-63 [FREE Full text] [doi: 10.1016/j.nepr.2017.10.031] [Medline: 29180228]
17. Love R, Carrington JM. Introducing telehealth skills into the Doctor of Nursing practice curriculum. J Am Assoc Nurse Pract 2020 Oct 07;33(11):1030-1034. [doi: 10.1097/JXX.0000000000000505] [Medline: 33038114]
18. O’Connor EA, Worman T. Designing for interactivity, while scaffolding student entry, within immersive virtual reality environments. J Educ Technol Syst 2018 Dec 11;47(3):292-317 [FREE Full text] [doi: 10.1177/0047239518817545]
19. Woodcock S, Sisco A, Eady M. The learning experience: training teachers using online synchronous environments. J Educ Res Inst 2015 Mar 04;5(1):52 [FREE Full text] [doi: 10.5590/jerap.2015.05.1.02]
20. Müller AM, Goh C, Lim LZ, Gao X. COVID-19 emergency eLearning and beyond: experiences and perspectives of university educators. Education Sciences 2021 Jan 05;11(1):19 [FREE Full text] [doi: 10.3390/educsci11010019]
21. Williamson B, Eynon R, Potter J. Pandemic politics, pedagogies and practices: digital technologies and distance education during the coronavirus emergency. Learn Media Technol 2020 May 21;45(2):107-114 [FREE Full text] [doi: 10.1080/17439884.2020.1761641]
22. Zhang W, Wang Y, Yang L, Wang C. Suspending classes without stopping learning: China’s education emergency management policy in the COVID-19 outbreak. J Risk Financial Manag 2020 Mar 13;13(3):55 [FREE Full text] [doi: 10.3390/jrfm13030055]

23. Nowell L, Dhingra S, Carless-Kane S, McGuinness C, Paolucci A, Jacobsen M, et al. A systematic review of online education initiatives to develop students remote caring skills and practices. Med Educ Online 2022 Dec 12;27(1):2088049 [FREE Full text] [doi: 10.1080/10872981.2022.2088049] [Medline: 35694798]
24. Creswell JW. A Concise Introduction to Mixed Methods Research, 2nd Edition. Thousand Oaks, CA: Sage Publications; 2020.
25. Rattray J, Jones MC. Essential elements of questionnaire design and development. J Clin Nurs 2007 Feb;16(2):234-243 [FREE Full text] [doi: 10.1111/j.1365-2702.2006.01573.x] [Medline: 17239058]
26. Topolovec-Vranic J, Natarajan K. The use of social media in recruitment for medical research studies: a scoping review. J Med Internet Res 2016 Nov 07;18(11):e286 [FREE Full text] [doi: 10.2196/jmir.5698] [Medline: 27821383]
27. Yuan P, Bare MG, Johnson MO, Saberi P. Using online social media for recruitment of human immunodeficiency virus-positive participants: a cross-sectional survey. J Med Internet Res 2014 May 01;16(5):e117 [FREE Full text] [doi: 10.2196/jmir.3229] [Medline: 24784982]
28. Polit DF, Beck CT. Nursing Research: Generating And Assessing Evidence For Nursing Practice, 11th Edition. New York, NY: Wolters Kluwer; 2013.
29. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
30. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis. Int J Qual Methods 2017 Oct 02;16(1):e007641 [FREE Full text] [doi: 10.1177/1609406917733847]
31. 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 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
32. Younas A, Durante A. Decision tree for identifying pertinent integration procedures and joint displays in mixed methods research. J Adv Nurs 2023 Jul;79(7):2754-2769 [FREE Full text] [doi: 10.1111/jan.15536] [Medline: 36524303]
33. Morse JM. Critical analysis of strategies for determining rigor in qualitative inquiry. Qual Health Res 2015 Sep;25(9):1212-1222 [FREE Full text] [doi: 10.1177/1049732315588501] [Medline: 26184336]
34. Carnevale FA. Authentic qualitative research and the quest for methodological rigour. Can J Nurs Res 2002 Sep;34(2):121-128. [Medline: 12425004]
35. Bolster MB, Chandra S, Demaerschalk BM, Esper CD, Genkins JZ, Hayden EM, Virtual CareMedical Educator Group. Crossing the virtual chasm: practical considerations for rethinking curriculum, competency, and culture in the virtual care era. Acad Med 2022 Jun 01;97(6):839-846. [doi: 10.1097/ACM.0000000000004660] [Medline: 35263303]
36. Cutri RM, Mena J. A critical reconceptualization of faculty readiness for online teaching. Distance Educ 2020 Aug 03;41(3):361-380 [FREE Full text] [doi: 10.1080/01587919.2020.1763167]
37. Cole AW, Lennon L, Weber N. Student perceptions of online active learning practices and online learning climate predict online course engagement. Interact Learn Environ 2019 May 23;29(5):866-880 [FREE Full text] [doi: 10.1080/10494820.2019.1619593]
38. Kuntz J, Manokore V. "I did not sign up for this": student experiences of the rapid shift from in-person to emergency virtual remote learning during the COVID pandemic. High Learn Res Commun 2022 Jul 01;12:110-146 [FREE Full text] [doi: 10.18870/hlrc.v12i0.1316]
39. Ortega G, Rodriguez JA, Maurer LR, Witt EE, Perez N, Reich A, et al. Telemedicine, COVID-19, and disparities: policy implications. Health Policy Technol 2020 Sep;9(3):368-371 [FREE Full text] [doi: 10.1016/j.hlpt.2020.08.001] [Medline: 32837888]
40. Shepherd L, McConnell A, Watling C. Good for patients but not learners? Exploring faculty and learner virtual care integration. Med Educ 2022 Dec;56(12):1174-1183 [FREE Full text] [doi: 10.1111/medu.14861] [Medline: 35732194]
41. Goldingay S, Boddy J. Preparing social work graduates for digital practice: ethical pedagogies for effective learning. Aust Soc Work 2016 Dec 11;70(2):209-220 [FREE Full text] [doi: 10.1080/0312407X.2016.1257036]
42. Liu C, Scott KM, Lim RL, Taylor S, Calvo RA. EQClinic: a platform for learning communication skills in clinical consultations. Med Educ Online 2016;21:31801 [FREE Full text] [doi: 10.3402/meo.v21.31801] [Medline: 27476537]
43. Pullen Jr RL, Silvers CA. Helping students embrace HIT. Nurs Manage 2018 Dec;49(12):17-21. [doi: 10.1097/01.NUMA.0000547841.89245.bd] [Medline: 30499855]
44. Rutledge C, Hawkins EJ, Bordelon M, Gustin TS. Telehealth education: an interprofessional online immersion experience in response to COVID-19. J Nurs Educ 2020 Oct 01;59(10):570-576 [FREE Full text] [doi: 10.3928/01484834-20200921-06] [Medline: 33002163]

Edited by E Borycki; submitted 19.07.24; peer-reviewed by D Patel, S Otero; comments to author 23.09.24; revised version received 11.11.24; accepted 28.11.24; published 15.01.25.

Please cite as:
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 Nursing 2025;8:e64548
URL: https://nursing.jmir.org/2025/1/e64548
doi:10.2196/64548
PMID:39608377

Effectiveness of Patients' Education and Telenursing Follow-Ups on Self-Care Practices of Patients With Diabetes Mellitus: Cross-Sectional and Quasi-Experimental Study

Mohammed Alsahli[1], PhD; Alaa Abd-alrazaq[2], PhD; Dalia M Fathy[3,4], PhD; Sahar A Abdelmohsen[5,6], PhD; Olfat Abdulgafoor Gushgari[7], PhD; Heba K Ghazy[3], PhD; Amal Yousef Abdelwahed[8,9], PhD

[1]Health Informatics Department, College of Health Sciences, Saudi Electronic University, Riyadh, Saudi Arabia
[2]AI Center for Precision Health, Weill Cornell Medicine-Qatar, Al Luqta St, Ar-Rayyan, PO Box 5825, Doha, Qatar
[3]Community Health Nursing Department, Faculty of Nursing, Kafrelsheikh University, Kafrelsheikh, Egypt
[4]Nursing Department, North Private College of Nursing, Arar, Saudi Arabia
[5]Department of Nursing Science, College of Applied Medical Sciences, Prince Sattam Bin Abdulaziz University, Wadi Aldawaser, Saudi Arabia
[6]Department of Medical-Surgical Nursing, Faculty of Nursing, Assiut University, Assiut, Egypt
[7]Public Health Department, College of Health Sciences, Saudi Electronic University, Jeddah, Saudi Arabia
[8]Public Health Department, College of Health Sciences, Saudi Electronic University, Dammam, Saudi Arabia
[9]Community Health Nursing Department, Faculty of Nursing, Damanhour University, Damanhour, Egypt

Corresponding Author:
Alaa Abd-alrazaq, PhD
AI Center for Precision Health, Weill Cornell Medicine-Qatar, Al Luqta St, Ar-Rayyan, PO Box 5825, Doha, Qatar

Abstract

Background: Information and communications technology can be used in telenursing to facilitate remote service delivery, thereby helping mitigate the general global nursing shortage as well as particular applications (eg, in geographically remote communities). Telenursing can thus bring services closer to end users, offering patient convenience and reduced hospitalization and health system costs, enabling more effective resource allocation.

Objective: This study aims to examine the impact of patients' education and telenursing follow-ups on self-care indicators among patients with type I and type II diabetes mellitus (DM).

Methods: In phase I, a cross-sectional descriptive analysis was conducted to evaluate the self-care practices of 400 patients with DM at Kafr El Sheikh University Hospital in Egypt. In phase II, a pretest-posttest experiment was applied with a selected group of 100 patients purposively recruited from phase I due to their low self-care practice knowledge to ascertain the impacts of a 4-week intervention delivered via telenursing. They were reminded via telephone follow-up communication of the importance of adhering to recommendations on physical activity, nutritional intake, and the management of blood sugar (ie, insulin). Data collection was undertaken using a structured quantitative questionnaire, encompassing sociodemographic characteristics, medical symptoms and history, and knowledge of DM. Paired *t* test analysis was applied to study pre- and postintervention self-care behaviors.

Results: Participants had a mean age of 49.7 (SD 11.5) years. More than one-third received their DM diagnosis over a decade previously (135/400, 33.8%) and were obese (147/400, 36.8%). Almost half (176/400, 44%) received insulin, and the majority had cardiac disease (231/400, 57.7%) and the DM symptom of elevated blood sugar levels while fasting (365/400, 91.3%). A relatively high score of DM knowledge was reported (255/400, 63.7%). Males exhibited significantly lower knowledge levels (102/200, 51%) compared to females (153/200, 76.5%; $P<.001$). The intervention was effective in improving knowledge of DM ($t_{99}=30.7$, two-tailed; $P<.001$), self-care practices ($t_{99}=53.7$, two-tailed; $P<.001$), and self-care skills ($t_{99}= 47$, two-tailed; $P<.001$) among patients with DM.

Conclusions: The emergent evidence suggests that patients' education and telenursing follow-ups have the potential to improve self-care behavior in patients with DM. The delivery of frequent nursing reinforcement via telenursing enables improved self-management while contemporaneously reducing the need for patients to visit clinical settings (ie, improving patient condition and reducing net health system costs). The outcomes of this research underscore the need to integrate telenursing within conventional care for DM, and more research is needed to longitudinally assay its efficacy and sustainability over the long term and in different clinical and geographical contexts.

(JMIR Nursing 2025;8:e67339) doi:10.2196/67339

KEYWORDS
diabetes mellitus; education; knowledge; self-care; telenursing

Introduction

Diabetes mellitus (DM) is a chronic metabolic disorder characterized by elevated blood glucose levels resulting from defects in insulin secretion, insulin action, or both. Insulin is a hormone produced by the pancreas that regulates blood sugar levels and facilitates the uptake of glucose into cells for energy. There are 3 main types of DM (type I DM, type II DM, and gestational DM), each with etiological and pathological characteristics. Type I DM is a condition of the autoimmune system, arising from the lack of functioning beta cells generating insulin. Type II DM is more common and is generally attributable to lifestyle attributes and nutritional factors (eg, sedentary behavior and high sugar consumption), albeit genetic predispositions are also instrumental. Gestational DM occurs during pregnancy and typically resolves after delivery, although it increases the mother's risk of developing type II diabetes later in life.

DM poses a significant threat to the safety of hundreds of millions of people worldwide, with disconcertingly escalating prevalence. It is estimated that 643 million people will be diagnosed by 2030, rising to 783 million by 2045, up from 537 in 2021 [1]. This estimated increase can be associated with global population growth and the rising prevalence of diabetes due to unhealthy lifestyle-related factors and aging populations. The prevalence is significantly higher in certain regions, including the Middle East, where more than 70 million people are currently affected by DM. According to the International Diabetes Federation, Egypt ranks ninth globally for DM prevalence. In early 2020, there were approximately 8.85 million people with DM in the country, representing a prevalence rate of 15.2% [2].

DM entails direct costs in itself, and it also entails secondary costs related to interlinked conditions (which may themselves be causative or reciprocally exacerbated by DM). DM is often associated with complications such as vision impairment and blindness, cardiovascular diseases, and kidney failure and may require foot amputation [3]. In order to mitigate the more serious impacts of the condition and enable patients to have a better quality of life, DM must be managed with a strong autonomous role of patients themselves, including consistent adherence to practices recommended for self-care, such as frequent monitoring of their blood glucose levels, appropriate nutritional intake, recommended levels of physical activity, and medication compliance [3].

While patients tend to be aware of the imperatives associated with such positive behaviors, they commonly struggle to implement them in their daily lives, especially as metabolic disorders and DM itself commonly arise from a knowing lack of compliance with positive behaviors (ie, the general public typically knows that eating large amounts of processed sugar and having a sedentary lifestyle will predispose them to DM, yet they continue to indulge in such behaviors, leading to or exacerbating diabetes) [3]. In low-income countries, research has consistently shown that a large proportion of patients with DM typically adhere to negative self-care, essentially manifesting poor control of their glycemic index and a commensurately elevated propensity toward serious resultant issues [4,5].

The accelerating development and adoption of many useful technological solutions in health care services over the last 2 decades have led to greatly expanded opportunities for the more effective management of chronic illness, including DM. Telenursing, which is defined as the use of technological channels (eg, telephone or video calls) to provide nursing services to individuals in remote locations, has offered ways in which to reduce the distance between health care services and patients, as well as reducing the need for some patients to attend traditional care venues (thereby reducing pressure on limited resources) [6]. Its obvious advantages include increasing health providers' interaction with service users, including for symptoms monitoring and educating service users without expensive and burdensome face-to-face clinical appointments.

Telenursing fundamentally increases the ability of health care professionals to deliver services remotely, which has obvious implications for more frequent monitoring of patient symptoms and escalating interventions where appropriate, with personalized assistance for service users in the comfort of their homes and everyday lives [7]. As DM management is particularly sensitive to general lifestyle factors, the telenursing paradigm can be particularly useful to extend the reach of health care providers to give patients with DM additional support and encouragement in their daily lives, especially with engagement for reminders and follow-up on particular issues [7].

It should be noted that telenursing benefits encompass important clinical outcomes in addition to practice expedience in communication; the more frequent and direct communication engendered by telenursing formats enables increased patient adherence to medication, self-care, and other outcomes, which intrinsically comprises improved quality of care and contributes to optimized patient prognosis [8]. A systematic review found that telenursed patients displayed statistically significant enhancement in their glycemic control, with 0.5% reduced HbA_{1c} (glycated hemoglobin A_{1c}) levels over half a year, alongside decreased BMI in some studies that effectively leveraged "combined" interventions [9].

Additionally, telenursing mitigates the burden placed on health services by obviating in-person (face-to-face) attendance at traditional care delivery venues, which is especially valuable in resource-constrained contexts, such as low-income countries or remote geographical regions [10]. In areas suffering from a dearth of conventional health care resources, telenursing offers essential care delivery channels for patients with DM, preventing the escalation of patients' conditions and reducing net health care costs (eg, timely telenursing interventions can reduce the need for hospital admission) [11].

Among the particular services that can be enhanced by telenursing, limited research has explored its potential to play a role in improving DM patients' capacity to undertake self-care practices. It appears to offer notable advantages, but differing results have been found in practice, with some studies reporting tangible positive outcomes, and others identifying substantive barriers in terms of technological issues and the stakeholder engagement, which can hamper the long-term sustainability of telenursing services [8]. A recent narrative review of 18 randomized controlled trials (RCTs) and 5 quasi-experimental studies worldwide concerning telenursing for DM care reported that a telenursing intervention of weekly telenursing contact over 3 months achieved no significant influence on BMI or weight loss, while a 6-month telenursing program attained no significant differences in either BMI or HbA_{1c} [8]. A systematic review of adherence to medication regimens among patients with DM found that there was no study that had reported consistent improvement due to telenursing [12]. Such negative findings are contrary to expectations, given the potential promise of telenursing; thus, further studies are needed to ascertain telenursing impacts on self-care practices among patients with DM in numerous different and varied health care settings.

This study seeks to fill this research gap by ascertaining the impacts of patients' education and telenursing follow-ups on self-care practices among patients with DM at an Egyptian tertiary hospital. Using a single-group pretest-posttest design and cross-sectional analytical approaches, this research sought to evidence telenursing's scope to enhance self-care practices, thereby improving the quality of care and outcomes for patients with DM. The outcomes can guide practice in clinical contexts and advance emerging studies on digital health solutions for the management of chronic diseases, especially DM. The insights gained from this research are particularly important in considering the intervention impacts to improve self-care practices among patients with DM, especially for contexts where conventional care is limited or hard to access.

Methods

Study Design

This 2-phase study encompassed a cross-sectional assessment of self-care (phase I) and a single-group quasi-experimental pretest-posttest design to assess the impacts of telenursing education on patients' knowledge, skills, and self-care (phase II).

Study Setting

The research setting was the outpatient clinic for diabetes care at Kafr El Sheikh University Hospital. This is the main diabetes care hub for the whole governorate. The sessions for patient education were delivered in specially allocated locations within the clinic, and the phase II follow-up interventions were delivered remotely using WhatsApp or SMS text messages.

Sampling

Inclusion Criteria

To be eligible for the study, participants had to (1) be adult patients with a diagnosis of DM for at least 1 year, (2) be aged between 18 and 65 years, (3) have access to and the ability to use a smartphone, (4) have HbA_{1c} level greater than 7, and (5) express interest in and willingness to participate in the study's interventions. Patients were excluded if they had psychological illnesses, speech or hearing impairments, or failed to respond to mobile phone contact for 2 weeks.

Sample Size and Sampling Technique

For phase I, clinical records for outpatients during 2020 were analyzed. The outpatient clinic records for the year 2020 were reviewed to determine the patient population. Using the Roasoft calculation program with a 50% response rate, a 95% CI, and a 5% margin of error, the required sample size was calculated to be 384. In order to attain more robust data, we purposively selected 400 eligible patients who met the inclusion criteria (above), comprising 200 males and 200 females.

The preliminary analysis of the data collected in phase I showed that 255 (63.7%) of 400 patients were categorized as having poor knowledge and poor self-care practice. Based on the inclusion criteria, a purposive sample of 100 patients was selected for phase II, focusing specifically on those with the lowest scores in both knowledge and self-care practices, as they were identified as the patients most in need of educational intervention.

Data Collection Tools

Sociodemographic and Medical Data Questionnaire

This tool gathered data on sociodemographic features such as age, educational level, and marriage status, and clinical attributes such as time since diabetes diagnosis, presence and type of comorbidities, fasting blood glucose levels, and BMI.

Knowledge Assessment Questionnaire

Participants' knowledge about DM (hereinafter "knowledge") was gathered using 23 open-ended questions divided into 8 categories: basic knowledge about diabetes and its complications (10 questions), treatment regimens (3 questions), physical exercise (2 questions), the importance of follow-up visits (2 questions), dietary patterns (2 questions), foot care (2 questions), bad habits that worsen the disease (1 question), and sources of knowledge (1 question). Responses were scored using a system where correct and complete answers received 5 marks, correct but incomplete answers received 4 marks, incomplete answers received 3 marks, incorrect answers received 2 marks, and answers of "don't know" received 1 mark. Each subsection score was averaged, with total knowledge scores ranging from 23 to 115 marks. Scores were then classified into poor knowledge (less than 60%, ≤69 marks) and fair knowledge (60% or more, ≥70 marks).

Self-Care Practices Questionnaire

This questionnaire focused on self-care practices among patients with DM and covered 43 different practices, which were categorized into 6 areas: nutritional practices and adherence to the DM dietary regimen (12 practices), practices related to medication regimen (5 practices), practices related to glucose monitoring (6 practices), practices related to physical activity (8 practices), practices to avoid complications (6 practices), and practices related to foot care (6 practices). The self-care practices were assessed using a 3-point Likert scale, where responses were rated as "always" (2 marks), "sometimes" (1 mark), and "rarely" (0 marks). Scores for each subsection were summed, and the total scores were classified into 3 categories: poor practices (less than 60%), fair practices (60% to <75%), and good practices (75% or more).

Self-Care Skills Checklist

The researchers used the self-care skills checklist in phase II to evaluate participants' practical self-care skills through direct observation. This assessment focused on 3 key tasks: preparing and injecting insulin (comprising 9 and 7 steps, respectively) and testing glucose levels in urine (9 steps). Conducting direct observations and assessments postintervention enhanced objectivity, providing a more reliable evaluation compared to patients' subjective self-ratings. The checklist used a 3-point scoring system for each step: 3 points for correctly performed steps, 2 points for incorrectly performed steps, and 1 point for steps not performed. Subtotal scores for each skill were calculated, and participants' performance was categorized as either satisfactory (≥60%) or unsatisfactory (<60%) based on the total score for each individual skill and the overall score.

Language of Data Collection Tools

Questionnaires in this study were used to accommodate the linguistic and practical needs of the participants and researchers. All questionnaires, except the self-care skills checklist, were in Arabic to ensure clear and effective communication with the study participants, who are native Arabic speakers. Delivering the questionnaire in their native language facilitated accurate comprehension of the questions and reliable responses, minimizing the risk of misinterpretation. The self-care skills checklist was in English, as it was designed for and completed by the researchers, all of whom possess a high level of proficiency in English. Using English for the researcher-administered questionnaire allowed for precision in recording and interpreting data while maintaining consistency with standard scientific and academic conventions. This dual-language approach ensured that both participants and researchers could engage effectively with the study materials, optimizing the validity and reliability of the data collected.

Piloting and Validation

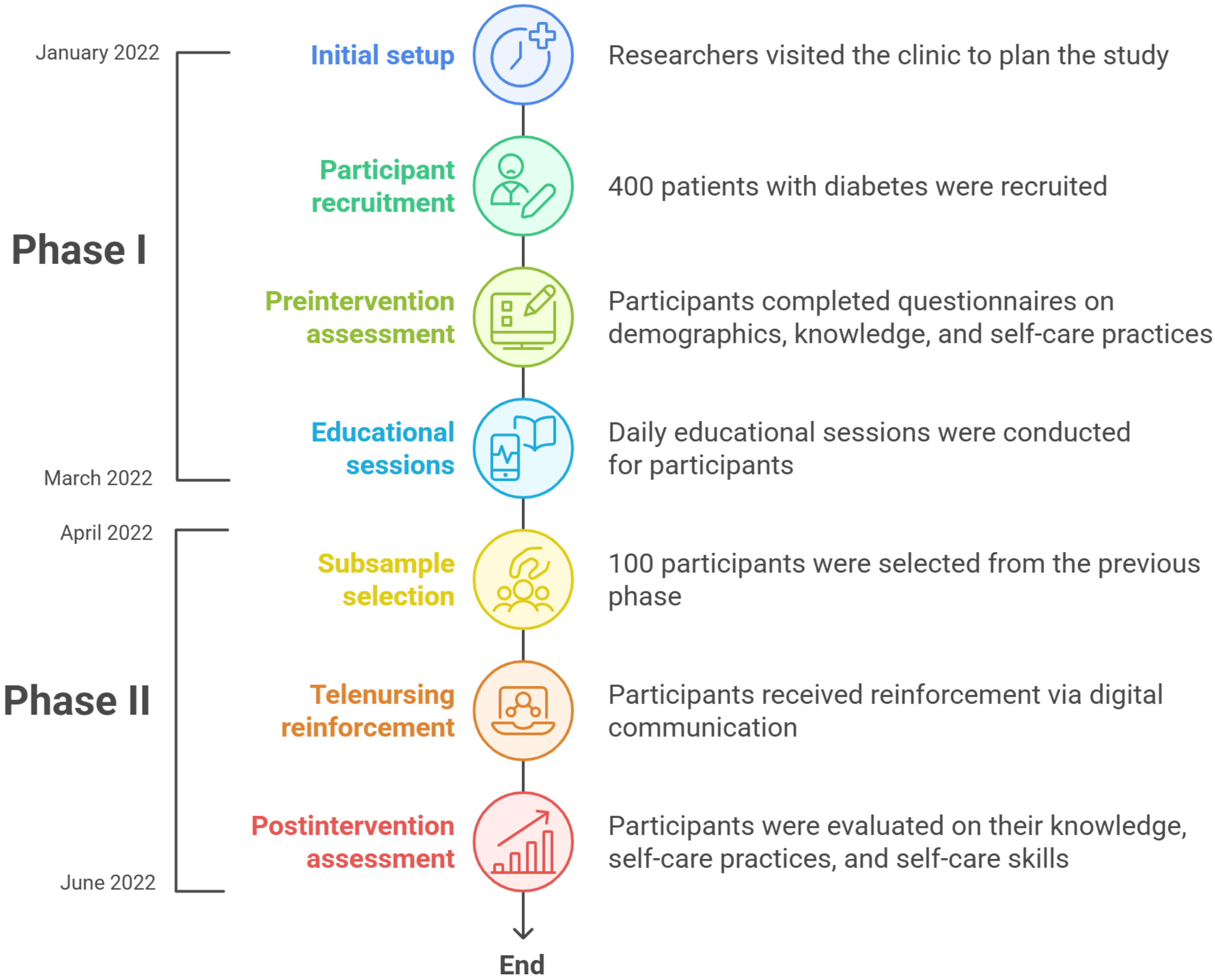
The developed tools were validated by a panel of experts from the Faculty of Nursing at Kafr El-Shiekh University. The panel consisted of 5 experts: 2 professors of medical-surgical nursing, 1 professor of medicine from the Faculty of Medicine, 1 assistant professor, and 1 lecturer of medical-surgical nursing from the Faculty of Nursing. The tool underwent both face and content validity assessments. The content validity focused on evaluating the clarity, appropriateness, applicability, wording, and comprehensiveness of the tool. To assess the internal consistency of the tool, the Cronbach α test was used. The results showed a Cronbach α of 0.78 for the knowledge assessment questionnaire, 0.8 for the reported self-care practice scale, and 0.88 for the diabetic self-care practice checklist. The same group of experts also validated the scientific content of the educational program.

After incorporating the experts' recommendations, the questionnaire was pilot tested. The pilot study was conducted over 3 weeks and included 10% (40/400) of the sample size (40 patients with DM in phase I and 10 patients in phase II). The purpose of the pilot study was to evaluate the clarity, applicability, and comprehensiveness of the tools and to assess the feasibility of the study process. Based on the findings from the pilot study, necessary modifications were made, such as the omission or addition of certain questions, to enhance the content, improve simplicity and clarity, and ensure the tools were concise and focused. The patients who participated in the pilot study were excluded from the main study sample.

Data Collection Process

Data collection spanned approximately 6 months, from January 2022 to the end of June 2022, and consisted of 2 phases as shown in Figure 1.

Figure 1. Data collection process.



Phase I

The researchers first visited the diabetic clinic to discuss the research objectives and methods with nursing leaders. During this visit, they coordinated meetings with potential participants and identified private spaces for conducting interviews and delivering the initial intervention sessions. A sample of 400 patients with DM, meeting the previously described inclusion criteria, was selected to explore their knowledge of DM and self-care practices. Potential participants were informed about the study's purpose and invited to participate if they were interested in receiving the intervention. Data collection took place for those who agreed to participate, with each interview lasting between 15 and 30 minutes (approximately 10 - 14 participants were interviewed per day, one-on-one).

Participants were grouped according to their outpatient appointments, and the educational intervention sessions were conducted in groups ranging from 10 to 15 members. Sessions were held daily, excluding Fridays, with each session lasting 30 - 40 minutes. Each session began with a welcome and icebreaker, followed by an explanation of the session's objectives and topics, and concluded with a recap and time for participants to ask questions. The sessions ended with an open discussion, allowing participants to address any clarifications, and handouts related to the content were distributed for participants to read at their convenience.

Phase II

As mentioned earlier, a purposive sample of a hundred patients from phase I was selected to undergo phase II. Participants underwent telenursing reinforcement of the educational intervention content from phase I via calls, SMS text messages, WhatsApp messages, videos, and voice notes. This was undertaken over 4 weeks (details of this intervention are provided in the following section). One month after the telenursing reinforcement, a posttest assessment was carried out at the outpatient clinic in Kafr Elshiekh Hospital. During this assessment, patients were interviewed to evaluate their knowledge, reported self-care practices, and self-care skills.

Intervention

Intervention Design

The intervention in this study was developed by the researchers, all of whom were diabetes nursing specialists, to enable patients with DM to enhance their self-care practices in response to the needs of patients. The intervention included educational sessions and telenursing follow-ups as described below.

Educational Sessions

The researchers delivered 6 educational sessions to the 400 participants in phase I at Kafr El Sheikh University Hospital's outpatient diabetes clinic. The educational sessions were held daily, excluding Fridays, with each session lasting 30 - 40 minutes. The educational sessions involved 3 theoretical and 3 practical sessions on diabetes. Specifically, the theoretical sessions covered essential topics such as basic knowledge of DM and its management (including medical treatment, physical exercise, dietary management, foot care, follow-up, and lifestyle habits that exacerbate the disease). The practical sessions focused on promoting healthy lifestyles (eg, dietary practices and physical exercises) and self-care practices (eg, insulin preparation and injection, glucose testing in urine, blood glucose monitoring, medication schedules, prevention of complications, and foot care practices).

To reinforce the learning experience, the educational content was compiled into a booklet distributed to participants after the sessions, serving as a reference for the information provided. A variety of pedagogical methods were used during the sessions, including practical demonstrations, abstract lectures, group discussions, and role-playing activities. These diverse teaching styles were designed to accommodate different learning preferences and build participants' confidence and adherence to the intervention. Additionally, visual aids such as images, physical models, and PowerPoint presentations were used to enhance understanding and engagement throughout the sessions.

Telenursing Follow-Up

During phase II, 100 participants were purposively selected for a 4-week telenursing follow-up. This intervention aimed to reinforce the educational content provided in phase I and support patients in adopting effective self-care practices. The follow-up schedule included daily 10 - 15-minute calls in the first week, twice-weekly calls in the second week, and weekly calls in the final 2 weeks. These personalized interactions focused on revisiting the educational material and addressing any questions or challenges faced by the participants.

During the follow-up period, participants received daily health education through various channels, including SMS text messages, WhatsApp messages, voice notes, and videos, all of which reiterated the information provided in the educational booklet. To further encourage adherence to self-care practices, daily reminders and audio recordings were sent to prompt actions such as blood glucose self-assessment, medication compliance, foot care, physical activity, and following the recommended diet plan. The content of the telenursing follow-up was meticulously developed by the researchers, drawing on insights from phase I, evidence-based diabetes self-care guidelines, and input from diabetes nursing specialists and clinical researchers. This ensured the content was accurate, culturally relevant, and aligned with the specific needs of the participants.

Statistical Analysis

Data collection, coding, and analysis were undertaken using SPSS (version 20, IBM Inc). Mean and SD values were used to report continuous data (with independent sample *t* testing to compare group differences) and frequency and percentage values for categorical data (with chi-square and Fisher exact probability tests to determine intervariable relationships). Pretest-posttest differences were determined using paired sample *t* tests and the McNemar test, indicating binary categorical variables' changes following the intervention. The robustness of the contingency table analyses was assured using Monte Carlo simulations. The application of these methods of statistical analysis affirmed cross-comparison results' reliability and mitigated risks of erroneously rejecting null hypotheses, as presented in the following section. After adjustments, a P value of ≤.05 was assumed to indicate statistical significance.

Ethical Considerations

Kafr El Sheikh University granted ethical approval for this study (KFIRB200-9). Studied patients gave verbal consent to taking part after full disclosure of the nature and scope of the research and their rights, including their ability to decline to take part or to subsequently withdraw without any consequences for their health care services or statutory rights. They were assured of their right to confidentiality, and that all data are reported anonymously in this study, with coding. All participants were informed that the data related to their participation would only be used for the current research purpose as per ethical guidelines for participant protection.

Results

Overall Findings

As described in the following subsections, significant shortcomings were discovered in participants' knowledge and skills at baseline, especially for female patients. The results after the intervention revealed significant enhancements in self-care practice and knowledge scores (P<.001). All of the patients were able to ascend from "unsatisfactory" to "satisfactory" scores in relation to skills for DM self-care, underscoring the efficacy of the intervention in enabling patients to achieve improved self-management of DM.

Sociodemographic Characteristics

The total studied sample comprised 400 people with DM, who had a mean age of 49.7 (SD 11.5) years. As shown in Table 1, the largest cohort (143/400, 35.7%) was aged 55 - 65 years, while over a fifth (86/400, 21.5%) each were aged 35 - 44 and 45 - 54 years. The vast majority of patients resided in family residences (399/400, 99.7%) and were married (325/400, 81.3%). A large minority (149/400, 37.2%) reported being illiterate, while almost a third (130/400, 32.5%) cited secondary school as their highest educational level and a negligible proportion (12/400, 3%) reported being university-educated. Females were significantly more likely to be illiterate (108/200, 54%) than males (41/200, 20.5%; P<.001). Furthermore, a negligible proportion (1/200, 0.5%) of male participants were unemployed, while none of the female participants were employed.

Table . Participants' sociodemographic characteristics.

Variables	Male (n=200)	Female (n=200)	Total (N=400)	χ^2 [a]	*P* value
Age (years), n (%)				4.6	<.001
29 - 34	42 (21)	16 (8)	58 (14.5)		
35 - 44	41 (20.5)	45 (22.5)	86 (21.5)		
45 - 54	40 (20)	46 (23)	86 (21.5)		
55 - 64	69 (34.5)	74 (37)	143 (35.7)		
65 - 72	8 (4)	19 (9.5)	27 (6.8)		
Mean (SD)	47.1 (11.4)	52.3 (11.1)	49.7 (11.5)		
Marital status, n (%)				9.2	.1
Single	8 (4)	3 (1.5)	11 (2.7)		
Married	170 (85)	155 (77.5)	325 (81.3)		
Widowed	22 (11)	42 (21)	64 (16)		
Living alone?, n (%)				-	.317[b]
Yes	0 (0)	1 (0.5)	1 (0.3)		
No	200 (100)	199 (99.5)	399 (99.7)		
Education, n (%)				59.2	<.001
Illiterate	41 (20.5)	108 (54)	149 (37.2)		
Literate	41 (20.5)	39 (19.5)	80 (20)		
Preparatory	21 (10.5)	8 (4)	29 (7.3)		
Secondary	87 (43.5)	43 (21.5)	130 (32.5)		
University	10 (5)	2 (1)	12 (3)		
Work status, n (%)				-	<.001[c]
Working	199 (99.5)	0 (0)	199 (49.7)		
Not working	1 (0.5)	200 (100)	201 (50.3)		

[a]Two-tailed.

[b]*P* value for Living alone? is based on Fisher exact test.

[c]*P* value for Work status is based on Monte Carlo exact test.

Clinical Characteristics

As Table 2 shows, one-third (135/400, 33.8%) of patients in this study received their DM diagnosis over 10 years previously, and the majority (231/400, 57.7%) had the comorbidity of cardiac disease. Concerning the latter condition, males (82/200, 41%) were significantly less likely to have it than females (149/200, 74.5%; *P*<.001). Almost half (176/400, 44%) of patients just received insulin treatment, while almost a quarter (93/400, 23.3%) additionally received oral hypoglycemic medications. The vast majority of patients exhibited elevated blood glucose (365/400, 91.3%), and most were overweight (146/400, 36.5%) or obese (147/400, 36.8%); females were disproportionately more prone to obesity (135/200, 67.5%) than their male counterparts (12/200, 6%).

Table . Participants' medical symptoms.

Medical data	Male (n=200), n (%)	Female (n=200), n (%)	Total (N=400), n (%)	χ^2 [a]	*P* value
Disease onset (years)				22.8	<.001
<1	28 (14)	42 (21)	70 (17.5)		
1 - 5	34 (17)	49 (24.5)	83 (20.7)		
5 - 10	48 (24)	64 (32)	112 (28)		
10+	90 (45)	45 (22.5)	135 (33.8)		
Other chronic diseases					
None	117 (58.5)	51 (25.5)	168 (42)	5.9	<.001
Cardiac disease	82 (41)	149 (74.5)	231 (57.7)	4.2	<.001
Hypertension	19 (9.5)	6 (3)	25 (6.2)	1.7	.541
Renal disease	0 (0)	2 (1)	2 (0.5)	0.51	.814
Rheumatic disease	2 (1)	1 (0.5)	3 (0.7)	0.5	.885
Liver disease	5 (2.5)	4 (2)	9 (2.2)	0.1	.924
Type of diabetes treatment regimen				28.2	<.001
Oral hypoglycemic drugs	65 (32.5)	66 (33)	131 (32.7)		
Insulin	107 (53.5)	69 (35.5)	176 (44)		
Both	28 (14)	65 (32.5)	93 (23.3)		
Commitment to follow-up schedule					
Always	200 (100)	200 (100)	400 (100)	N/A[b]	N/A
Fasting blood glucose				0.83	.662
Below normal	1 (0.5)	1 (0.5)	2 (0.5)		
Normal	14 (7)	19 (9.5)	33 (8.2)		
Above normal	185 (92.5)	180 (90)	365 (91.3)		
BMI				94.2	<.001
Normal	79 (39.5)	28 (14)	107 (26.7)		
Overweight	109 (54.5)	37 (18.5)	146 (36.5)		
Obese	12 (6)	135 (67.5)	147 (36.8)		

[a]Two-tailed.

[b]N/A: not applicable.

Baseline Knowledge Scores

In terms of knowledge, the majority (255/400, 63.7%) exhibited poor knowledge at baseline, albeit this was significantly less pronounced among males (102/200, 51%) than females (153/200, 76.5%; *P*<.001), as shown in Table 3. About half (98/200, 49%) of male participants had "fair" knowledge, while less than a quarter (47/200, 23.5%) of females did. Consequently, the outcomes underscore major differences in baseline knowledge among males and females, especially concerning comprehension of appropriate DM management practices, as affirmed by results on actual practices (discussed below), indicating the necessity of specific educational interventions targeted to females.

Table . Baseline diabetes mellitus knowledge.

	Male (n=200), n (%)	Female (n=200), n (%)	Total (N=400), n (%)	χ^2	*P* value
Knowledge level				28.1	<.001
Poor (<60%)	102 (51)	153 (76.5)	255 (63.7)		
Fair (≥60%)	98 (49)	47 (23.5)	145 (36.3)		

Baseline Self-Care Practices

At the beginning of the intervention, most patients (248/400, 62%) exhibited inadequate baseline self-care practices, albeit this was significantly lower among males (96/200, 48%; $P<.001$) than females (152/200, 76%), as shown in Table 4. "Good" practices for self-care were only reported among 24% (48/200) of males and 8% (16/200) of females. The lowest adherence was noted for blood glucose monitoring (316/400, 79%), physical exercise (296/400, 74%), and the prevention and management of acute complications (268/400, 67%). Critical shortfalls in self-care behaviors were thus observed, especially with regard to females, which indicates that more targeted interventions are needed to enhance essential self-care among female service users (in addition to the general need for improved self-care among DM patients in general).

Table . Pretest self-care practice scores.

	Male (n=200), n (%)	Female (n=200), n (%)	Total (N=400), n (%)	χ^2	*P* value
Total practice score				35.2	<.001
Good (≥75%)	48 (24)	16 (8)	64 (16)		
Fair (60 - 74%)	56 (28)	32 (16)	88 (22)		
Poor (<60%)	96 (48)	152 (76)	248 (62)		

Intervention Impacts on DM Knowledge

Table 5 demonstrates that the intervention achieved significant enhancements of patients' DM management knowledge for all studied domains ($P<.001$). The biggest improvements were seen concerning physical exercise knowledge, which saw a mean increase of 6.6 points, and tangible improvements were seen in knowledge of dietary choices and regimens of treatment.

Table . Mean of diabetes mellitus knowledge scores before and after the intervention (n=100).

Knowledge domains	Score, mean (SD)		Mean change[a]	*t* test[b]	*P* value
	Preintervention	Postintervention			
Basic knowledge about DM[c]	12.1 (3.7)	40.3 (7.4)	28.2	31.8	<.001
Treatment regimen	6.2 (1.9)	10.5 (1.5)	4.3	9.5	<.001
Physical exercise	1.2 (0.5)	8.2 (1.9)	7	37.4	<.001
Importance of follow-up visits	5 (0)	5.8 (0.4)	0.8	21.9	<.001
Dietary knowledge	3 (1.3)	6.5 (1.2)	3.5	22.9	<.001
Foot care knowledge	1.4 (0.6)	3.8 (0.6)	2.4	25.6	<.001
Knowledge of bad habits increasing DM severity	2.1 (0.5)	3.8 (0.5)	1.7	23.5	<.001
Total knowledge score	29.6 (3.9)	71.8 (13)	42.2	30.7	<.001

[a]Mean change = Posttest score – Pretest score.

[b]Two-tailed paired sample *t* test.

[c]DM: diabetes mellitus.

Intervention Impacts on Self-Care Practices

As shown in Table 6, the applied intervention achieved statistically significant enhancements of practices for self-care for all studied domains ($P<.001$). Mean increases of 8.15 points each were attained for the practices of "foot care" and "blood glucose monitoring," with a more modest increase in exercise practices of 3.45 points. These outcomes indicate that the intervention successfully improved participants' self-care behaviors for improved DM management.

Table . Mean of self-care practice scores before and after the intervention (n=100).

Practice domains	Score, mean (SD)		Mean change[a]	*t* test[b]	*P* value
	Preintervention	Postintervention			
Nutritional practices	8.4 (2.2)	21.2 (2.3)	12.8	40.4	<.001
Treatment regimen adherence	2.3 (0.6)	7 (0)	4.7	55.4	<.001
Monitoring of blood glucose level	1.6 (0.9)	10.3 (2.5)	8.7	36.9	<.001
Physical activities	3.8 (1.7)	14.2 (1.1)	10.4	55.1	<.001
Practices to avoid complications	5.2 (2)	10.2 (1.2)	5	21.9	<.001
Foot care practices	1.6 (0.9)	10.3 (2.5)	8.7	36.9	<.001
Total practice score	20.8 (4.5)	59.9 (7)	39.1	53.7	<.001

[a]Mean change = (Posttest score – Pretest score)/Pretest score.

[b]Two-tailed paired sample *t* test.

Intervention Impacts on Self-Care Skills

Significant improvements were seen following the intervention in patients' self-care skills, as shown in Table 7. Every participant went to "satisfactory" postintervention from "unsatisfactory" at baseline (400/400, 100%; *P*<.001), as reflected in the baseline scores for insulin preparation (mean 12.3, SD 2.1), self-injection (mean 11.8, SD 2.5), and glucose testing (mean 10.5, SD 1.9) increasing to 25.4 (SD 3), 24.8 (SD 2.8), and 23.5 (SD 2.4), respectively. These results underscore the effectiveness of the intervention to empower patients with prerequisite DM management skills, demonstrating the efficacy of practical instruction and reminders and reinforcement via telenursing, with the possibility of scalability for different and varied populations.

Table . Level of self-care skills before and after the intervention.

Self-care skills	Preintervention	Postintervention	χ^2 [a]	*P* value
Insulin preparation (n=58), n (%)			47	<.001
Unsatisfactory	58 (100)	0 (0)		
Satisfactory	0 (0)	58 (100)		
Insulin injection (n=58), n (%)			47	<.001
Unsatisfactory	58 (100)	0 (0)		
Satisfactory	0 (0)	58 (100)		
Urine glucose testing (n=100), n (%)			51.9	<.001
Unsatisfactory	100 (100)	0 (0)		
Satisfactory	0 (0)	100 (100)		

[a]McNemar test for related groups.

Discussion

Main Outcomes

Summary of Key Findings

This study on patient education and telenursing impacts concerning self-care practices among DM patients produced statistically significant outcomes, encompassing quantifiable improvements in skills, knowledge, and practices. Consequently, the intervention was effective in improving DM self-management and mitigating risks, as described below.

Improved DM Knowledge

The intervention resulted in patients with DM attaining significantly improved DM knowledge, especially concerning physical exercise, nutrition, and compliance with treatment. A more in-depth understanding of physical exercise was reflected in the postintervention increase in mean knowledge about physical exercise (and its impact on blood glucose) by 6.6 points [13]. This was striking, as education for patients with DM often lacks sufficient attention to physical exercise, despite its fundamental place in managing blood glucose and avoiding DM complications [13]. Improved knowledge scores concerning regimens and nutrition were also significant, and these outcomes are essential for the strategy of managing diabetes.

The intervention analyzed in this research effectively addressed existent educational needs among DM patients, offering them accurate and clear information they could apply, via easy-to-use formats (eg, SMS text messages, telephone calls, and WhatsApp). The ease of access enabled patients to effectively manage their conditions, which was particularly useful for the subset recruited for phase II, due to their particularly poor knowledge and self-care determined in the preliminary assessment. Although individual needs of specific patients were not targeted by the studied intervention, it was directed to commonly identified barriers and needs among patients requiring such services, offering scope for genuine enhancements in patients' outcomes and self-care behaviors.

Improved Self-Care Practices

Self-care practices significantly increased participants' scores for practices following the intervention, including monitoring blood glucose, undertaking appropriate foot care, and physical exercise. For monitoring blood glucose and foot care, participants achieved a mean improvement of 8.15 points each, highlighting the efficacy of the intervention in terms of encouraging positive practices to avoid long-term complications and deteriorating health conditions, including serious ones commonly affecting patients with DM due to a dearth of appropriate self-care (eg, neuropathy and foot ulcers) [14]. Physical exercise-related self-care practices also yielded an improvement of 3.45 points, showing more likelihood of undertaking exercise after the intervention. This addresses a core aspect of the management of diabetes, enhancing sensitivity to insulin and lowering the risk of cardiovascular damage [15].

Improved Self-Care Skills

The effects of the intervention on participants' self-care skills were substantial; all 100% (400/400) had "unsatisfactory" skills preintervention, and 100% (400/400) had "satisfactory" skills after it, in terms of preparing and injecting insulin and testing glucose in urine. This demonstrates the potentially remarkable effectiveness of patient education and telenursing follow-ups to enable patients with DM or other serious conditions to more proactively improve and maintain positive skills and behaviors, thereby improving their health outcomes (and substantially reducing costs for health systems).

Relation to Existing Literature

The outcomes of this study affirm those of the broader literature on positive telenursing impacts on the management of chronic diseases, such as DM [6,16]. Previous studies have extensively demonstrated particular impacts of telenursing in terms of enhanced engagement and medication adherence among patients, which ultimately contribute to improved prognosis [17,18]. This research contributes to the literature by presenting how a holistic telenursing intervention combining educational with skills-based content delivered via modern telecommunications (eg, WhatsApp messages) can facilitate major breakthroughs for patients in terms of increased self-care practices and DM management knowledge. This notably goes beyond most DM-related research, which tends to prioritize fundamental biomedical indicators of telenursing effectiveness (eg, HbA_{1c} and BMI), without commensurate attention to the holistic dimensions of DM care and self-management for patients (eg, exercise) [8,19].

Implications for Practice and Research

The intervention used in this research achieved notable benefits for patients, offering broader potential impacts for health practice and studies. For practitioners, the outcomes of this study affirm the effectiveness of patient education and telenursing follow-ups to improve diabetes care services, and personalized support and education delivered remotely via modern technologies, which are increasingly ubiquitous, can enlarge patient access to education and improve medication and healthy behavior adherence. Such impacts reduce demand for conventional clinical resources and avoid the escalation of negative DM-related conditions, thereby improving quality of care (ie, patient health and satisfaction) while achieving maximum resource deployment efficiency for health systems, which is essential for contexts with limited resources (eg, in low-income countries or remote geographical areas).

This research suggests that the effectiveness of patient education and telenursing follow-ups can be enhanced by adopting a patient-centered approach that addresses specific gaps in skills and knowledge among particular patient groups or individual patients. A personal paradigm considering each patient's particular requirements, as applied in this study, can enable patient education and telenursing follow-ups to offer its full benefits, reducing net costs on conventional health care resources, especially for chronic and serious conditions requiring improved self-management by patients, such as DM. In terms of implications for research, this study leaves open the requirement to investigate longitudinal effects of patient education and telenursing follow-ups to see if the advantages for self-care practices and DM knowledge recorded after a few weeks in this research can be sustained over time, and the extent to which they affect health indicators over the longer term (eg, reduced rates of DM complications and improved glycemic control). Furthermore, the cost efficiency of patient education and telenursing follow-ups in various potential applications can be compared to enable policy development to optimally deploy such initiatives for the maximum benefits. Finally, future studies should consider conducting 2-arm RCTs to compare patient education and telenursing follow-ups with standard care. Research involving diverse populations would also help determine the broader applicability of this approach. Additionally, integrating mobile health applications with automated reminders could enhance communication and patient adherence, potentially improving self-care related outcomes.

Limitations

The foremost limitation of this study pertains to its reliance on patients' own self-perceived and rated performance for some of the tools used, which is obviously subject to various forms of bias (including social desirability bias concerning self-care practices when reporting data in health care contexts). Furthermore, it was not possible to objectively measure patient indicators outside of clinical settings, including their exercise habits, nutrition, and blood glucose levels; such data would have offered improved, robust proof concerning the positive impacts of the intervention.

The used design, with a single group and pretest-posttest format, precludes the use of a control group, which consequently reduces the confidence with which observed changes can be solely attributed to the intervention. It should also be remembered that many patients in real clinical contexts lack access to the internet, smartphones, etc, due to socioeconomic and geographical barriers and digital literacy, which can affect the applicability of this and other interventions, undermining the equity of telenursing care. The single setting from which participants were recruited is also an issue that reduces generalizability. It is advised that researchers use objective methods of measuring patients' clinically relevant data in studies of their self-care behaviors, use control groups, and recruit participants from multiple contexts in order to generate more generalizable feedback about patient needs and the efficacy of interventions. Researchers should also always consider accessibility issues, including with regard to the use of digital technologies to deliver care.

Conclusions

The results affirm that an appropriately designed educational telenursing intervention can achieve significantly improved patient knowledge, self-care practices, and skills among patients with DM. Delivered via numerous modern methods of telecommunication, the intervention was successful in targeting essential issues in DM management to prevent complications, including monitoring blood glucose, physical exercise, and appropriate care for the foot. These results buttress calls for telenursing inclusion in conventional care for patients with DM, especially in contexts where conventional resources are not optimally accessible for all patients. This study's outcomes highlight the potential of patient education and telenursing follow-ups as an effective and scalable intervention to enable improved self-care practices, skills, and knowledge for patients with DM. The statistically significant enhancements demonstrated by this research support the use of patient education and telenursing follow-ups to help address the expanding costs of diabetes care, especially in contexts with limited resources. Nevertheless, more studies are required to ascertain whether the outcomes of this study are similar across different service user populations and to assay the long-term clinical and economic sustainability of education and telenursing solutions for such care.

Acknowledgments

The publication of this article was funded by the Weill Cornell Medicine – Qatar Health Sciences Library.

Conflicts of Interest

None declared.

References

1. Hossain MJ, Al-Mamun M, Islam MR. Diabetes mellitus, the fastest growing global public health concern: early detection should be focused. Health Sci Rep 2024 Mar;7(3):e2004. [doi: 10.1002/hsr2.2004] [Medline: 38524769]
2. Abouzid MR, Ali K, Elkhawas I, Elshafei SM. An overview of diabetes mellitus in Egypt and the significance of integrating preventive cardiology in diabetes management. Cureus 2022;14(7):e27066. [doi: 10.7759/cureus.27066]
3. Matoori SAO. Diabetes and its complications. ACS Pharmacol Transl Sci 2022 Aug 12;5(8):513-515. [doi: 10.1021/acsptsci.2c00122] [Medline: 35983272]
4. Ahmad F, Joshi SH. Self-care practices and their role in the control of diabetes: a narrative review. Cureus 2023;15(7):e41409. [doi: 10.7759/cureus.41409]
5. Opoku R, Ackon SK, Kumah E, et al. Self-care behaviors and associated factors among individuals with type 2 diabetes in Ghana: a systematic review. BMC Endocr Disord 2023 Nov 22;23:256. [doi: 10.1186/s12902-023-01508-x] [Medline: 37993843]
6. Mun M, Park Y, Hwang J, Woo K. Types and effects of telenursing in home health care: a systematic review and meta-analysis. Telemed J E Health 2024 Sep;30(9):2431-2444. [doi: 10.1089/tmj.2023.0188] [Medline: 37707998]
7. Dehghani A, Pourfarid Y, Hojat M. The effect of telenursing education of self-care on health-promoting behaviors in patients with multiple sclerosis during the COVID-19 pandemic: a clinical trial study. Mult Scler Relat Disord 2023 Feb;70:104507. [doi: 10.1016/j.msard.2023.104507] [Medline: 36682241]
8. 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]
9. Eberle C, Stichling S. Effect of telemetric interventions on glycated hemoglobin A1c and management of type 2 diabetes mellitus: systematic meta-review. J Med Internet Res 2021 Feb 17;23(2):e23252. [doi: 10.2196/23252] [Medline: 33595447]
10. Gajarawala SN, Pelkowski JN. Telehealth benefits and barriers. J Nurse Pract 2021 Feb;17(2):218-221. [doi: 10.1016/j.nurpra.2020.09.013] [Medline: 33106751]
11. Marlina TT, Haryani H, Widyawati W, Febriani D. The effectiveness of telenursing for diabetes self-management education: a scoping review. Open Nurs J 2023;17:e187443462307190. [doi: 10.2174/18744346-v17-230815-2023-38]

12. Teo V, Weinman J, Yap KZ. Systematic review examining the behavior change techniques in medication adherence intervention studies among people with type 2 diabetes. Ann Behav Med 2024;58(4):229-241. [doi: 10.1093/abm/kaae001] [Medline: 38334280]
13. Schubert-Olesen O, Kröger J, Siegmund T, Thurm U, Halle M. Continuous glucose monitoring and physical activity. Int J Environ Res Public Health 2022;19(19):12296. [doi: 10.3390/ijerph191912296] [Medline: 36231598]
14. Creber A, Leo DG, Buckley BJR, et al. Use of telemonitoring in patient self-management of chronic disease: a qualitative meta-synthesis. BMC Cardiovasc Disord 2023;23:469. [doi: 10.1186/s12872-023-03486-3] [Medline: 37726655]
15. Cannata F, Vadalà G, Russo F, Papalia R, Napoli N, Pozzilli P. Beneficial effects of physical activity in diabetic patients. J Funct Morphol Kinesiol 2020;5(3):70. [doi: 10.3390/jfmk5030070] [Medline: 33467285]
16. Liang HY, Hann Lin L, Yu Chang C, Mei Wu F, Yu S. Effectiveness of a nurse ‐ led tele ‐ homecare program for patients with multiple chronic illnesses and a high risk for readmission: a randomized controlled trial. J Nurs Scholarsh 2021 Mar;53(2):161-170. [doi: 10.1111/jnu.12622] [Medline: 33507626]
17. Ghoulami-Shilsari F, Esmaeilpour Bandboni M. Tele-nursing in chronic disease care: a systematic review. Jundishapur J Chronic Dis Care 2019;In Press(In Press):e84379. [doi: 10.5812/jjcdc.84379]
18. 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]
19. Shahsavari A, Bakhshandeh Bavarsad M. Is telenursing an effective method to control BMI and HbA1c in illiterate patients aged 50 years and older with type 2 diabetes? A randomized controlled clinical trial. J Caring Sci 2020 Jun;9(2):73-79. [doi: 10.34172/JCS.2020.011] [Medline: 32626668]

Abbreviations

DM: diabetes mellitus
HbA_{1c}: glycated hemoglobin A_{1c}
RCT: randomized controlled trial

Edited by E Borycki; submitted 10.10.24; peer-reviewed by C Li, H Monkman; revised version received 08.12.24; accepted 08.12.24; published 21.03.25.

Please cite as:
Alsahli M, Abd-alrazaq A, Fathy DM, Abdelmohsen SA, Gushgari OA, Ghazy HK, Abdelwahed AY
Effectiveness of Patients' Education and Telenursing Follow-Ups on Self-Care Practices of Patients With Diabetes Mellitus: Cross-Sectional and Quasi-Experimental Study
JMIR Nursing 2025;8:e67339
URL: https://nursing.jmir.org/2025/1/e67339
doi:10.2196/67339

Original Paper

Detailed Analysis and Road Map Proposal for Care Transition Records and Their Transmission Process: Mixed Methods Study

Elisabeth Veronica Mess[1], MA; Matthias Regner[1], MEng; Sabahudin Balic[1], MSc; Lukas Kleybolte[1], MSc; Lisa Daufratshofer[2], MSc; Andreas Mahler[2], MBA; Sabrina Tilmes[2], BSc; Viktor Werlitz[1], MSc; Claudia Reuter[1], PhD; Alexandra Teynor[1], PhD

[1]Institute for agile Software Development, Technical University of Applied Sciences Augsburg, Augsburg, Germany

[2]Digitization Staff Unit, University Hospital Augsburg, Augsburg, Germany

Corresponding Author:
Alexandra Teynor, PhD
Institute for agile Software Development
Technical University of Applied Sciences Augsburg
An der Hochschule 1
Augsburg, 86161
Germany
Phone: 49 82155863060
Email: alexandra.teynor@tha.de

Abstract

Background: The digitalization of health care in Germany holds great potential to improve patient care, resource management, and efficiency. However, strict data protection regulations, fragmented infrastructures, and resistance to change hinder progress. These challenges leave care institutions reliant on outdated paper-based workflows, particularly for patient data transmission, despite the pressing need for efficient tools to support health care professionals amid a nursing shortage and rising demand for care.

Objective: This paper aims to analyze Germany's care transition record (CTR) and CTR transmission process as part of transition management and suggests improvements toward a seamless digital solution.

Methods: To understand the current challenges of manual CTR transfers, we used a mixed methods approach, which included a web-based questionnaire with nursing professionals, field observations, business process model and notation modeling, semantic and frequency analysis of CTR entries, and user story mapping.

Results: A web-based questionnaire involving German nursing professionals (N=59) revealed considerable delays in patient care due to manual, patient-transferred CTRs. Of the 33 usable responses (n=33), 70% (n=23) of the respondents advocating for digital transmission to improve efficiency. Observations (N=11) in care facilities (n=5, 45%) and a hospital (n=6, 55%) confirmed the high administrative burden, averaging 34.67 (SD 10.78) minutes per CTR within a hospital and 44.6 (SD 20.5) minutes in care facilities. A semantic analysis of various CTRs (N=4) highlighted their differences and complexity, stressing the need for standardization. Analyzing a new CTR standard (care information object CTR) and manually mapping an existing CTR to it showed that the procedure was ambiguous, and some associations remained unclear. A frequency analysis of CTR entities revealed which were most used. In addition, discussions with care staff pointed out candidates for the most relevant entities. On the basis of the key findings, a stepwise transition approach toward a road map proposal for a standardized, secure transfer of CTRs was conceptualized. This road map in the form of a user story map, encompassing a "CTR transformer" (mapping of traditional CTRs to a new standard) and "care information object CTR viewer/editor" (in short, CIO-CTR viewer and editor; a new standard for viewing, editing, and exporting), shows a possibility to bridge the transition time until all institutions fully support the new standard.

Conclusions: A future solution should simplify the overall CTR transmission process by minimizing manual transfers into in-house systems, standardizing the CTR, and providing a secure digital transfer. This could positively impact the overall care process and patient experience. With our solutions, we attempt to support care staff in their daily activities and processes until nationwide state regulations are implemented successfully, though the timeline for this remains uncertain.

(JMIR Nursing 2025;8:e60810) doi:10.2196/60810

KEYWORDS

care transition record; transmission management; observations; process modeling; telematics infrastructure; TI; Fast Healthcare Interoperability Resources; FHIR; Health Level 7; HL7; medical information object; MIO; care information object care transition record; CIO-CTR; Pflegerisches Informationsobjekt-Überleitungsbogen; PIO-ULB; artificial intelligence; AI

Introduction

Digitalization in Health Care in Germany

Digitalization has emerged as a transformative force across various sectors, fundamentally altering organizational operations and service delivery. Health care is one sector benefiting significantly from digitalization as it can support patient care, resource management, and overall efficiency [1,2].

The growing shortage of qualified nursing personnel and the rising number of people needing care signify the need for more efficient, high-quality processes and tools to support health care professionals. Digital solutions offer a pathway to address these challenges by automating administrative tasks, improving communication between health care providers, and freeing up valuable time for direct patient care [3-5]. In Europe, policy makers, researchers, and health care practitioners are working to enhance health care infrastructure and promote interoperability to foster more efficient and coordinated care [6]. However, in Germany, the digital transformation of health care remains slow and faces significant obstacles [3,7,8].

Stringent data protection regulations for the processing of personal data (eg, the European Union's General Data Protection Regulation [GDPR] and Germany's Patient Data Protection Act, derived from the GDPR [9]) and fragmented technical infrastructures combined with the resistance to change make it difficult to integrate new tools or adjust existing processes [7,10]. In addition, the lack of a unified digital strategy further hinders the seamless implementation of digital health solutions [7,10].

Ultimately, the complexity of implementing digital solutions in the German health care system stems from balancing innovation with regulatory compliance, data security, and protecting patient privacy.

Care Transmission Process in Germany

A critical challenge within health care digitalization is ensuring the seamless transition of patient information between health care institutions. Paper-based workflows, still prevalent in many facilities, often cause delays and data loss during the transfer process due to the lack of standardized formats and the inability to share data in time.

Our research focuses on streamlining parts of the care transition record (CTR) transmission process to address this issue. The project's goal is to improve the transfer of patient data across care institutions, which currently suffer from time-consuming manual data entry, format inconsistencies, and delays in the arrival of crucial patient information.

State of the Art

Health Care Data Exchange

A security-conformant approach for digital transfer is the use of a dedicated health data (transfer) network. In Europe, such a service must be conformant to the GDPR, that is, legal compliance (ensuring data privacy and security), patient data control (data consent management for patients), data security (only access by authorized users, protection against breaches), and interoperability (fostering data exchange between different health care providers across various platforms) [9].

The telematics infrastructure (TI) is Germany's digital health data network designed to connect all health care providers, enabling the exchange of medical data across institutions [11]. It integrates various applications to streamline communication between health care entities such as physicians, hospitals, and pharmacies.

A specific way to exchange health data within a health data network is via an electronic health record. An electronic health record represents the digital version of a patient's medical history maintained over time by health care providers. It includes key clinical data relevant to patient care, such as medical history, diagnoses, medications, treatment plans, immunization dates, allergies, radiology images, and laboratory results. It is possible to share the patient data with other health care stakeholders, including the patient [12,13].

Several countries have made significant progress in this area, for example, the electronic patient dossier from Switzerland [14,15], electronic health record (Elektronische Gesundheitsakte) from Austria [16], MyKanta from Finland [17,18], and Mon espace santé from France [19]. They offer structured consent management for patients, meet the high security standards of the European Union, and foster interoperability by using the standard Fast Healthcare Interoperability Resources Health Level 7 (HL7), for exchanging electronic health care data.

The implementation of Germany's electronic patient record (ePA) [20] is progressing; however, it faces challenges. Many health care providers are not yet integrated, and patients must manually upload data. Technical and privacy issues, including interoperability concerns and strict data protection laws, continue to hinder broader adoption and use [21,22].

Another component within the TI is Kommunikation im Medizinwesen (KIM). It is a communication service with which health data can be exchanged directly by care providers, such as via email [23]. Nationwide implementation of the TI has been slow due to interoperability challenges. Adoption has lagged, primarily due to concerns over complexity, costs, and workflow disruptions. Health care professionals are hesitant to fully transition to digital tools because of these technical difficulties and the perceived burden of TI integration. There are 2 model projects in Germany [24,25] piloting and evaluating the TI and including components (eg, KIM and ePA). Unfortunately, no detailed evaluation reports have been published yet.

Standardization of Health Care Data

Standardization is a significant aspect that can improve the transfer of patient data in terms of reducing potential manual data entry and format inconsistencies. Standardizing CTRs is a potential possibility for improving the CTR transmission process. In Germany, 2 subsequent projects have focused on this issue: the first project is the ePflegebericht.

The ePflegebericht Project (electronic nursing report) began in 2002 when the Network for Continuity of Care in the Osnabrück Region [26] developed the concept for an electronic nursing report [27]. Insights from testing this software and its transition forms were gathered in a project under the patronage of the German Nursing Council starting in 2006. These insights were generalized beyond local use, placed in an international context [28], and aligned regionally and nationally [29]. The result was then submitted for approval as an HL7 standard [30].

The ePflegebericht served as a data exchange format for sharing information between care facilities and hospitals. It is based on the HL7 Clinical Document Architecture standard described in the study by Flemming et al [31]. The study validated the HL7 Clinical Document Architecture–based ePflegebericht and confirmed that it could cover all relevant nursing data compared with 114 paper-based nursing summaries used by 806 health care facilities in Germany. The ePflegebericht provided a comprehensive structure for transferring nursing information, demonstrating its applicability during care transitions. It improved the transmission of nursing data compared to paper-based methods, adding details such as social and homecare information, leading to more holistic documentation. Technically, advancements such as reusable templates were also introduced. These updates led to the relaunch of the ePflegebericht, with slight modifications, and in 2019, it was again up for approval.

The introduction of the ePflegebericht marks a significant advancement in the standardization of CTRs in Germany. A nationwide initiative aiming to develop a standard format for a variety of health-related documents (ie, medical information objects [32]) used the ePflegebericht as a foundational model for their CTR format: Pflegerisches Informationsobjekt-Überleitungsbogen (PIO-ULB). In this paper, the authors refer to the PIO ULB as care information object (CIO) CTR. This initiative was commissioned by the German government and overseen by the Gesetzliche Krankenversicherung Spitzenverband (central representative body of the statutory health and nursing care funds in Germany), and the mio42 GmbH (organization that develops medical information objects on behalf of the National Association of Statutory Health Insurance Physicians [Kassenärztliche Bundesvereinigung]). Furthermore, it involved the collaboration of the Deutscher Berufsverband für Pflegeberufe eingetragender Verein (DBfK; German Professional Association for Nursing Professions eingetragender Verein), and the Deutscher Pflegerat eingetragender Verein (German Nursing Council eV) [32]. CIO-CTR uses HL7 Fast Healthcare Interoperability Resources datasets, as described in the publication by mio42 GmbH [33], and was completed by the end of 2022. However, there is still uncertainty regarding the swift implementation of this new standard, primarily due to the financial and human resource challenges faced by health care software manufacturers who must adapt their existing products to comply with the specification, which spans approximately 2000 pages (XML code), as shown in mio42 GmbH [34]. The CIO-CTR will be effective at the beginning of 2025 [34] but without legal obligation for software manufacturers to implement it.

Objectives

This paper aims to analyze and address the challenges of the CTR transmission process in Germany. On the basis of a review of the current situation and possible approaches, a road map toward a fully digital, seamless solution is to be proposed. The overall goal is to improve the transfer of patient care data across care institutions.

Methods

Several methods were used to assess the satisfaction of nursing staff in the context of patient data transfer in care facilities in Germany. These include the creation of a web-based questionnaire, conducting field observations and contextual inquiries, business process model and notation (BPMN) modeling, semantic and frequency analysis of existing CTRs, and user story mapping. The findings are presented in this paper.

Web-Based Questionnaire

A web-based survey was conducted to identify challenges and preferences related to the CTR transmission process. The survey targeted nurses, nursing assistants, and trainees working in ambulatory, acute inpatient (eg, hospitals), or long-term care settings familiar with the CTR process. Participation was solicited through various channels, including the Bavarian State Ministry of Health and Prevention and the professional networks of project members. Due to a low initial response rate, the survey period was extended, and multiple reminders were issued. Using LimeSurvey, the survey ran from February 11, 2022, to April 30, 2022.

The questionnaire, developed iteratively by the project team (developers; care managers; and ethical, legal, and social issue experts), was based on literature and included custom questions and items from the validated Copenhagen Psychosocial Questionnaire tool [35]. Copenhagen Psychosocial Questionnaire items covered 12 domains, such as sociodemographic information (eg, gender, age, and work setting). Additional items focused on the experience with CTR creation and transmission, error rates, and attitudes toward digitalization. The 24-item questionnaire primarily used 4-point Likert scales, supplemented by nominal, metric scale, and open-ended questions. Respondents could opt out at any time, and all data were anonymized. A pretest with 7 participants from 2 independent institutions (implementation and nursing sciences) identified several structural and technical issues, which were addressed in a second pretest round. The same individuals tested the final version and did not reveal any issues.

Data analysis was conducted using SPSS (version 28.0.0.0; IBM Corp). Responses to open-ended questions were categorized using Microsoft Excel, and the data were checked for erroneous entries before being analyzed, focusing on descriptive statistics.

Field Observations and Contextual Inquiries

Overview

Field observations and contextual inquiries were conducted in a hospital and inpatient care facilities to understand the CTR transmission process thoroughly. These methods focused on the activities of care staff in their natural work environments, providing foundational insights for process modeling and research. The CTR transmission process in this study refers to all activities involved in creating a CTR at the sending facility and integrating it into the in-house system at the receiving facility, including the use of computer equipment, work tools, and telephone calls, while accounting for potential confounding factors. The observations aimed to clarify whether staff entered all data from the CTR at once or alternated between tasks.

Field Observation

Field observation, a qualitative research method, involves systematically observing participants in their natural settings to collect rich, contextual data on behaviors, interactions, and the surrounding environment [36]. An observation protocol was established to ensure consistency across sites and sessions, focusing on key areas such as activities performed, use of aids (eg, software and hardware), how information was handled and transferred, and any special features or abnormalities. Unobtrusive observation techniques were used to minimize observer effect, and detailed field notes were recorded, capturing both activities and nonverbal cues.

Contextual Inquiry

Contextual inquiry, a user-centered design method, was used to observe participants in their natural work environments while engaging in informal conversation to ask questions or clarify processes. This approach provided a deep understanding of the context in which tasks were performed and the challenges faced by users [37,38]. These inquiries, which were conducted primarily in participants' offices, allowed researchers to ask questions during task performance, facilitating an exploration of thought processes and decision-making, particularly with complex systems.

Execution

The observations and inquiries were conducted by 2 researchers, one with a medical background and the other specializing in user-centered design, ensuring comprehensive documentation and minimizing potential biases. The field observations and the contextual inquiries followed the same protocol. Thematic analysis [39] was applied to the data, with the researchers collaboratively reviewing and coding field notes to identify relevant patterns that informed the process modeling. In less formalized care facility environments, contextual inquiries were preferred, with researchers assuming an apprentice role to ask clarifying questions without disrupting workflows.

The observations were restricted to on-site care staff and did not include patients or external personnel (eg, patient transport). Observations occurred between 2020 and 2022, with no specific temporal or spatial restrictions within the facilities. Each observation was planned for 1 hour each.

Ethical Considerations

All studies adhered to ethical guidelines, and informed consent was obtained from participants. All data were anonymized. No incentives were offered. Ethics approval for the study was granted by the joint ethics committee of the Universities of Bavaria (GEHBa-202107-V-028).

BPMN Modeling of CTR Transmission Process

BPMN is an established and widely used graphical representation for modeling business processes. It is a standard developed by the Object Management Group (OMG) and has been adopted as an International Organization for Standardization (ISO) standard.

In BPMN, a process is represented as a sequence of activities or events, ordered in a flow that can be split or merged using gateways, directing the flow into one or multiple paths. Due to its simplicity, business process managers have widely used this standard in many application domains. Despite not being explicitly designed for clinical processes, BPMN has proven its value in the health care domain, allowing an easy-to-understand representation of clinical processes [40,41].

Semantic Analysis of CTRs

Semantic analysis is a good approach to extract and interpret the meaning of terms and sentences in detail. In the discipline of computer science, it is a fundamental component of natural language processing [42,43].

For semantic analysis, CTRs (empty and filled with fictive patient data) from cooperation facilities (n=4) were analyzed and compared in detail to better understand their structure, similarities, and differences. For clarification of any questions (eg, exact meaning, relevance, or scope of a specific category or word and overall comprehension), 1 meeting per facility with care staff was held. Given the semistructured to unstructured nature of the CTRs, it was critical to determine which data elements hold the same or different information compared to another facility. The meetings (n=4) lasted approximately 60 minutes.

Afterward, the CTRs were mapped to the new CIO-CTR standard. For this, the CTR entries were subdivided into entities and values and afterward mapped with pen and marker to the new standard format CIO-CTR.

Frequency Analysis of CTR Entities

Frequency analysis [44] is a method used to determine how often specific elements occur within a dataset, both in absolute terms and as a proportion of the total data. In this study, frequency analysis was applied to assess the occurrence of individual CTR entities to determine which pieces of information are most included. This helped inform the design of the proposed digital solution, ensuring that it prioritizes the most frequent CTR entries.

User Story Mapping

User story mapping [45] is a user-centric bottom-up technique used to outline a product or product feature. The output, known as a story map, provides a global view of the product, detailing the steps a user takes to achieve a specific outcome. This method helps prioritize tasks, identify dependencies, and adapt to changes.

Story maps are organized along 2 dimensions: the backbone (horizontal axis), which represents the user's activities step by step, and the release dimension (vertical axis), which defines the scope of the product and its various stages of development. A commonly used format for user stories is the role-feature-reason format: "As a <user>, I want to <feature> so that <value>" [45]. While a story backlog lists user stories in isolation, user story mapping provides a structured, global view of the entire application, fostering a common understanding between developers and stakeholders. This method also encourages communication, helping to eliminate misunderstandings early in the development cycle.

In the story mapping workshop, results from previous requirement analysis—including user feedback, product vision, and initial process modeling—are used to create actionable user stories. The key objectives of the workshop included understanding the user's perspective, identifying potential gaps, prioritizing, and release planning.

In total, 2 workshops were conducted, involving a total of 7 participants. These participants were part of the core research project team, bringing diverse expertise from various disciplines: health care (n=2, 29%), computer science (n=3, 43%), design (n=1, 14%), and IT security (n=1, 14%). All 7 (100%) participants attended both workshops, ensuring continuity and consistency in the discussions and decisions.

Results

Web-Based Questionnaire

A total of 59 participants participated in the web-based survey to determine the experiences and needs of nursing professionals regarding care transition reports, of which 35 (59%) met the inclusion criteria. Of the 35 participants, 2 (6%) did not finish the survey, resulting in 33 usable datasets. In Table 1, specific sociodemographic information about the participants is provided. An overview of the systems or software used is also provided in Table 2.

Table 1. Sociodemographic information of participants (n=33).

Characteristics	Participants, n (%)
Gender	
Women	22 (67)
Men	10 (30)
Nonbinary	1 (3)
Age group (y)	
18-24	2 (6)
25-34	13 (40)
35-44	7 (21)
45-54	8 (24)
>55	3 (9)
Care setting	
Short-term care (outpatient)	2 (6)
Acute inpatient care (hospital)	28 (85)
Long-term care (care facility)	3 (9)
Federal state (within Germany)	
Bavaria	32 (97)
Berlin	1 (3)

Table 2. Information about the system or software used.

Information	Participants, n (%)
System or software used for the creation of CTRs[a]	
I use software	20 (61)
I do not know	4 (12)
I use a paper form	5 (15)
I use a paper form and software	3 (9)
Not specified	1 (3)
Specific software used	
ORBIS (by Dedalus)	16 (49)
C&S	1 (3)
SAP	1 (3)
Sic Pflegeassistent (by CGM SYSTEMA SIC)	1 (3)
SnapAmbulant (by euregon)	1 (3)
Sorian	1 (3)
Not specified	12 (36)

[a]CTR: care transition record.

The high percentage of female participants (22/33, 67%) reflects the well-established predominance of women in nursing. The concentration of participants in the 25 to 34 age group suggests that the web-based survey may have been more appealing or accessible to younger adults. In addition, during the COVID-19 pandemic, care professional faced more stress and work, which might have led to a discouragement of answering a questionnaire that does not benefit their daily work.

The overwhelming representation of acute inpatient care (28/33, 85%) indicates a strong representation of hospitals in the questionnaire.

Of all the federal states in Germany, approximately all participants were from Bavaria (32/33, 97%) and only very few from Berlin (1/33, 3%). The overall overwhelming representation from Bavaria is probably due to the location of the research team, indicating that the recruiting efforts were particularly successful in this region despite numerous efforts to reach other care facilities and hospitals.

The results of the system or software used (Table 2) show that most (20/33, 61%) participants used software to create CTRs. Only 15% (5/33) of the participants used the paper form. Most (16/33, 49%) of the participants used the software ORBIS, reflecting the very high percentage of participants from hospitals, as ORBIS is a hospital information system. Sorian (1/33, 3%) is also a hospital information system. The other software listed (C&S, SAP, Sic Pflegeassistent, and Snap Ambulant) are documentation software used in the care facilities setting, which underlines the variety of software used.

Additional findings from the web-based questionnaire revealed that the CTRs were mainly transferred via the patient (27/33, 82%). This means that in these cases, the nurse gave the CTR to the patient as a printout, and the patient or the relatives were responsible for ensuring that it reached the next care facility. As a result, the nursing staff at the receiving facility has limited time to fully prepare for the patient in advance. Preparations and admission begin once the patient arrives at the facility, which can lead to waiting times. This is consistent with the results from the field observations that were conducted. The remaining 18% (5/33) transferred the CTR via fax, patient file, or telephone.

This gives the nursing staff more time to prepare for the patient, for example, preparing for isolation, special therapy treatment, or similar. According to the survey, the manual transfer of the CTR into the in-house system takes an average of 45 minutes, and 61% (20/33) of care staff perceived the transfer process as time-consuming. Manual transfer means that the care professional copies the information from the printout using their hand (typing on the keyboard) and transfers it to their care software. This step is necessary to add further information to the patient file, for example, information from patient conversations and decisions on care measurements.

This process can be time-consuming, as the care staff alternates between referring to the printout and typing the information into the system. During that time, confounding factors such as telephone ringing, colleagues, or technical issues can arise, prolonging the process.

Due to the use of different software in various facilities, the information is often displayed or organized differently, resulting in additional work.

Regarding the digital transmission of CTRs (cross-institutional dispatch and automatic integration into the in-house system), most (23/33, 70%) participants expressed no concerns. However, 30% (10/33) of them raised issues, such as concerns about possible threats to patient data protection (4/33, 12%). Most (24/33, 72%) respondents hope digital CTR transmission will reduce administrative effort. Some (18/33, 55%) participants indicated that they favored the standardization of CTRs because standardization of CTRs would result in relevant information being found more quickly in the future. On the basis of the responses, the primary consideration in developing a new solution should ensure, for example, that receiving, sending, and creating a CTR is less time-consuming for nurses than in the current process.

All (33/33, 100%) participants stated that CTR standardization would help them a lot as the CTRs they work with are usually different in structure and semantics.

Concerning the essential information in CTRs, all (33/33, 100%) participants agreed that patient information, medication, aids, and last bowel movement are considered to be very relevant regarding a potential standardization of CTRs. Finally, their opinion on automatic data integration was asked; they were curious as to whether something like this is possible so that they do not need to copy and paste information manually.

Field Observations and Contextual Inquiries

Field Observations

The observations focused on the receiving side of the CTR, that is, the creation of a CTR in the in-house primary system. This means that the scenario of a receiving facility was always observed. This focus on the receiving facility was agreed upon through collaboration with the facilities due to the COVID-19 pandemic, as stricter visitor restrictions prevented parallel observation in both the sending and receiving facilities. In all cases, the transfer of a patient was announced in advance.

The observation occurred from the moment the nurse sat down at their computer to either create the patient case or fill it in. At the hospital, the cases are already created by the administration and contain information that is necessary for billing but does not influence the nursing documentation any further. One nurse was observed during every observation, but it was not necessarily every time the same as it depended on their schedule. While at the hospital, both field observations and contextual inquiries were conducted; only contextual inquiries took place in the care facility.

The results of field observations at University Hospital Augsburg (UHA; n=6) in 2020, showed a high administrative time burden for nurses (refer to Table 3 for the CTR transmission process). Manual recording of CTRs resulted in an average time expenditure of 34 minutes. The observations showed that the CTRs were not sent in advance but arrived with the patient. While entering the data into their in-house system, the care staff mentioned that they could not prepare adequately for the patient in advance (eg, by preparing medications and nursing aids). The field observation also showed that the nursing specialist endures many interruptions while entering the CTR (relatives, colleagues, physicians, telephone, patients, or emergency calls), forcing them to switch between different tasks very often. Therefore, the nurse had to refocus on the CTR repeatedly.

Table 3. Overview of care transition record (CTR) transmission process observations at the hospital.

Observation	Observation duration (min)	Software	Interruptions, n	Type of interruptions	Resource used for transferring data	CTR present (print)
1	50	ORBIS	5	Relatives, telephone, colleagues, physician, and missing information	Computer and telephone	Present
2	25	ORBIS	3	Telephone and missing information	Computer and telephone	Present
3	40	ORBIS	5	Relatives, telephone, colleagues, and ambulances	Computer and telephone	Present
4	45	ORBIS	4	Colleagues, physician, telephone, and patients	Computer	Present
5	25	ORBIS	1	Physician	Computer	Present
6	23	ORBIS	3	Emergency calls and colleagues	Computer	Present

There was no direct association between observation duration and the care need of a patient; rather, it depended on the overall setting, for example, completeness of the available information, number of interruptions, and the length of the CTR.

The CTRs encountered in the field observations were all from different nursing homes with different lengths (approximately 12 and 30 pages). The information is primarily unstructured, that is, free text. Structured elements were primarily areas with checkboxes.

During the observations, many of the observed nurses complained that the manual transfer of the CTR was time-consuming. After the observation, the care staff were asked additional questions regarding the relevance of a digital CTR process, the most important information to be transferred, and their opinion about the fully automatic integration of the CTR data into the in-house system. The questions were open ended and digitally documented by the observer. In terms of relevance, all participants (n=6) said that the early, preferably digital, transfer of the CTR would hold immense value. It would help them to prepare in advance and obtain, for example, missing information and medication beforehand. It would reduce their administrative workload. In total, 3 (50%) of the 6 participants stated that the current process is frustrating as patients are often transferred before the weekend without medication, physician’s notes, or aids. Without these things, they have to come up with makeshift solutions to care for the patients over the weekend.

After conducting the field observations at the hospital, it became evident that specific questions remained unanswered and could not be answered fully in the follow-up discussion. These questions were about the specific functionalities of the software used and also specific work-arounds that were conducted by the care staff but not remembered after the observation. Therefore, one additional contextual inquiry was conducted.

Contextual Inquiries

Contextual inquiries (n=5) were conducted in 2021-2022 at 2 care facilities and UHA (Table 4). The results of the contextual inquiries provided valuable insights into the observation duration, the confounding factors, and the aids used. For most observations, documents about the patient (eg, physician's letter, medication plan, and CTR) were available as printouts. These were either sent with the patient or faxed to the referred institution. The latter could occur during registration or after inquiries about missing CTRs or information. The duration of 4 complete observations in care facilities and 1 hospital (excluding observation 2 because no input happened) averaged 47 minutes. In observation 2, it took 33 minutes to determine that no CTR was present, and it could not be sent from the sending facility. However, this required the nurse to make internal and external phone calls. She also needed to delegate procurement tasks to colleagues in the facility (eg, ask colleagues to check if the CTR might not be in the facility after all). In other cases, the CTR was handed out to the patient upon discharge but was not necessarily available right after the patient arrived at the receiving facility when the data were entered into the system.

Table 4. Overview of care transmission record (CTR) transmission process contextual inquiries, care facilities, and hospital. Care facilities are divided into facility 1 (CF1) and facility 2 (CF2).

Observation	Facility	Observation duration (min)	Software	Number of interruptions	Type of interruptions	Resource used for transferring data	CTR present (print)
1	CF1	55	Connext Vivendi NG+PD	1	Telephone, colleagues, and technical problems	Computer, smartphone, telephone, pen, and fax	Present
2	CF2	33	None used	0	No CTR present	Computer, smartphone, telephone, and fax	Not present
3	CF1	78	Connext Vivendi NG+PD	5	Telephone and colleagues	Computer, smartphone, telephone, paper, and pen	Present
4	Hospital	20	ORBIS	2	None	Laptop, paper, and patient	Present
5	CF2	37	Sic Pflege-assistent	3	Telephone	Telephone, paper, and pen	Present

All nurses involved in the contextual inquiries noted that the transfer process was time-consuming, particularly if they needed to retrieve missing information and also because they had to refocus on CTR data input due to interruptions.

Another interesting observation was that the nurses at the care facilities combined information from the CTR, physician's letter, medication plans, and the initial interview with the patient and entered these in free-text fields.

After the contextual inquiries, the same questions were asked as in the field observations. The nurses responded very similarly.

Comparing the average observation duration of all care facilities (excluding observation 2 because no input happened) with the hospital shows that the care staff requires approximately 56 minutes in the care facilities and only 20 minutes at the hospital.

BPMN Modeling

Overview

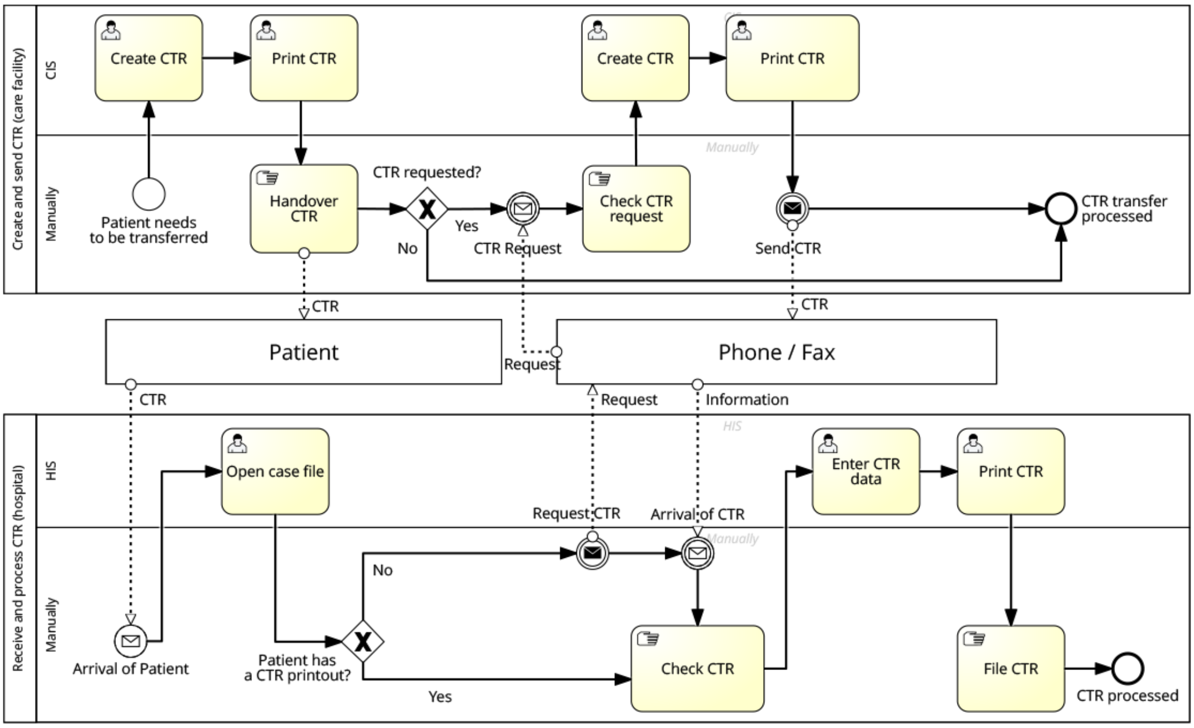
On the basis of the findings of the observations, BPMN models were created to better understand the various CTR activities (creating, sending, and receiving). These were discussed with the respective facility and detailed in the previous publication. After discussion, it was determined that the process models of the 2 care facilities can be combined into 1 process, as the activities are identical. Furthermore, the models were divided into different lanes, making it easier to understand which activities are manual and which are software based (human-computer interaction).

Transfer Process From Care Facility to Hospital

The process starts with the patient's need to be transferred (see Figure 1, Create and Send [Care Facility]). The nurse at the care facility creates a CTR, prints it, and usually hands it to the patient. Then, the patient arrives at the hospital, and the nurse at the hospital opens the patients' case file (see Figure 1, Receive and Process CTR Hospital). Afterward, she checks if the CTR printout is available and whether it is complete and error free (referring to the content of the CTR). Patients' case file is a digital file that contains the basic information of the patient for billing. As these files are prepared by the administration upon arrival of the patient, the care staff do not need to prepare those themselves.

If the CTR is complete and error free, she transfers the CTR into the hospital information system, prints the CTR in its specific structure, and files the CTR manually. After this, the CTR is processed, and the process is complete. If the CTR is unavailable, the nurse calls the sending care facility. The request is then processed there. If a CTR is missing, the sending facility creates a CTR, prints it, and sends it via fax to the hospital. Next, the nurse checks the document (eg, the correct CTR for the patient). After that, the CTR is transferred into the hospital information system, printed, and manually filed. Then, the process is complete.

Figure 1. Process modeling: care transition record (CTR) data transfer from the care facility to the hospital. CIS: case information system; HIS: hospital information system.

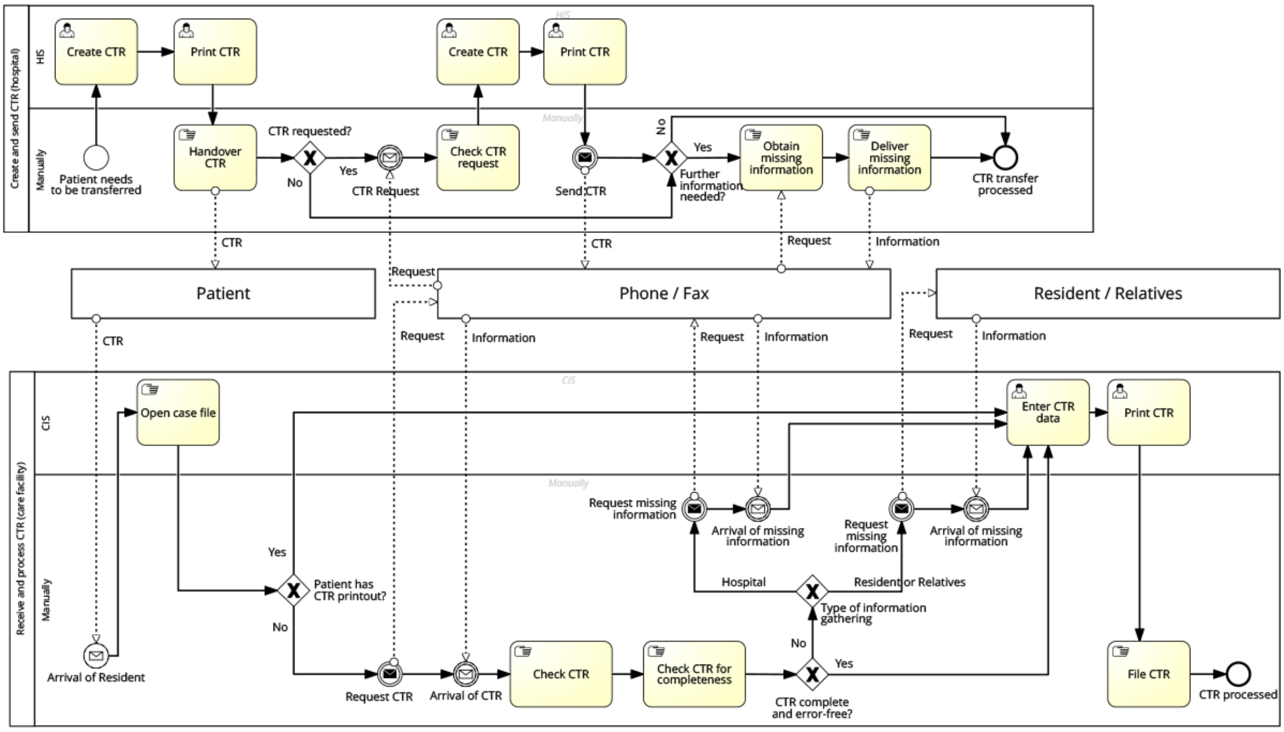


Transfer Process From Hospital to Care Facility

A patient is transferred from the UHA (Figure 2) to a care facility. At the UHA, the nurse creates a CTR, prints it, and hands it to the patient. Upon arrival of the patient (now called resident), the nurse at the care facility logs into their care information system and checks if the resident has a printed CTR. If so, they start transferring the CTR into the system. Afterward, they print the document in a proprietary file format and file it manually, then the process is complete.

If the CTR is unavailable, the nurse requests it from the UHA via phone. In the UHA, the request is checked and processed. A CTR is created, printed, and transferred via fax. Upon receipt of the missing CTR, the nurse checks whether it is the correct CTR for the resident and verifies its completeness and validity. If there is missing or incorrect information, the nurse requests the missing information either from the UHA via telephone or directly through the resident or relatives (this option does not exist in the process at the hospital). After the arrival of the missing information, the nurse starts transferring the CTR, prints it from the care information system, and manually files it. The CTR is processed, and the process is complete.

Figure 2. Process modeling: care transition record (CTR) data transfer from hospital to care facility. CIS: case information system; HIS: hospital information system.



Semantic Analysis of CTRs

A total of 4 CTRs of cooperation facilities were analyzed regarding their structures, similarities, and differences. The analysis highlighted their different structures (eg, bowel movement on the front page or second or third page) or different wordings (eg, movement or mobility). The analysis and follow-up meetings with care staff revealed that this makes it challenging to work with CTRs effectively, as some of the most important fields are located at the end of the report. The meetings also revealed that the CTRs from the hospital are typically shorter (≤8 pages) and hold more structured information (checkboxes) than free-text fields. In comparison, CTRs from the care facilities are usually longer (≤20 pages) and include more free-text fields. The front pages of each analyzed CTR are shown in Figure 3, showcasing their different structure.

In the next step, a semantic analysis, including the mapping of CTRs to the new CIO-CTR standard, was conducted. This was done by assigning parts of the CTR to the data structure of the CIO-CTR (Figure 4). The green box represents a CIO-CTR resource, the white box represents the specification of the resource, and the red box represents the information of the CTR.

Throughout the process, it was realized that the mapping often cannot be done straightforwardly. There were some entities (eg, diagnosed diseases, deafness, aphasia, and limited vision) that could not be assigned to a single field in the CIO-CTR. This was mainly because some of the resources of the CIO-CTR format were too similar to each other. Most of the issues with overlapping assignability were resolved by further study of the CIO-CTR standard and discussion with the research team. For uncertain cases, meetings with mio42 GmbH (originator of the format) were held. Decisions regarding mappings were then based on their feedback. Nevertheless, in some cases, an assignment was still not possible. There was no resource element that provided information about whether the patient or resident had been transferred within a facility (internal; transmission, eg, within a hospital from one to another department) or outside (external; transmission from another facility).

Furthermore, some fields in the CIO-CTR are implemented as free-text fields, which makes unambiguous, error-free mapping difficult.

An excerpt of the mapping of site-specific CTRs of 2 facilities to the new standard CIO-CTR is shown subsequently. Mapping 1 focuses on the CTR of the UHA (Figure 5), and mapping 2 focuses on the CTR of 1 care facility (Figure 6). The visuals illustrate the overall complexity and difficulty of mapping each entity correctly. In mapping 1, it was possible to assign 147 (99.3%) of the 148 information objects from CTR to the CIO-CTR; in mapping 2, it was only possible to map 114 (91.2%) of the 125 information objects.

This raises the question of what should be done with the information that could not be mapped. One possibility would be to add it to the free-text fields.

Figure 3. First page of care transition records (CTRs) from one hospital (1) and 3 care facilities (2-4).

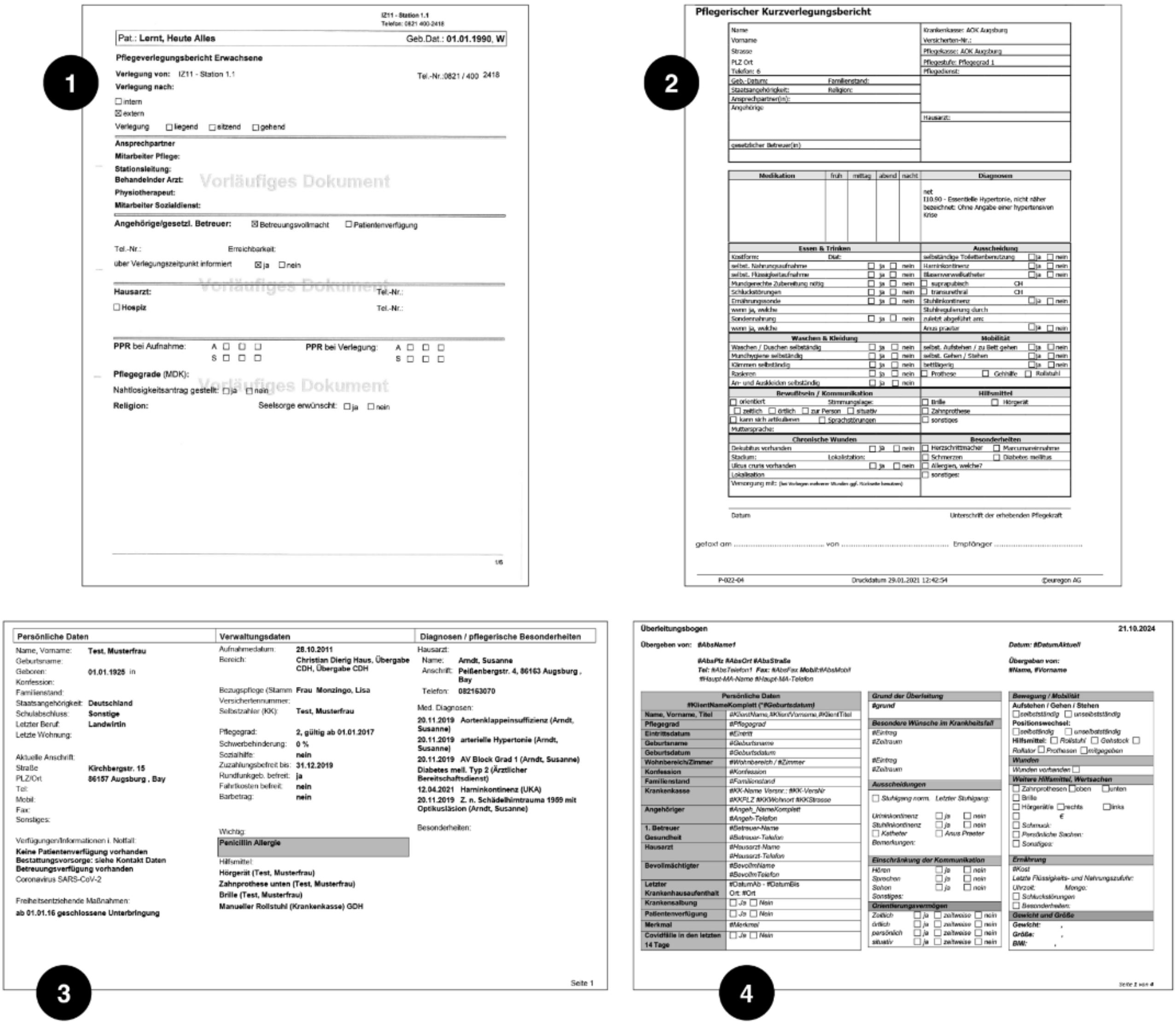


Figure 4. PDF care transition record (CTR) on the right with a free-text field (A, red rectangle) mapped to corresponding resources on the left (B, black and green rectangle, C).

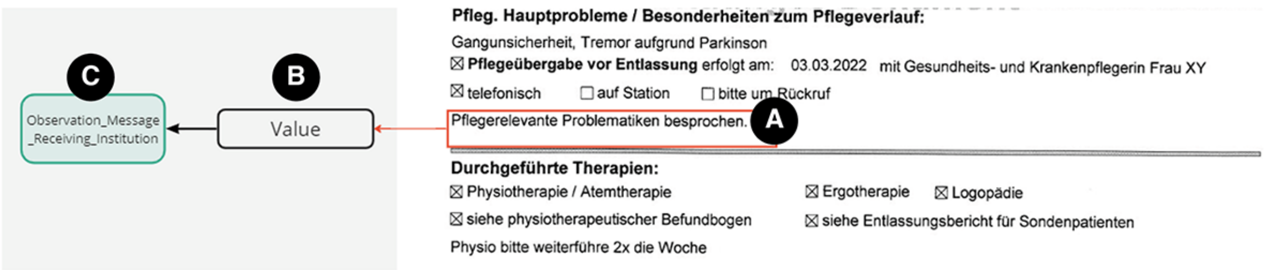


Figure 5. Mapping 1: excerpt of the mapping of a care transition record (CTR) of the University Hospital Augsburg to the care information object (CIO) CTR standard (1 of the 6 pages). The X shows that one piece of information could not be mapped.

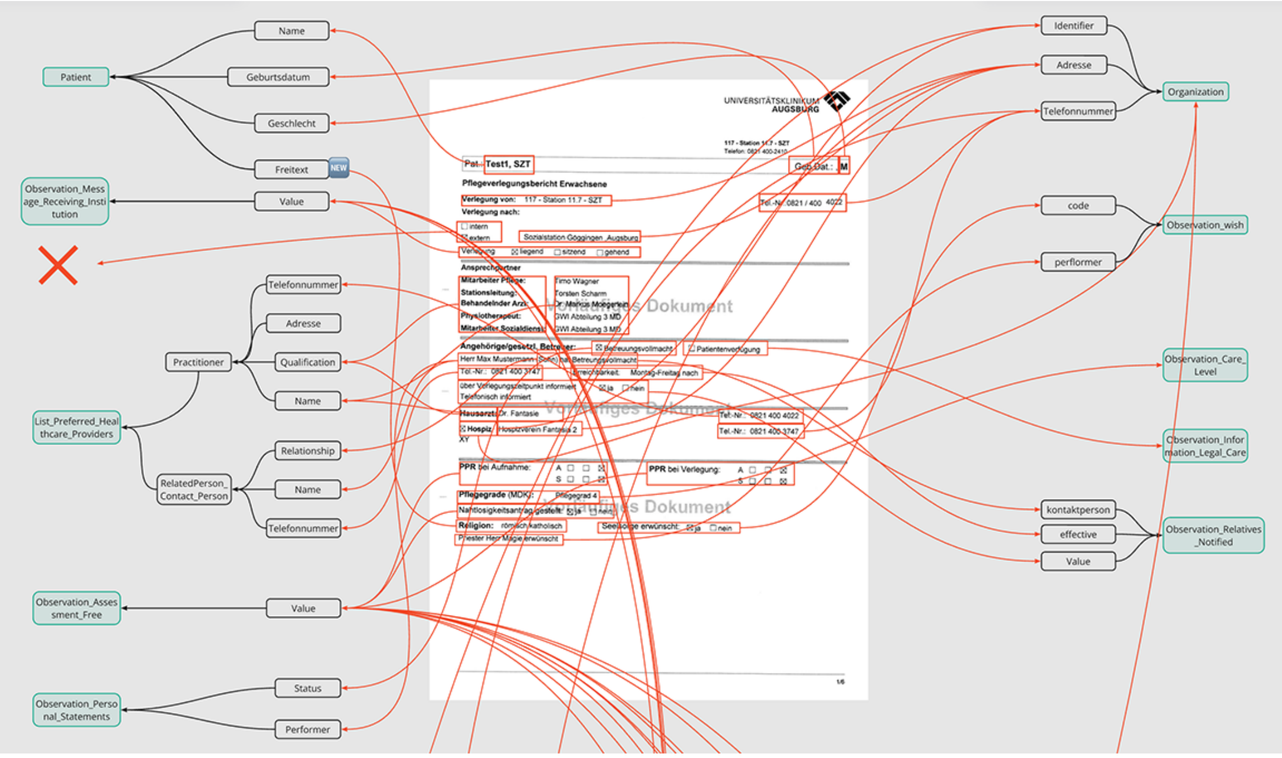
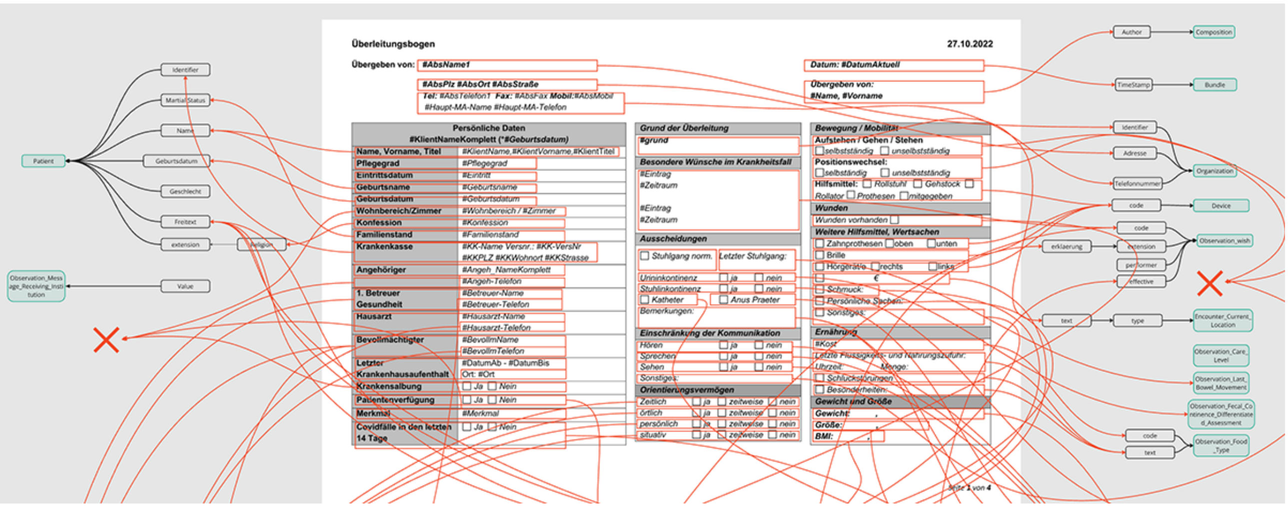


Figure 6. Mapping 2: excerpt of the mapping of a care transition record (CTR) of a care facility to the Pflegerisches Informationsobjekt (PIO) standard (1 of the 4 pages). The X shows that two pieces of information could not be mapped.



Frequency Analysis of CTR Entities at UHA

The occurrence of individual data entities in 204 CTRs of UHA and 54 CTRs from care facilities was analyzed to find out their frequencies. As comparable field entries are needed for processing, only the CTRs of UHA were used for subsequent processing, as this dataset was the biggest.

An entity is understood as a single piece of information represented in the CTR by its input field.

On the basis of these results, a percentage for each entity was computed (entity is filled or not filled), and a frequency range was created (commonly used, occasionally used, and rarely used). These ranges estimate the frequency of entities in the nursing transition process and are shown as follows: (1) 100% to 50%: commonly used entities, (2) 49% to 25%: occasionally used entities, and (3) 24% to 0%: rarely used entities

The results of each entity were presented to care staff (n=2) at UHA who are involved in the CTR process for discussion. An extract of the results is presented in Table 5. It is important to note that the frequency analysis was limited to data that did not include personal information about patients (eg, date of birth, primary care physician, contact options, and religious affiliation), as the UHA anonymized the CTRs before further processing. However, during the discussion, the nursing staff stated that all personal data could be classified as very relevant.

Table 5. Extract from the frequency analysis from University Hospital Augsburg care transition records (N=204).

	Frequency, n (%)
Very relevant (100%-50%)	
Ability of self-body care	201 (98.5)
Orientation ability	198 (97.1)
Dressing	197 (96.6)
Medication: reference to physician’s letter	195 (95.6)
State of consciousness	190 (93.1)
Nutrition	188 (92.2)
Mobility	179 (87.7)
Presence of pain	177 (86.8)
Main diagnosis	167 (81.9)
Last bowel movement	149 (73.0)
Items brought along (suitcase)	121 (59.3)
The degree of care	115 (56.4)
Relevant (49%-25%)	
Nursing-relevant secondary diagnoses	69 (33.8)
The location of the pain	68 (33.3)
The special features of the care process	58 (28.4)
Less to not relevant (24%-0%)	
Medication: reference to a medication plan	23 (11.3)
Free-text field about pain	19 (9.3)
Seamless request (yes or no)	17 (8.3)
Pastoral care requested (yes or no)	3 (1.5)
Aids ordered and their retailers	1 (0.5)
Items brought along	
Valuables	23 (11.3)
Insurance card	22 (10.8)
Identification	5 (2.5)
Patient passport	0 (0.0)

Although the information about the main diagnosis (167/204, 82.3%), state of consciousness (190/204, 93%), and nutrition (188/204, 92%) occurs with high frequency in the dataset, their placement in the paper-based CTR is inadequate, as they appear relatively late in the document.

Another finding is that bowel movement is rated as an essential piece of information (149/204, 73%), but 55/204 (27%) do not include it in the CTR.

Medication information was also expected to be present more frequently; however, because this information is usually included in the physician’s letter rather than in the CTR, the occurrence was only 11% (23/204).

Regarding items brought along, many selection possibilities were given in the UHA’s CTR. Valuables (23/204, 11%) and insurance cards (23/204, 11%) had the highest frequency among them. However, no additional information about the individual items could be provided.

User Story Mapping

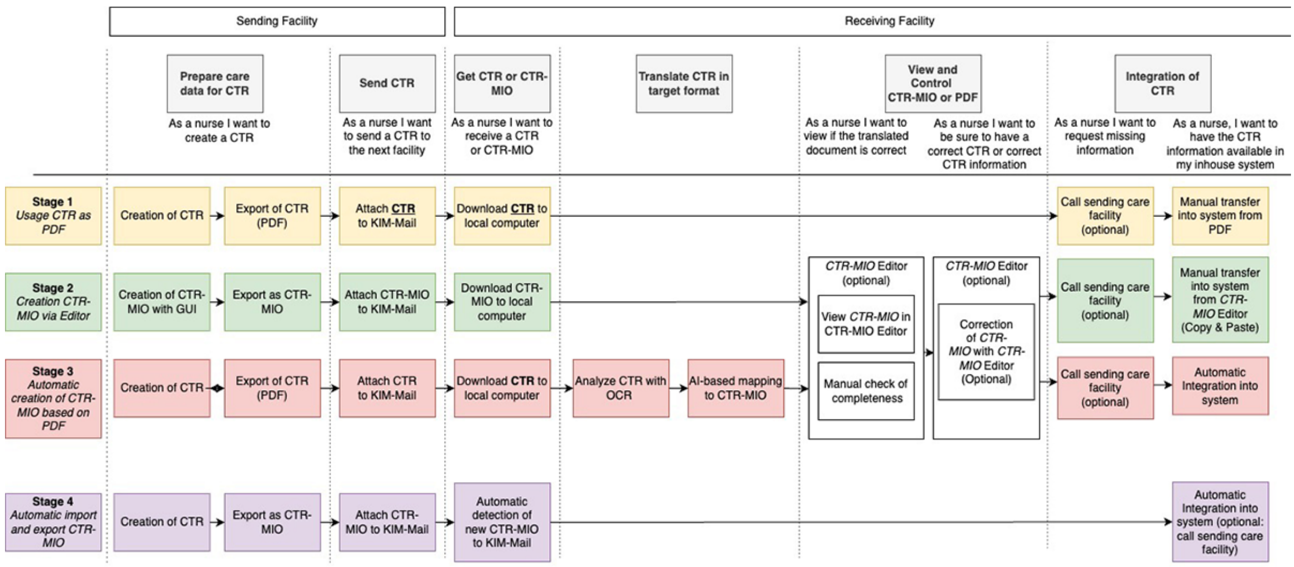
There were 2 workshops conducted, involving a total of 7 participants. These participants were part of the core research project team, bringing diverse expertise from various disciplines: health care (n=2, 29%), computer science (n=3, 43%), design (n=1, 14%), and IT security (n=1, 14%). All 7 (100%) participants attended both workshops, ensuring continuity and consistency in the discussions and decisions. During the first workshop (hybrid, due to COVID-19 restrictions), participants used both physical materials (paper and whiteboards) and digital tools (Zoom [Zoom Communications] and chat) to record potential user stories. The process involved writing down ideas and then engaging in a collaborative card-sorting exercise to discuss and prioritize these stories. A whiteboard was used to document the structured user journey, which was shared with online participants via camera, ensuring everyone had equal access to the visual information.

Between the first and second workshops, participants had approximately 2 weeks to reflect on the identified user stories and their potential impact on the development process. This period allowed the team to refine their understanding and prepare for the next stage of discussion, which focused on a stepwise implementation plan (release planning).

The second workshop was conducted entirely online using the tools Zoom and Miro. Miro is an online collaborative platform developed by RealtimeBoard Inc. The participants had time to share their reflections on the previous work, discuss it, and refine their understanding of the user journey. Afterward, the participants focused on creating release stages to guide the upcoming development. They collaboratively designed the user story map due to the workshops, which was continuously refined throughout the project. The final result, a road map proposal, can be seen in Figure 7. The Backbone section describes the backbone and the proposed solution's release stages.

In Figure 7, four release stages are shown on the left and divided into 4 colors to provide a better division throughout the user stories. Some story cards do not have color, as they apply to multiple release stages. The subsequent sections describe the use dimensions, backbone, release stages, and implementation scenarios.

Figure 7. Final user story mapping with 4 release stages (on the left). Colors are used to better distinguish between the stages. Some story cards do not have color, as they apply to multiple stages. AI: artificial intelligence; CIO: care information object; CTR: care transition record; KIM: Kommunikation im Medizinwesen.



Backbone

The horizontal axis of the user story map shows the main activities that have to be performed sequentially to achieve care data exchange between facilities. These activities are referred to as epics and are listed in the top row of Figure 7. The user stories concretize the epics. One activity (translating CTR in target format) has to be performed only in an intermediate release stage, and it becomes obsolete as soon as all facilities use the standardized target format.

Release Stages

A brief overview of all 4 release stages can be seen in Figure 7. A description of each stage is given in the subsequent sections.

Release Stage 1: Use of CTR in PDF Format

The first release requires the least implementation effort but already meets one basic requirement: timely, digital transfer via the TI. The functionality is limited to conventional CTRs, typically in PDF format. This release requires both the sending and receiving care facilities to be connected to the TI. As usual, the sending facility creates a CTR in its facility-specific layout and transmits it using the KIM service. The receiving facility can then retrieve the CTR from its KIM mailbox.

Release Stage 2: Creation of CIO-CTR via Editor

At the beginning of a transition process to the new standardized CIO-CTR format, few or no in-house systems will support the new standard. To remain independent from software manufacturers, a dedicated software module that can create, read, and edit CTRs in the new format ("CIO-CTR editor") would be beneficial. Nurses could use this editor to create CIO-CTRs and send them to the receiving institution via KIM. Particular emphasis should be placed on the user-centered design of the interface, particularly regarding the structure of the input options and how information is compiled. This could serve as a blueprint for later implementation in the proprietary software systems.

Release Stage 3: Automatic Creation of CIO-CTR Based on PDF

Another, more complex way to create and transfer a CTR in the new standard format is to transform the conventional, proprietary CTR using an automatic artificial intelligence (AI)–based tool. On the basis of the previous analysis of CTRs, it can be assumed that most CTR data will be unstructured and provided in PDF format. A transformer service could analyze this structure using AI and extract text sections with an optical character recognition module. The extracted content is then mapped to the CIO-CTR format. This approach would be relevant if a receiving facility is already capable of processing CIO-CTRs but receives a nonstandardized CTR via KIM. With a transformer, the new CTR-CIO format can be generated and imported with little extra effort.

Release Stage 4: Automatic Export and Import of CIO-CTR

In this final stage, the care staff can create a CTR in the in-house system, export it as a CIO-CTR, and transfer it via the TI. After receiving the CIO-CTR, the receiving facility can then integrate it directly into their in-house system. The benefit is that neither a transformer service nor a separate editor would be needed, resulting in the least effort for the care staff. This requires the software manufacturers of the various care and medical information systems to fully support the new CIO-CTR format; however, it is unclear when this will happen.

Discussion

Principal Findings

Despite years of efforts toward digitalization in health care in Germany, our research shows that the creation and transmission of CTRs remain highly time-consuming, averaging 34.67 (SD 10.78) minutes at hospitals and 44.6 (SD 20.5) minutes in care facilities (findings from observations).

Semantic Interoperability of CTRs Between Institutions

As health care systems transition toward digital formats, it becomes increasingly important to enable different institutions to exchange, understand, and use the transmitted data seamlessly. From the perspective of nursing science, discharge management has long been recognized as a crucial aspect of patient care. Efficient discharge processes ensure that patients receive continued care, reduce readmission rates, and improve overall patient outcomes. The CIO-CTR standard, introduced in December 2022, marks a significant step toward a fully digital exchange of CTR data. However, our study reveals that this progress has been hampered by a lack of widespread implementation and resistance. Because the CIO-CTR is not legally binding and the necessary updates are resource intensive for software manufacturers, they prefer to concentrate on more urgent issues. Thus, we propose an iterative, stepwise implementation approach that could gradually improve the situation.

Iterative Implementation Approach

The user story map with the resulting release stages offers a step-by-step approach toward a seamless digital solution. As the overall issue is complex, changes cannot be expected simultaneously at all ends. A quick, early solution is the mere digital transfer of CTRs in existing, proprietary formats via a digital infrastructure (stage 1). For this, the institutions only have to be connected to the health data network (TI), as they are obliged by law in Germany by July 1, 2025 (according to §341 (8) SGB V [46]) and a KIM account is set up. Sending CTRs in the institutions' traditional formats does not require them to have updated software that can read or export the new CIO-CTR format. At this stage, the time-consuming manual data transfer into the in-house systems is still required. The goal is for all software systems in all institutions to directly import and export CTRs in the new format, and for all the information to be integrated automatically into in-house systems (stage 4). During a transition time, when only some of the systems can process CTRs in the new format, certain incompatibilities will occur, which we want to address with interim solutions: the CIO-CTR viewer or editor (stage 2) and the CTR-transformer (stage 3).

For stage 1 (data transfer via TI), we accompany and assist our cooperating partner institutions in installing the necessary infrastructure to connect to the TI. In this regard, we plan to offer experience reports, which could lower the entry hurdle, particularly for care facilities.

For stage 2, we are developing an open-source software where CIO-CTRs can be created, viewed, and edited. This has several benefits: (1) developing an editor with a concrete suggestion for a user interface visualizing the CIO-CTR standard provides a figurative basis for discussion between developers, care professionals, and regulatory institutions; (2) bridging the gap for continuous digital transfer if not all institutions support the new digital standard; and (3) serving as a blueprint for software manufacturers who want to implement the new CIO-CTR.

Stage 3 introduces an automated process to convert CTRs from proprietary formats (eg, scanned PDFs) into the CIO-CTR format, using AI-based mapping. This solution is applicable when an institution that can process CIO-CTRs but receives a nonstandardized CTR. Of course, this automatic transfer would have to be reliable, and creating such a component would be complex, as many different proprietary formats exist, and as seen in the semantic mapping, a direct transfer is not possible in all cases.

Stage 4 represents the most desirable solution. Nurses would be able to work with an improved process without manual transfer of CTR data, potentially leading to a minimization of disruptions. The primary responsibility for implementing the CIO-CTR falls on system manufacturers. To facilitate this transition, the manufacturers could actively be supported by providing a reference implementation for the new standard, for example, conducting workshops and organizing related events. This collaborative effort would support a smooth and efficient integration of the CIO-CTR into existing systems while minimizing the burden on health care providers.

Contributions

Our research used an iterative, user-centered methodological approach to develop a road map that helps overcome the current challenges in the CTR transmission process in Germany. This road map offers a practical, phased approach toward digital solutions, particularly valuable in settings where full-scale adoption of digital standards is not yet feasible. It provides health care providers with a flexible pathway to transition toward digital care processes without requiring immediate, costly system changes. This road map is more than the mere definition of a new format; it supports gradual digital integration.

Future Implications and Work

If the adoption of digital standards remains voluntary and lacks regulatory support, the duration required to establish a standard data format is likely to be prolonged. Without more vigorous regulatory enforcement and widespread buy-in from all stakeholders, the vision of seamless care transitions may remain out of reach. Therefore, future efforts must focus not only on technological solutions but also on fostering collaboration between regulators, software providers, and health care institutions to ensure the long-term success of health care digitalization.

Our current solution uses KIM as a means of transport within the TI. As soon as the ePA is more widely adopted in Germany, an exchange of the CTR via this means might be preferable over KIM. Future work has to investigate this further.

Limitations

The COVID-19 pandemic posed significant challenges for data collection, particularly in gaining access to cooperative facilities. The necessary planning and multiple postponements due to visitor restrictions limited our ability to observe the complete care transition process.

The pandemic may have also introduced a selection bias in our web-based questionnaire. Nurses facing higher technical barriers or those under significant stress due to pandemic-related demands may have been less likely to participate, which could skew the findings toward participants who were more technically adept or had fewer pandemic-related pressures.

Furthermore, the sample size of the web-based questionnaire (n=33 usable datasets), while offering valuable qualitative insights, limits the generalizability of our findings.

Our study used a qualitative research approach to gain in-depth, context-specific insights. The combination of field observations, contextual inquiries, and questionnaire data from hospital and care facilities provided a rich understanding of the practical barriers and opportunities in transitioning to digital CTR processes. The relatively small sample sizes for field observations (n=6) and contextual inquiries (n=5) were sufficient for this research's detailed, exploratory nature but could limit the robustness of the conclusions.

Conclusions

A future solution should simplify the overall CTR transmission process by minimizing the manual transfers into the in-house systems, standardizing the CTR, and providing a secure digital transfer. Doing so could positively impact the overall care process and patient experience. With our suggestion for a stepwise solution, we attempt to make the complex task feasible, ultimately supporting care staff with their daily activities and processes.

Acknowledgments

This research is part of the project CARE REGIO, funded by the Bavarian State Ministry of Health, Care and Prevention. The authors would like to thank their partners from 3 local health care institutions for participating in the stakeholder workshop, allowing them to conduct field observations and contextual inquiries at their sites, and discussing their care transition records. The authors would also like to thank the University of Applied Sciences in Neu-Ulm (Institut DigiHealth) for its support regarding legal issues and ethics. Writing artificial intelligence (AI) assistants ChatGPT-3.5, 4.0 (OpenAI); Grammarly (Grammarly Inc); and DeepL (DeepL SE) were used for language improvement and text shortening. These or any other AI systems created no scientific content in this manuscript.

Authors' Contributions

EVM, AM, CR, and AT were involved in the conceptualization. EVM, SB, LK, and MR were involved in formal analysis. EVM, SB, MR, LK, LD, ST, VW, CR, and AT were involved in the investigation. EVM, SB, MR, LK, LD, ST, and AT were involved in data curation. EVM, SB, MR, LK, and AT were involved in writing the original draft. EVM, SB, VW, and AT were involved in reviewing and editing the draft. EVM and LK were involved in visualization. AT was involved in supervising the study.

Conflicts of Interest

None declared.

References

1. Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. J Am Med Inform Assoc 2020 Jun 01;27(6):957-962 [FREE Full text] [doi: 10.1093/jamia/ocaa067] [Medline: 32311034]
2. Lupton D. The digitally engaged patient: self-monitoring and self-care in the digital health era. Soc Theory Health 2013 Jun 19;11(3):256-270. [doi: 10.1057/sth.2013.10]
3. Gerke S, Stern AD, Minssen T. Germany's digital health reforms in the COVID-19 era: lessons and opportunities for other countries. NPJ Digit Med 2020 Jul 10;3(1):94 [FREE Full text] [doi: 10.1038/s41746-020-0306-7] [Medline: 32685700]
4. Implementing the European health data space across Europe. ThinkTank. URL: https://eithealth.eu/wp-content/uploads/2024/04/EIT_Health_ThinkTank_Implementing_the_EHDS_across_Europe_23.04.24.pdf [accessed 2024-04-29]

5. Sauermann S, Herzberg J, Burkert S, Habetha S. DiGA - a chance for the German healthcare system. J Eur CME 2022 Dec 23;11(1):2014047 [FREE Full text] [doi: 10.1080/21614083.2021.2014047] [Medline: 34992948]
6. Baltaxe E, Czypionka T, Kraus M, Reiss M, Askildsen JE, Grenkovic R, et al. Digital health transformation of integrated care in Europe: overarching analysis of 17 integrated care programs. J Med Internet Res 2019 Sep 26;21(9):e14956 [FREE Full text] [doi: 10.2196/14956] [Medline: 31573914]
7. Nohl-Deryk P, Brinkmann J, Gerlach F, Schreyögg J, Achelrod D. Hürden bei der Digitalisierung der Medizin in Deutschland – eine Expertenbefragung. Gesundheitswesen 2018 Nov 04;80(11):939-945. [doi: 10.1055/s-0043-121010] [Medline: 29301149]
8. Hansen A, Herrmann M, Ehlers JP, Mondritzki T, Hensel KO, Truebel H, et al. Perception of the progressing digitization and transformation of the German health care system among experts and the public: mixed methods study. JMIR Public Health Surveill 2019 Oct 28;5(4):e14689 [FREE Full text] [doi: 10.2196/14689] [Medline: 31661082]
9. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (general data protection regulation) (text with EEA relevance). Official Journal of European Union. URL: https://op.europa.eu/en/publication-detail/-/publication/3e485e15-11bd-11e6-ba9a-01aa75ed71a1/language-en [accessed 2024-04-29]
10. Pohlmann S, Kunz A, Ose D, Winkler EC, Brandner A, Poss-Doering R, et al. Digitalizing health services by implementing a personal electronic health record in Germany: qualitative analysis of fundamental prerequisites from the perspective of selected experts. J Med Internet Res 2020 Jan 29;22(1):e15102 [FREE Full text] [doi: 10.2196/15102] [Medline: 32012060]
11. Die Telematikinfrastruktur. gematik GmbH. URL: https://www.gematik.de/telematikinfrastruktur [accessed 2024-10-20]
12. Hoerbst A, Ammenwerth E. Electronic health records. Methods Inf Med 2018 Jan 17;49(04):320-336. [doi: 10.3414/me10-01-0038]
13. Kataria S, Ravindran V. Electronic Health Records: A Critical Appraisal of Strengths and Limitations. Journal of the Royal College of Physicians of Edinburgh 2020 Sep 01;50(3):262-268. [doi: 10.4997/jrcpe.2020.309]
14. Gesundheitsinfos, online verfügbar. Elektronisches Patientendossier. URL: https://www.patientendossier.ch/fachpersonen [accessed 2024-10-20]
15. Schweizerische Eidgenossenschaft, 81 Gesundheit, 810 Medizin und Menschenwürde, 816 Patientendossier. Fedlex, Die Publikationsplattform des Bundesrechts. URL: https://www.fedlex.admin.ch/de/cc/internal-law/81#816 [accessed 2024-10-20]
16. Elektronische Gesundheitsakte. Bundesministerium für Soziales, Gesundheitswesen, Pflege und Konsumentenschutz. URL: https://www.gesundheit.gv.at/gesundheitsleistungen/elga.html [accessed 2024-10-20]
17. Jormanainen V, Lindgren M, Keskimäki I, Kaila M. Use of My Kanta in Finland 2010-2022. Stud Health Technol Inform 2023 Jun 29;305:448-451. [doi: 10.3233/SHTI230528] [Medline: 37387062]
18. MyKanta. Kanta Services, The Social Insurance Institution of Finland. URL: https://www.kanta.fi/en/mykanta [accessed 2024-10-20]
19. Mon Espace Santé – Mon Espace Santé is a personal space where users manage their health data. Ministerial eHealth Delegation and Ministry of Health (FR MoH). URL: https://gnius.esante.gouv.fr/en/regulations/regulation-profiles/mon-espace-sante [accessed 2024-01-24]
20. The electronic patient record. Federal Ministry of Health. URL: https://gesund.bund.de/en/topics/electronic-health-record-epa [accessed 2024-10-20]
21. Möglichkeiten und Herausforderungen der elektronischen Patientenakte. aerzteblatt. URL: https://www.aerzteblatt.de/nachrichten/150563/Moeglichkeiten-und-Herausforderungen-der-elektronischen-Patientenakte [accessed 2024-10-20]
22. Zeitz G. Elektronische Patientenakte: Droht Chaos beim Start? Hausärztinnen- und Hausärzteverband Hessen. URL: https://www.hausaerzte-hessen.de/aktuelles/news/801-interview-epa [accessed 2024-10-20]
23. Kommunikation im Medizinwesen (KIM). gematik GmbH. URL: https://www.gematik.de/anwendungen/kim [accessed 2024-10-20]
24. TI-Modellregion Franken. Medical Valley EMN e.V. URL: https://gesundheitsnetz-franken.de/ti-modellregion-franken/ [accessed 2024-10-20]
25. TIMO TI-Modellregion Hamburg and Umland. GfgA – Gesellschaft für geschäftliche Angelegenheiten UG. URL: https://timo-hamburg-umland.de/ [accessed 2024-10-20]
26. Projekte und Publikationen: ePflegebericht und eWundbericht. Netzwerk Versorgungskontinuität in der Region Osnabrück e.V. URL: https://www.hs-osnabrueck.de/netzwerk-versorgungskontinuitaet/projekte-und-publikationen/#c10209075 [accessed 2024-10-20]
27. Giehoff C, Hübner UH. Der elektronische Pflegebericht des "Netzwerks Versorgungskontinuität in der Region Osnabrück" - Evaluationsergebnisse und ihre Konsequenzen. ResearchGate. URL: https://tinyurl.com/2zd8h6mk [accessed 2024-04-29]
28. Hübner U, Flemming D, Heitmann KU, Oemig F, Thun S, Dickerson A, et al. The need for standardised documents in continuity of care: results of standardising the eNursing summary. Stud Health Technol Inform 2010;160(Pt 2):1169-1173. [Medline: 20841868]
29. Flemming D, Giehoff C, Hübner U. Entwicklung eines Standards für den elektronischen Pflegebericht auf Basis der HL7 CDA Release 2. In: Proceedings of the 2008 Annual meeting of the German Society for Medical Informatics, Biometry

and Epidemiology. 2008 Presented at: GMDS '08; September 15-18, 2008; Düsseldorf, Germany p. 682 URL: https://www.egms.de/static/en/meetings/gmds2008/08gmds182.shtml
30. Flemming D, Hübner K, Heitmann F, Thun S. Implementierungsleitfaden „ePflegebericht "auf Basis der HL7 Clinical Document Architecture Release 2 für das deutsche Gesundheitswesen-draft v06. HL7.de. URL: http://wiki.hl7.de/index.php/IG:Pflegeberich [accessed 2024-10-20]
31. Flemming D, Schulte G, Hübner U. Evaluation des Deutscher HL7 CDA basierten elektronischen pflegeberichts. In: Proceedings of the 2013 Conference on Health Informatics meets eHealth. 2023 Presented at: eHealth '13; May 23-24, 2013; Vienna, Austria p. 1-7 URL: https://www.hs-osnabrueck.de/fileadmin/HSOS/Homepages/KeGL/Artikel_Evaluation_des_deutschen_HL7_CDA_basierten_elektronischen_Pflegeberichts.pdf
32. MIO - Medizinische Informationsobjekte. mio42 GmbH. URL: https://mio.kbv.de/site/mio [accessed 2024-03-20]
33. FHIR-Spezifikation. mio42 GmbH. URL: https://mio.kbv.de/pages/viewpage.action?pageId=273318266 [accessed 2024-10-20]
34. Pio-Festlegung: Überleitungsbogen. Kassenärztliche Bundesvereinigung and GKV-Spitzenverband. URL: https://www.kbv.de/temp/Anlage_1_PIO-Festlegung.pdf [accessed 2024-04-29]
35. Burr H, Berthelsen H, Moncada S, Nübling M, Dupret E, Demiral Y, international COPSOQ Network. The third version of the Copenhagen Psychosocial Questionnaire. Saf Health Work 2019 Dec;10(4):482-503 [FREE Full text] [doi: 10.1016/j.shaw.2019.10.002] [Medline: 31890332]
36. Angrosino M. Doing Ethnographic and Observational Research. Thousand Oaks, CA: Sage Publications; 2007.
37. Beyer H, Holtzblatt K. Contextual Design: Defining Customer-Centered Systems. New York, NY: Morgan Kaufmann Publishers; 1997.
38. Goodman E, Kuniavsky M, Moed A. Observing the User Experience. 2nd edition. Berlin, Germany: Morgan Kaufmann; 2012.
39. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
40. Ruiz-Fernández D, Marcos-Jorquera D, Gilart-Iglesias V, Vives-Boix V, Ramírez-Navarro J. Empowerment of patients with hypertension through BPM, IoT and remote sensing. Sensors (Basel) 2017 Oct 04;17(10):2273 [FREE Full text] [doi: 10.3390/s17102273] [Medline: 28976940]
41. De Ramón Fernández A, Ruiz Fernández D, Sabuco García Y. Business process management for optimizing clinical processes: a systematic literature review. Health Informatics J 2020 Jun 04;26(2):1305-1320 [FREE Full text] [doi: 10.1177/1460458219877092] [Medline: 31581880]
42. Jurafsky D, Martin JH. Speech and language processing: an introduction to natural language processing, computational linguistics, and speech recognition. University of California San Diego. URL: https://tinyurl.com/z4h5k99v [accessed 2024-02-20]
43. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed representations of words and phrases and their compositionality. arXiv Preprint posted online October 16, 2013 [FREE Full text]
44. Mayring P. Qualitative Inhaltsanalyse: Grundlagen und Techniken. 13th eddition. Berlin, Germany: Beltz; 2022.
45. Patton J. User Story Mapping: Discover the Whole Story, Build the Right Product. Berlin, Germany: O'Reilly Media; 2014.
46. Sozialgesetzbuch (SGB) Fünftes Buch (V) - Gesetzliche Krankenversicherung - (Artikel 1 des Gesetzes v. 20. Dezember 1988, BGBl. I S. 2477). Bundesministerium der Justiz. URL: https://www.gesetze-im-internet.de/sgb_5/BJNR024820988.html [accessed 2024-04-29]

Abbreviations

AI: artificial intelligence
BPMN: business process model and notation
CIO: care information object
CTR: care transition record
ePA: electronic patient record
GDPR: General Data Protection Regulation
HL7: Health Level 7
KIM: Kommunikation im Medizinwesen
TI: telematics infrastructure
UHA: University Hospital Augsburg

Edited by E Borycki; submitted 23.05.24; peer-reviewed by PDD Flemming, RA El Arab; comments to author 22.07.24; revised version received 23.10.24; accepted 12.11.24; published 21.02.25.

Please cite as:
Mess EV, Regner M, Balic S, Kleybolte L, Daufratshofer L, Mahler A, Tilmes S, Werlitz V, Reuter C, Teynor A
Detailed Analysis and Road Map Proposal for Care Transition Records and Their Transmission Process: Mixed Methods Study
JMIR Nursing 2025;8:e60810
URL: https://nursing.jmir.org/2025/1/e60810
doi:10.2196/60810
PMID:39982779

Advancing Clinical Chatbot Validation Using AI-Powered Evaluation With a New 3-Bot Evaluation System: Instrument Validation Study

Seungheon Choo[1]; Suyoung Yoo[2], RN; Kumiko Endo[3], PhD; Bao Truong[3], MD; Meong Hi Son[2,4], MD

[1]Research Institute for Future Medicine, Samsung Medical Center, Seoul, Republic of Korea

[2]Department of Digital Health, Samsung Advanced Institute for Health Sciences and Technology (SAIHST), Sungkyunkwan University, Seoul, Republic of Korea

[3]Med2Lab Inc, San Francisco, CA, United States

[4]Department of Emergency Medicine, Samsung Medical Center, Seoul, Republic of Korea

Corresponding Author:
Meong Hi Son, MD
Department of Digital Health, Samsung Advanced Institute for Health Sciences and Technology (SAIHST), Sungkyunkwan University, Seoul, Republic of Korea

Abstract

Background: The health care sector faces a projected shortfall of 10 million workers by 2030. Artificial intelligence (AI) automation in areas such as patient education and initial therapy screening presents a strategic response to mitigate this shortage and reallocate medical staff to higher-priority tasks. However, current methods of evaluating early-stage health care AI chatbots are highly limited due to safety concerns and the amount of time and effort that goes into evaluating them.

Objective: This study introduces a novel 3-bot method for efficiently testing and validating early-stage AI health care provider chatbots. To extensively test AI provider chatbots without involving real patients or researchers, various AI patient bots and an evaluator bot were developed.

Methods: Provider bots interacted with AI patient bots embodying frustrated, anxious, or depressed personas. An evaluator bot reviewed interaction transcripts based on specific criteria. Human experts then reviewed each interaction transcript, and the evaluator bot's results were compared to human evaluation results to ensure accuracy.

Results: The patient-education bot's evaluations by the AI evaluator and the human evaluator were nearly identical, with minimal variance, limiting the opportunity for further analysis. The screening bot's evaluations also yielded similar results between the AI evaluator and human evaluator. Statistical analysis confirmed the reliability and accuracy of the AI evaluations.

Conclusions: The innovative evaluation method ensures a safe, adaptable, and effective means to test and refine early versions of health care provider chatbots without risking patient safety or investing excessive researcher time and effort. Our patient-education evaluator bots could have benefitted from larger evaluation criteria, as we had extremely similar results from the AI and human evaluators, which could have arisen because of the small number of evaluation criteria. We were limited in the amount of prompting we could input into each bot due to the practical consideration that response time increases with larger and larger prompts. In the future, using techniques such as retrieval augmented generation will allow the system to receive more information and become more specific and accurate in evaluating the chatbots. This evaluation method will allow for rapid testing and validation of health care chatbots to automate basic medical tasks, freeing providers to address more complex tasks.

(JMIR Nursing 2025;8:e63058) doi:10.2196/63058

KEYWORDS

artificial intelligence; patient education; therapy; computer-assisted; computer; understandable; accurate; understandability; automation; chatbots; bots; conversational agents; emotions; emotional; depression; depressive; anxiety; anxious; nervous; nervousness; empathy; empathetic; communication; interactions; frustrated; frustration; relationships

Introduction

Faced with a projected shortfall of 10 million health care workers by 2030 [1], the health care sector urgently requires innovative solutions to sustain patient care and education. Artificial intelligence (AI) automation in low- to mid-level tasks like patient education and initial therapy screening emerges as a strategic response to mitigate this shortage, reallocating medical staff to higher-priority tasks [2,3].

The advent of advanced multimodal large language models (LLMs) such as GPT-4 introduces a paradigm shift, promising scalable, cost-effective chatbot solutions, which are particularly helpful for tasks that require the provider to interact with the patient [4]. GPT-4 and similar models offer a more dynamic, conversational approach, tailoring information to individual patient needs with minimal logistical or financial overhead for health care institutions. This technological evolution promises not only to fill the imminent workforce gap but also to enhance the quality and accessibility of health care services, leveraging AI's capacity for on-demand, personalized patient support [4-7]. It has been reported that LLMs have the cognitive capacity to role-play the character as portrayed in the dialogue prompt [8]. Shao et al [9] showed that GPT-3.5 can be used to score the believability of LLM role-playing. Finally, Yang et al [10] pointed to the high potential that medical chatbots have in clinical settings, while Gilbert et al [11] warned of the need to extensively test health care chatbots.

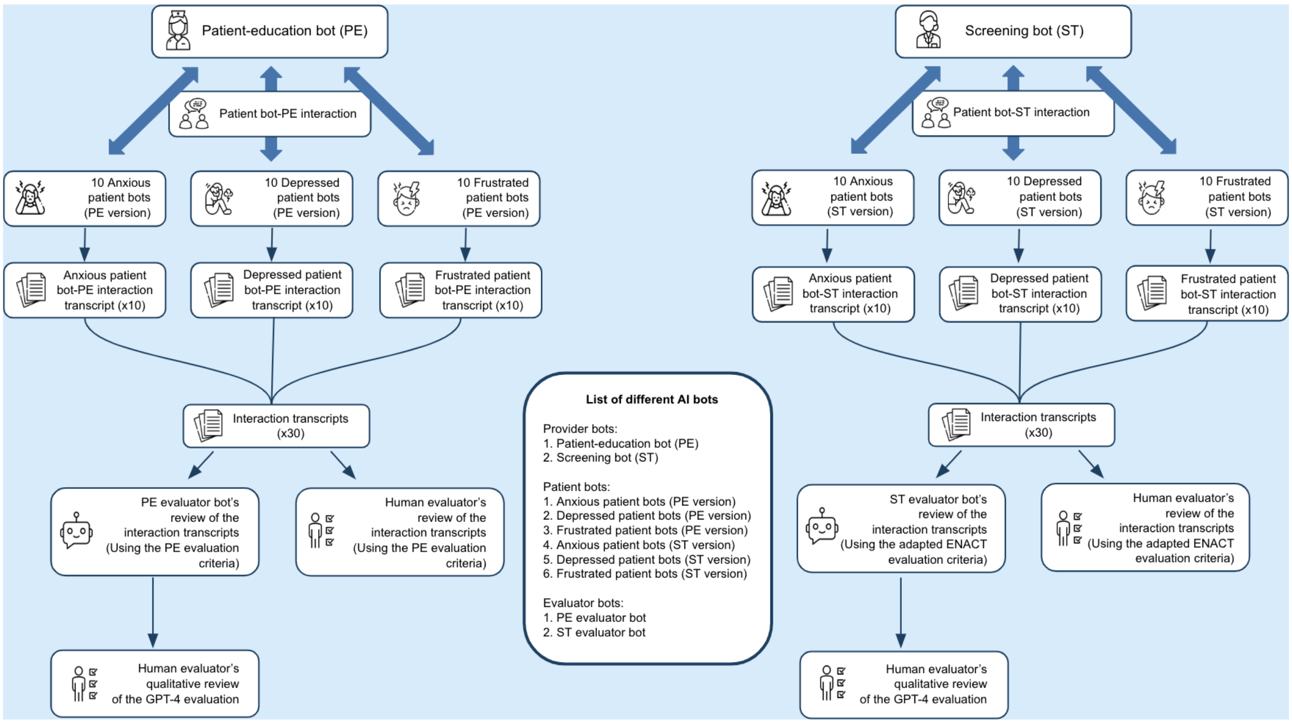
However, current methods of creating and evaluating early-stage health care bots face steep development costs due to the high level of human involvement in each phase of the development process. In this study, we present a novel, bot-driven method of developing, testing, and evaluating automated health care chatbots. At the center of this strategy is the use of the LLM as an "evaluator agent" to iteratively review and provide feedback on the dialog between the health care bot being evaluated and a set of "digitally simulated patients" also role-played by the LLM. This approach provides a fully automated system that will not only reduce the amount of time and effort required to develop the chatbots but also provide a feasible way to continuously monitor the performances of health care chatbots in different clinical settings.

Methods

Study Design

This study introduces a novel bot-driven method to evaluate the abilities of LLMs in health care tasks. In this approach, LLMs were configured to perform as a patient-education bot, a pretherapy screening bot, patient bots, and evaluator bots. The patient bots simulated distinct emotional personas—depressed, anxious, and frustrated—to test the adaptability and competency of the provider bots. The evaluator bots assessed the interactions based on predefined criteria. Results from the AI evaluations were cross-referenced with human expert reviews for accuracy and reliability (Figure 1).

Figure 1. Interaction and evaluation methodology for the patient-education bot and the initial screening bot. Includes a list of all the different bots used. AI: artificial intelligence; ENACT: Enhancing Assessment of Common Therapeutic Factors.



Setup

To demonstrate the system, 2 AI provider bots were developed using GPT-4 in collaboration with an experienced oncology nurse and a licensed cognitive behavior counselor. One provider bot emulated a patient-education nurse, delivering medical information with clarity and empathy. The second bot acted as a mental health therapist, modeled on acceptance and commitment therapy and mindfulness practices, to provide nonpharmacological mental health support.

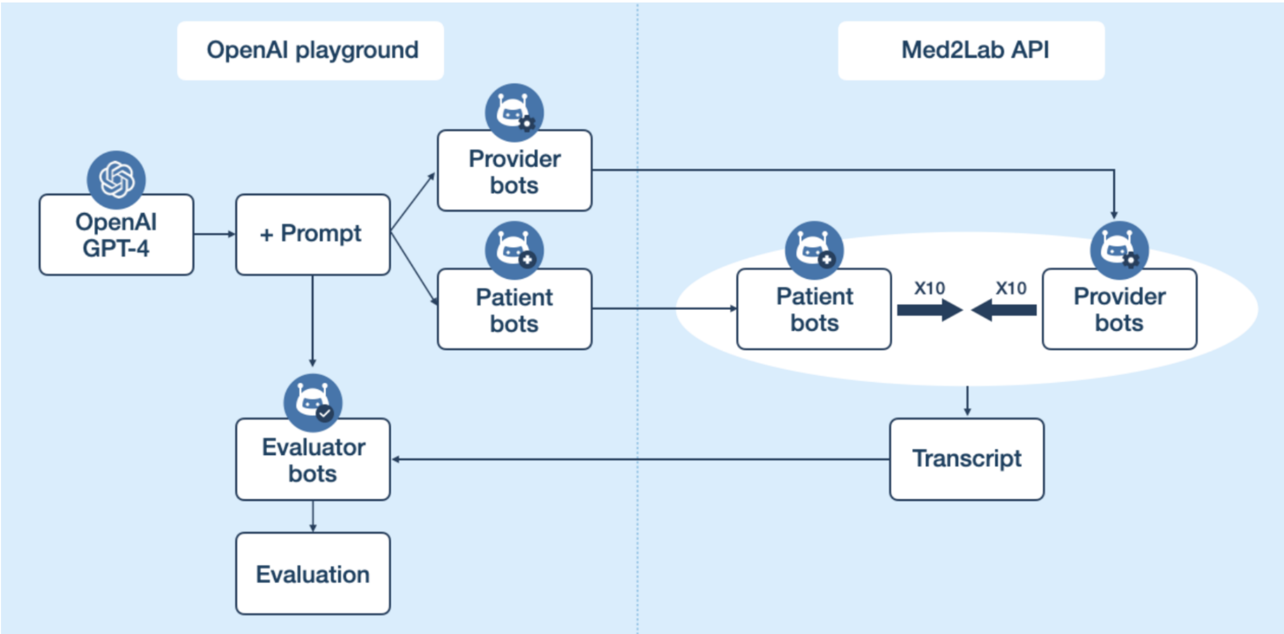
AI patient bots, also developed using GPT-4, were programmed to represent 40-year-old male patients with lung cancer with 1 of 3 emotional personas: depressed, anxious, and frustrated. In total, 30 patient bots (10 per persona) were created, with each provider bot engaging in 30 interactions. The patient bots' responses were unique due to GPT-4's stochastic generation processes, even with consistent prompts.

Evaluator bots were created for each provider bot to assess their performance based on predefined criteria, offering scores and qualitative feedback. These AI-generated evaluations were subsequently reviewed by human experts in relevant fields to ensure validity.

Once the evaluator bots reviewed each provider-patient transcript, human experts in each field reviewed the transcripts, scored the interaction using the same criteria as the GPT-4 bots, commented on the provider's overall performance, and then reviewed the evaluator bot's assessment.

The patient-education bot was reviewed by the same pediatric hematology-oncology nurse who helped create the patient-education approach, while the pretherapy screening bot was reviewed by a PhD in IT psychology as well as by the cognitive behavior counselor (Figure 2).

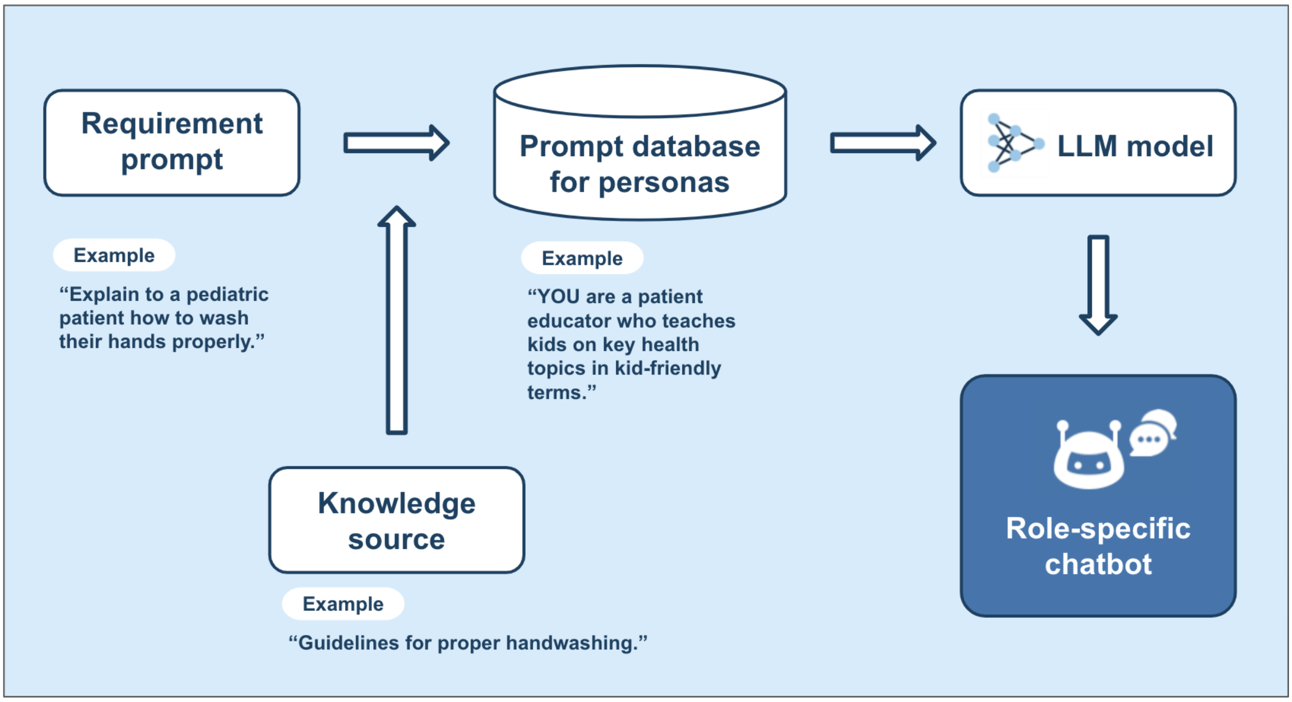
Figure 2. Outline of the bot-bot interactions and evaluations. API: application programming interface.



Prompting

To ensure that the bots adhered strictly to their designated roles while mitigating unwanted behaviors, explicit, role-specific instructions were incorporated into their prompts. This design approach balanced general AI capabilities with task-specific requirements, ensuring consistent and contextually appropriate responses (Figure 3).

Figure 3. Chatbot architecture. LLM: large language model.



Patient-Education Bot Prompting

The patient-education bot was designed to emulate the role of a "patient-education nurse" tasked with educating patients with cancer about medical vocabulary, procedures, and treatment options. The bot was prompted with detailed instructions, emphasizing clarity, empathic expressions in communication, and a patient-centered approach. The following guidelines were incorporated into its prompt: (1) adopt a teaching role tailored to patients with limited medical knowledge; (2) provide accurate, comprehensive explanations of medical terms and procedures in simple, relatable language; (3) exhibit empathy and warmth while refraining from making medical recommendations outside the scope of a patient-education nurse; and (4) ensure consistency in tone and responsiveness to patient questions while maintaining a clear boundary of professional role.

Screening Bot Prompting

The screening bot was designed to act as a "therapist" specializing in supporting patients with cancer dealing with fear, anxiety, depression, or other stress-related conditions. The prompt emphasized its role in fostering emotional well-being and therapeutic rapport. Key instructions included (1) respond as a therapist practicing nonjudgmental support, inspired by principles from acceptance and commitment therapy and mindfulness practices; (2) reduce patient stress by validating emotions, exploring coping mechanisms, and encouraging hope for change; (3) avoid clinical diagnostic language or prescribing treatments, focusing instead on promoting self-reflection and stress management strategies; and (4) engage the patient through open-ended questions and supportive dialogue, tailored to the specific emotional state of the patient persona.

Patient Bot Prompting

The patient bots were modeled to represent 3 distinct emotional personas—anxious, depressed, and frustrated—and were designed to simulate real-life patient interactions. Each patient bot was assigned the role of a 40-year-old male patient with lung cancer undergoing treatment. Detailed persona-specific instructions were included to guide their interactions:

- Persona-specific emotional states:
  - Anxious persona: Expresses uncertainty and seeks detailed explanations.
  - Depressed persona: Exhibits low engagement and responds with shorter, less optimistic answers.
  - Frustrated persona: Displays irritability and impatience in responses.
- Respond consistently with the designated persona's characteristics throughout the dialogue.
- Do not use or understand high-level medical terms unless explicitly explained by the provider bot.
- For the patient-education bot interactions:
  - Frame responses as questions about unclear cancer-related terms, procedures, or treatments.
- For the screening bot interactions:
  - Actively participate in therapy sessions, responding to the therapist bot's efforts to reduce stress while maintaining the persona's emotional tone.

Evaluator Bot Prompting

Evaluator bots were designed to act as "supervisors," assessing the interactions between a provider and a patient. They evaluated transcripts based on a scoring scale (1=poor and 3=excellent) tailored to the respective provider bot's role.

For the patient-education bot, the following five criteria (maximum score: 15) were used: (1) medical information accuracy, (2) clarity and simplicity of explanations, (3) expressions of empathy and warmth, (4) explanation of purpose or importance of procedures, and (5) adherence to professional role boundaries.

For the screening bot, fourteen criteria (adapted from the Enhancing Assessment of Common Therapeutic Factors tool [12]; maximum score: 42) were used: (1) verbal communication: open-ended questions, summarization, and clarification; (2) relationship building; (3) exploration and normalization of emotions; (4) expressions of empathy and warmth; (5) assessment of functioning and life evaluation; (6) exploration of social support; (7) incorporation of coping mechanisms; (8) evaluation of recent life events; (9) assessment of mental health; (10) collaborative goal-setting; (11) promotion of realistic hope for change; (12) use of simple, jargon-free language; (13) problem-solving steps and processes; and (14) integration of feedback.

Criteria unsuitable for chatbot interactions, such as nonverbal communication, were excluded with detailed reasons listed in Table 1.

Table . List of Enhancing Assessment of Common Therapeutic Factors (ENACT) factors removed with the reason for their removal.

ENACT factor	Reason for removal
Nonverbal communication and active listening	Therapist is a chatbot and therefore cannot display body language.
Therapist self-disclosure	Therapist is a chatbot and therefore has no real experiences to disclose.
Alcohol or drug and physical problems	Patient has cancer; therefore, physical, alcohol, or drug issues would need to go through their oncologist.
Involvement of family members or caregivers	Patient and therapist are chatbots; therefore, all sessions are assumed to be individual and one-on-one with no family involvement.
Confidentiality promotion	Therapist and patients are chatbots, so all conversations are assumed to be confidential and private.
Assessment of harm to self, harm to others, developing a collaborative response plan	For this study, patient chatbots were assumed not to have violent or suicidal tendencies.

Provider-Patient Interactions

Each provider bot engaged in 30 unique conversations, distributed evenly across 3 patient personas: anxious, depressed, and frustrated (10 conversations per persona). Conversations were facilitated through an application programming interface designed to streamline the flow of interactions. Each conversation consisted of 20 interactions, defined as 10 turns exchanged between the provider bot and the patient bot.

To simulate concise and realistic clinical exchanges, both provider and patient bots were programmed with the following parameters: a temperature setting of 0.7 (to ensure balanced creativity and consistency), Top P: 1, frequency penalty: 0, and presence penalty: 0. The token limit was removed to avoid interruptions, and each conversation was capped at 10 conversational turns to maintain brevity and clinical relevance. The resulting transcripts from these conversations were reviewed by evaluator bots using predefined criteria and subsequently cross-validated by human experts to ensure the reliability and validity of the evaluations.

Provider Bot Validation

A 2-step validation process was conducted. First, evaluator bots assessed the provider bots based on predefined criteria, generating scores and qualitative feedback. These results were then reviewed by human experts, a pediatric oncology nurse for the patient-education bot and a cognitive behavior counselor and PhD in IT psychology for the screening bot, using the same criteria used by the evaluator bots to ensure consistency and reliability.

Evaluator Bot Validation

Evaluator bots graded each interaction transcript based on predefined criteria, producing quantitative scores and qualitative comments. Human experts then reviewed the same transcripts, blind to the evaluator bot's results, and provided their own scores for comparison. The experts then reviewed the bots' evaluations to ensure that a consistent and reliable evaluation was carried out by the evaluator bot.

Statistical Analysis

Descriptive analyses were performed to evaluate interaction characteristics, including word count and sentiment trends. Cronbach α analysis was used to assess the reliability of evaluation criteria across evaluators. Differences in responses between GPT evaluators and human experts were analyzed using the Kruskal-Wallis test. ANOVA was used to identify significant variations in provider bot responses to different patient personas. All analyses were conducted using SPSS (version 24.0; IBM Corp).

Ethical Considerations

We did not have any human participants or animal subjects and therefore did not need to go before an ethics board.

Results

Evaluation of the Patient-Education Bots by AI and Human Evaluators

The patient-education bot, evaluated by both AI and human evaluators, exhibited remarkably consistent performance across interactions with patient bots displaying frustrated, depressed, and anxious personas. The patient-education bot consistently provided accurate medical information, as validated by an experienced oncology nurse, and delivered clear explanations that were fully understood by patient bots, with no instances of confusion reported. Specifically, the AI evaluator assigned perfect or near-perfect mean scores of 15 (SD 0.00), 14.9 (SD 0.31), and 15 (SD 0.00), respectively, while the human evaluator echoed these assessments with similarly high mean scores of 14.9 (SD 0.31), 14.9 (SD 0.31), and 15 (SD 0.00), respectively. The AI evaluator described the patient-education bot to have "... demonstrated excellent skills in providing education and support to the patient. The information provided was accurate, comprehensive, and clearly articulated, catering to the patient's understanding. The nurse exhibited great empathy and warmth throughout the interactions, which significantly contributed to patient comfort, trust, and engagement. The nurse did not overstep their boundaries by making specific medical recommendations, respecting the role of the patient's treatment team. Overall, the nurse demonstrated exceptional patient-education skills."

A singular point of contention arose from the AI evaluator's interpretation of the patient-education bot potentially recommending treatments beyond its scope. The AI evaluator stated that the nurse could benefit from "being cautious and mindful to avoid being perceived as providing personalized treatment suggestions." This was later clarified as a misunderstanding, attributing the issue to the AI evaluator's scoring framework rather than the patient-education bot's performance.

The patient-education bot was described by the human evaluator to be "correct" and "well-organized and explained," but the bot's "[constant expression] of empathy" was reported to "[feel] a bit mechanical." It was noted that this bot would "likely be helpful, as it can repeatedly explain medical concepts on behalf of medical staff members who do not always have enough time for explanations."

Evaluation of the Screening Bot by AI and Human Evaluators

The average AI evaluator bot's scores for the pretherapy screening bot when interacting with the frustrated, depressed, and anxious patient bots, respectively, were 40.1 (SD 1.28), 40.3 (SD 1.05), and 40.7 (SD 1.15), of a total possible score of 42. Across all 3 patient bots, the lowest scoring criterion was the evaluation of realistic hope for change, which had an average score of 2.53 out of 3 (SD 0.51). Human expert evaluators corroborated the AI evaluation results. The average human evaluator scores of the screening bot when interacting with the frustrated, depressed, and anxious patient bots, respectively, were 37.5 (SD 0.84), 37.6 (SD 0.96), and 36.9 (SD 2.60) for the first reviewer and 36.8 (SD 1.31), 36.9 (SD 1.10), and 36.2 (SD 2.09) for the second reviewer.

The AI evaluator, under the impression it was assessing a human, reported that the pretherapy screening bot excelled in maintaining effective communication, building a warm relationship, and demonstrating empathy. The evaluator bot identified several strengths of the screening bot, stating that it "... provides a warm and empathetic attitude and responds likewise to the patient's negative reactions and feelings and leads the conversation naturally." The most common areas for improvement mentioned in the final comments were "exploration of prior successful coping strategies and providing more explicit encouragement for feedback."

Human evaluators similarly concluded that the pretherapy screening bot excelled in "... [communicating] clearly," building a "warm and empathetic" relationship, and "[leading] the conversation naturally." The screening bot reportedly could improve upon "exploring prior coping strategies and patient history a little more deeply" and was occasionally reported to be too informational or talkative. It was reported to "[pass] to the next topic too quickly (possibly due to its large list of duties—which the therapist was prompted to do)." Overall, the human reviewers suggested that "the bot is useful for initial consultations—the AI fluently checks for components of the initial step of counseling." Furthermore, it was noted that "a more detailed score standard is required for the evaluator bot's prompt."

Statistical Analysis Result

Patient-Education Bot

For the patient-education bot, the evaluation scores from both AI and human evaluators were remarkably consistent, showing minimal variance. This uniformity limited the opportunity for further analysis, as the lack of significant differences between evaluator scores precluded more detailed statistical comparisons.

Screening Bot

The Kaiser-Meyer-Olkin and Bartlett sphericity test results indicate that the Kaiser-Meyer-Olkin value of 0.714 suggests that the sample is suitable for factor analysis, and the significance probability of Bartlett sphericity test is less than .001, indicating that the correlation between variables is significant. The results of the communality analysis show that all variables have a communality of 1, indicating that all variables explain the extracted factors well. The 5 extracted factors explain 66.327% of the total variance of the variables.

Significant findings from the ANOVA analysis indicate notable variations in group responses across several key evaluation criteria for the screening bot. This variability suggests that specific factors or treatments have a meaningful impact on participant responses, reflecting their efficacy or relevance in different contexts. Verbal communication (open-ended questions, summarization, and clarification) demonstrated a highly significant difference between groups (P<.001), suggesting that the approach to verbal communication significantly affects the responses. Assessment of functioning and life evaluation exhibited one of the highest significances (P<.001), pointing to the critical role this factor plays in differentiating responses among groups. Exploration of the patient's social support network also showed a highly significant difference (P<.001), indicating a strong effect of social support exploration on participant responses. Assessment of mental health highlighted the most substantial difference between group means (P<.001), underscoring the importance of mental health assessment in eliciting varied responses. Evaluation of recent events in the patient's life and evaluation of realistic hope for change both showed significant differences between groups (P<.001 for the former and P<.001 for the latter), suggesting these areas notably influence responses.

Other significant areas include relationship building and exploration, interpretation, and normalization of emotions, with P values of .004 and .002, respectively, indicating noticeable effects on the responses, albeit less pronounced compared to the areas mentioned earlier. Nonsignificant findings were observed in the expression of empathy, warmth, and genuineness and collaborative goal-setting and managing patient's expectations, with P values of .36 and .28, respectively. These results suggest that variations in group responses to these criteria might not be significantly influenced by the tested factors, potentially due to inherent similarities in the implementation or perception of these aspects.

The use of easy-to-understand vocabulary and integration of feedback, giving advice, and recommendations showed moderate significant differences (P=.04 and P=.02, respectively), indicating that these areas have a discernible but varied impact on participant responses.

The detailed ANOVA analysis underscores the nuanced impact of different therapeutic communication and evaluation strategies on participant responses. It highlights the areas where specific approaches significantly influence outcomes, offering insights into the effectiveness of various therapeutic and communicative techniques.

The ANOVA results highlight the variability in how different groups responded to the questions. Significant P values (P<.05) indicate that the groups do not share the same mean response to a question, suggesting that the factor or treatment being tested influences the responses. The strength of this effect varies among the questions, as evidenced by the range of F values and P values.

Discussion

Principal Findings

Overall, the insights gained from this research suggest that AI health care chatbots can be developed, tested, and validated within a relatively short time frame using the 3-bot system. The results of the 3-bot evaluation system suggest that this method can prove valuable for extensive testing of early-stage health care chatbots. The patient bots are able to mimic patient dialogue and provide a platform for the provider bots to output their responses, while the evaluator bot is able to comb through the interaction transcripts and flag any potentially inappropriate responses, greatly reducing the amount of work for researchers. Furthermore, this 3-bot system is highly customizable and can be adapted to fit the needs and cultural norms required by the developers. It is also highly scalable, as the basic requirements to perform the 3-bot evaluations are a computer system and access to an LLM. Performing more iterations of an evaluation only requires a marginal amount of researcher effort, and performing multiple, different evaluations can be accomplished simultaneously, given the computer system has enough processing power.

This study introduces a novel AI-powered health care chatbot validation system featuring 3 types of AI bots—provider, patient, and evaluator. This 3-bot AI system represents a novel methodology not previously explored in existing literature, extending beyond the importance of validation discussed by Bohr and Memarzadeh [2] in AI's rise in health care, which did not delve into the conversational capabilities between different AI systems in clinical simulations. To our knowledge, our method of testing and evaluating the performance of AI health care provider bots by having them interact with other patient bots and then reviewing the transcripts with an evaluator AI bot has never been reported before.

In our study, we created 2 health care provider bots as examples to demonstrate our system, a patient-education bot and a mental health screening bot. The provider bots were intricately designed to replicate the roles traditionally held by human health care providers, addressing the urgent need for scalable and effective patient care solutions highlighted by Patel et al [13]. These bots are intended to support the health care workforce, which, according to the World Health Organization, is expected to face a significant shortfall [14]. By automating routine tasks, these AI systems could alleviate some of the burdens placed on human staff, allowing them to focus on more complex and sensitive care activities. Already, several health care chatbots are in development, including those designed to answer patient questions and provide mental health therapy [14,15].

However, provider chatbots such as these still require extensive testing, traditionally done by enrolling patients as subjects, which negatively affects the speed and resource cost of developing these tools while running the risk of exposing the patients to unvalidated AI. Therefore, we created 3 types of AI patient bots with personas as examples to test our provider bots. In designing the patient bots, we drew inspiration from Fortin et al [16], who emphasize the importance of personalized and empathetic care in treatment outcomes. In previous studies, various digital patient bots were reported in medical education. In our study, the patient bots were imbued with diverse emotional and psychological states to test the adaptability and responsiveness of the provider bots in a controlled, yet realistic environment, simulating real-life patient interactions.

Comparison to Prior Work

These current methods require great human input during the iterative testing and evaluation phases, which requires researchers and developers to invest significant time and effort. In contrast, using the 3-bot validation method removes the need for separate human responders and human evaluators, greatly streamlining the initial testing and evaluation process and focusing work efforts on areas of the evaluated bot that require improvement.

Until now, bot-bot interactions were manually reviewed by human experts, which greatly slows the validation process. Current methods of evaluating health care chatbots include a human reviewer reviewing the health care chatbot's performance against a grading standard as seen in Lechien et al [17] and Goodman et al [18], a human reviewer grading the health care chatbot's performance against another pre-established chatbot as seen in Aljamaan et al [19], or a mix of the methods, as seen in Huang et al [20].

In this study, we created 2 AI evaluator bots to demonstrate the feasibility of using them as first-line evaluators in addition to human experts. The role of the evaluator bots was crucial in objectively assessing the quality of interactions between provider and patient bots, ensuring adherence to predefined criteria. This evaluation process mirrors the necessity of validation for AI systems before clinical application as emphasized by Kretzschmar et al [21]. By comparing the evaluations conducted by AI evaluator bots with assessments from human experts, we ensured the feasibility of our system, further grounding the study in rigorous scientific methodology. To date, AI bots have been used to review text messages and academic manuscripts, but this is the first study to review dialogue between 2 bots for the purposes of evaluation.

Limitations and Future Directions

While promising, this study has limitations that warrant consideration. First, the evaluation criteria used were relatively limited in scope, which may not have captured subtle differences in performance between AI and human evaluators. Future research should incorporate more comprehensive and granular criteria to enable more nuanced evaluations. Retrieval-augmented generation could further enhance the evaluator bots by enabling them to cross-verify provider bot responses against dynamic, vetted information sources, thereby increasing the accuracy and reliability of evaluations.

Second, the patient bots were prompted using relatively concise instructions due to the practical constraints of maintaining response speed. This may have limited the complexity and variability of their responses, potentially underrepresenting the breadth of emotions and behaviors seen in real-world patients. Future studies should explore more elaborate prompting strategies or advanced techniques like retrieval-augmented generation to overcome this limitation.

Third, the study used prioritized examples of clinically relevant patient personas (anxious, depressed, and frustrated), chosen for their significance in addressing common and challenging scenarios in clinical practice. While these personas are a high priority for an initial evaluation, they do not fully represent the diversity of patient interactions.

Finally, biases inherent in LLMs may have influenced the results, despite efforts to standardize demographic inputs across all patient bots. Nonessential demographic details were excluded to minimize biases related to race, political affiliation, sexual orientation, or socioeconomic status. Nonetheless, future research should explore the use of specialized LLMs with controlled training datasets to further mitigate such biases.

Conclusions

We underscore the successful development and implementation of a novel 3-bot evaluation system. This system, consisting of provider bots, patient bots, and evaluator bots, represents a pioneering approach to testing and validating AI functionalities without the need for real patient interactions. Our findings offer a practical solution and set a benchmark for future AI-driven health care services, providing a direction for subsequent research and development efforts.

Acknowledgments

This research was supported by a grant from the Korea Health Technology R&D Project from the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare, Republic of Korea (grant HI23C038700), as well as through the resources of Med2Lab, Inc. The authors would like to express their sincere gratitude to all the individuals who contributed to the successful completion of this study. The authors would like to thank Binh Nguyen and Ha Huynh for setting up the application programming interface and providing large language model–related advice, Eunju Cho for helping set up the patient-education nurse evaluation criteria and for evaluating the patient-education nurse, Yoojin Seo for setting up the screening therapist approach and for evaluating the screening therapist, and Jungyi Ma for evaluating the screening therapist.

Conflicts of Interest

BT and KE have disclosed a financial relationship with Med2Lab Inc., which includes board membership, employment, and equity or stock ownership. They confirm that they had no involvement or influence in the design of the methodology or the results of this study. All other authors declare that they have no known competing financial interests or personal relationships that could have influenced the work reported in this paper.

References

1. Global strategy on human resources for health: workforce 2030. World Health Organization. 2016. URL: https://iris.who.int/bitstream/handle/10665/250368/9789241511131-eng.pdf [accessed 2025-02-07]
2. Bohr A, Memarzadeh K. The rise of artificial intelligence in healthcare applications. In: Artificial Intelligence in Healthcare: Elsevier; 2020:25-60. [doi: 10.1016/B978-0-12-818438-7.00002-2]
3. Benbelkacem S, Kadri F, Atmani B, Chaabane S. Machine learning for emergency department management. Int J Inf Syst Serv Sect 2019 Jul;11(3):19-36. [doi: 10.4018/IJISSS.2019070102]
4. Allen MR, Webb S, Mandvi A, Frieden M, Tai-Seale M, Kallenberg G. Navigating the doctor-patient-AI relationship—a mixed-methods study of physician attitudes toward artificial intelligence in primary care. BMC Prim Care 2024 Jan 27;25(1):42. [doi: 10.1186/s12875-024-02282-y] [Medline: 38281026]
5. Puri A, Mathur R, Sindhu N. Harnessing the power of AI in healthcare: benefits, concerns, and challenges for medical personnel training. AHOAJ 2024;6(2):90-91. [doi: 10.15406/ahoaj.2024.06.00227]
6. Loh E. ChatGPT and generative AI chatbots: challenges and opportunities for science, medicine and medical leaders. BMJ Lead 2023 May 2:leader-2023-000797. [doi: 10.1136/leader-2023-000797] [Medline: 37192124]
7. Wilhelm TI, Roos J, Kaczmarczyk R. Large language models for therapy recommendations across 3 clinical specialties: comparative study. . 2023 p. e49324. [doi: 10.2196/49324]
8. Shanahan M, McDonell K, Reynolds L. Role play with large language models. Nature New Biol 2023 Nov;623(7987):493-498. [doi: 10.1038/s41586-023-06647-8] [Medline: 37938776]
9. Shao Y, Li L, Dai J, Qiu X. Character-LLM: a trainable agent for role-playing. Presented at: The 2023 Conference on Empirical Methods in Natural Language Processing; Dec 6-10, 2023; Singapore. [doi: 10.18653/v1/2023.emnlp-main.814]
10. Yang HS, Wang F, Greenblatt MB, Huang SX, Zhang Y. AI chatbots in clinical laboratory medicine: foundations and trends. Clin Chem 2023 Nov 2;69(11):1238-1246. [doi: 10.1093/clinchem/hvad106] [Medline: 37664912]
11. 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]
12. Kohrt BA, Jordans MJD, Rai S, et al. Therapist competence in global mental health: Development of the ENhancing Assessment of Common Therapeutic factors (ENACT) rating scale. Behav Res Ther 2015 Jun;69:11-21. [doi: 10.1016/j.brat.2015.03.009] [Medline: 25847276]
13. Paterick TE, Patel N, Tajik AJ, Chandrasekaran K. Improving health outcomes through patient education and partnerships with patients. Bayl Univ Med Cent Proc 2017;30(1):112-113. [doi: 10.1080/08998280.2017.11929552]
14. Görtz M, Baumgärtner K, Schmid T, et al. An artificial intelligence-based chatbot for prostate cancer education: design and patient evaluation study. Digit Health 2023;9:20552076231173304. [doi: 10.1177/20552076231173304] [Medline: 37152238]
15. Singh J, Sillerud B, Singh A. Artificial intelligence, chatbots and ChatGPT in healthcare—narrative review of historical evolution, current application, and change management approach to increase adoption. J Med Artif Intell 2023;6:30-30. [doi: 10.21037/jmai-23-92]

16. Fortin J, Leblanc M, Elgbeili G, Cordova MJ, Marin MF, Brunet A. The mental health impacts of receiving a breast cancer diagnosis: a meta-analysis. Br J Cancer 2021 Nov;125(11):1582-1592. [doi: 10.1038/s41416-021-01542-3] [Medline: 34482373]
17. Lechien JR, Maniaci A, Gengler I, Hans S, Chiesa-Estomba CM, Vaira LA. Validity and reliability of an instrument evaluating the performance of intelligent chatbot: the Artificial Intelligence Performance Instrument (AIPI). Eur Arch Otorhinolaryngol 2024 Apr;281(4):2063-2079. [doi: 10.1007/s00405-023-08219-y] [Medline: 37698703]
18. Goodman RS, Patrinely JR, Stone CA Jr, et al. Accuracy and reliability of chatbot responses to physician questions. JAMA Netw Open 2023 Oct 2;6(10):e2336483. [doi: 10.1001/jamanetworkopen.2023.36483]
19. Aljamaan F, Temsah MH, Altamimi I, et al. Innovation of referencing hallucination score for medical AI chatbots' and comparison of six large language models. JMIR Preprints. Preprint posted online on Nov 6, 2023. [doi: 10.2196/54345]
20. Huang RS, Benour A, Kemppainen J, Leung FH. The future of AI clinicians: assessing the modern standard of chatbots and their approach to diagnostic uncertainty. BMC Med Educ 2024 Oct 11;24(1):1133. [doi: 10.1186/s12909-024-06115-5] [Medline: 39394122]
21. Kretzschmar K, Tyroll H, Pavarini G, Manzini A, Singh I, NeurOx Young People's Advisory Group. Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. Biomed Inform Insights 2019;11:1178222619829083. [doi: 10.1177/1178222619829083] [Medline: 30858710]

Abbreviations

AI: artificial intelligence
LLM: large language model

Edited by E Borycki, K Cato; submitted 09.06.24; peer-reviewed by D Chrimes, E Bai, M Chatzimina; revised version received 21.12.24; accepted 02.01.25; published 27.02.25.

Please cite as:
Choo S, Yoo S, Endo K, Truong B, Son MH
Advancing Clinical Chatbot Validation Using AI-Powered Evaluation With a New 3-Bot Evaluation System: Instrument Validation Study
JMIR Nursing 2025;8:e63058
URL: https://nursing.jmir.org/2025/1/e63058
doi:10.2196/63058

Review

Examining the Role of AI in Changing the Role of Nurses in Patient Care: Systematic Review

Inas Al Khatib[1], MBA, MSc; Malick Ndiaye[1], DPhil

[1]Department of Industrial Engineering, College of Engineering, American University of Sharjah, Sharjah, United Arab Emirates

Corresponding Author:
Inas Al Khatib, MBA, MSc
Department of Industrial Engineering
College of Engineering
American University of Sharjah
University City
Sharjah, 26666
United Arab Emirates
Phone: 971 65155555
Fax: 971 65152200
Email: g00091914@aus.edu

Abstract

Background: This review investigates the relationship between artificial intelligence (AI) use and the role of nurses in patient care. AI exists in health care for clinical decision support, disease management, patient engagement, and operational improvement and will continue to grow in popularity, especially in the nursing field.

Objective: We aim to examine whether AI integration into nursing practice may have led to a change in the role of nurses in patient care.

Methods: To compile pertinent data on AI and nursing and their relationship, we conducted a thorough systematic review literature analysis using secondary data sources, including academic literature from the Scopus database, industry reports, and government publications. A total of 401 resources were reviewed, and 53 sources were ultimately included in the paper, comprising 50 peer-reviewed journal articles, 1 conference proceeding, and 2 reports. To categorize and find patterns in the data, we used thematic analysis to categorize the systematic literature review findings into 3 primary themes and 9 secondary themes. To demonstrate whether a role change existed or was forecasted to exist, case studies of AI applications and examples were also relied on.

Results: The research shows that all health care practitioners will be impacted by the revolutionary technology known as AI. Nurses should be at the forefront of this technology and be empowered throughout the implementation process of any of its tools that may accelerate innovation, improve decision-making, automate and speed up processes, and save overall costs in nursing practice.

Conclusions: This study adds to the existing body of knowledge about the applications of AI in nursing and its consequences in changing the role of nurses in patient care. To further investigate the connection between AI and the role of nurses in patient care, future studies can use quantitative techniques based on recruiting nurses who have been involved in AI tool deployment—whether from a design aspect or operational use—and gathering empirical data for that purpose.

(JMIR Nursing 2025;8:e63335) doi:10.2196/63335

KEYWORDS

artificial intelligence; AI; nursing practice; technology; health care; PRISMA

Introduction

Background

The science and engineering field of artificial intelligence (AI) is concerned with the theory and application of creating systems that display the traits we identify with intelligence in human behavior [1]. The years 2000 to 2015 saw an upward trend in the growth of AI. With dramatic revolutions influenced by both ideas and methodologies, the progress of AI has promoted the development of human civilization in our day and age. However, due to its interdisciplinary nature and rapid expansion, AI is a discipline that is challenging to fully comprehend and is getting more and more flexible from the standpoint of reference behavior [2].

The previous decade was defined by AI, and the upcoming one will most likely also be defined by it. Systems that exhibit intelligent behavior by analyzing their surroundings and acting with some autonomy to accomplish predetermined goals are referred to as AI systems. Greater accuracy is needed to have relevant and fruitful discussions on AI because it encompasses so many different methodologies and circumstances. Arguments regarding straightforward “expert systems” that serve advising functions, for instance, must be separated from those about sophisticated data-driven algorithms that make conclusions about specific persons automatically. Similarly, it is crucial to distinguish between arguments regarding hypothetical future advancements that may never materialize and those regarding actual AI that already has an impact on society today including the nursing practice [3].

Numerous ideas, including computing, developing software, and transmitting data, are built on AI. Machine learning (ML), deep learning, natural language processing (NLP), voice recognition, robots, and biometric identification are examples of technologies that use AI. AI is used in a wide range of industries, including the health care, industrial, and automotive sectors as well as corporate organizations. AI also provides several benefits that help it become increasingly popular across numerous industries. AI-powered machines are accurate and efficient, can do many tasks at once, and their work costs less than a human’s. However, there are other issues with AI that make it difficult to use. Technology, security, and data issues are common with AI, and if users do not comprehend the system, mishaps may occur. The expanded use of AI has changed several industries by improving organizational effectiveness and enabling data security [4].

AI in nursing is revolutionizing the field by enhancing patient care, improving efficiency, and reducing the workload on nurses. AI-powered tools and applications enable real-time monitoring of patient’s vital signs, predicting potential health deteriorations, and providing alerts for immediate intervention. AI algorithms can analyze large volumes of patient data to assist in accurate diagnosis and personalized care plans [5]. Moreover, AI chatbots and virtual assistants support administrative tasks, such as scheduling and documentation, allowing nurses to focus more on direct patient care [6]. By automating routine tasks and providing decision support, AI empowers nurses to deliver higher quality care with greater precision and efficiency [7].

Research Rationale and Aim

The rationale behind this research is to investigate how the increasing use of AI in health care affects the role of nurses in patient care. As AI technologies become more integrated into health care systems, understanding their impact on nursing practice is crucial. AI’s applications, ranging from clinical decision support to operational improvements, promise to transform various aspects of health care, including nursing. By examining whether AI has led to changes in the nursing role or is likely to do so, this research aims to provide insights into how these technologies influence nursing responsibilities and practices. The aim of this review is to explore the evolution of AI as a technology through its various developmental phases. In this systematic literature review, we examine the different applications and deployments of AI in the nursing field. The primary research question addressed is “How will AI transform the role of nurses in patient care?”

Research Significance

The review offers important perspectives on how AI is transforming the roles and duties of nurses in patient care. This understanding is essential for adapting nursing education, training, and practice to align with evolving technological advancements. By identifying how AI impacts nursing roles, the research can guide the effective implementation of AI tools in health care settings. It highlights the importance of involving nurses in the development and deployment of AI technologies to ensure that these tools enhance rather than disrupt nursing practice. The findings can inform health care policies and training programs by emphasizing the need for ongoing professional development and support for nurses as they integrate AI into their workflows. This ensures that nurses are prepared to leverage AI effectively while maintaining high standards of patient care. The study contributes to the existing body of knowledge on AI in health care and sets the stage for future research. It opens avenues for quantitative studies and empirical data collection to further explore the relationship between AI and nursing roles, providing a foundation for evidence-based practice and decision-making. Overall, this research is important for its potential to enhance the understanding of AI’s impact on nursing practice, guide effective technology integration, and shape the future of nursing education and policy.

Methods

Overview

A well-defined review protocol (Textbox 1) was established at the outset of the research to guarantee that the review process is transparent, reproducible, methodical, and provides a clear roadmap for conducting and reporting the review.

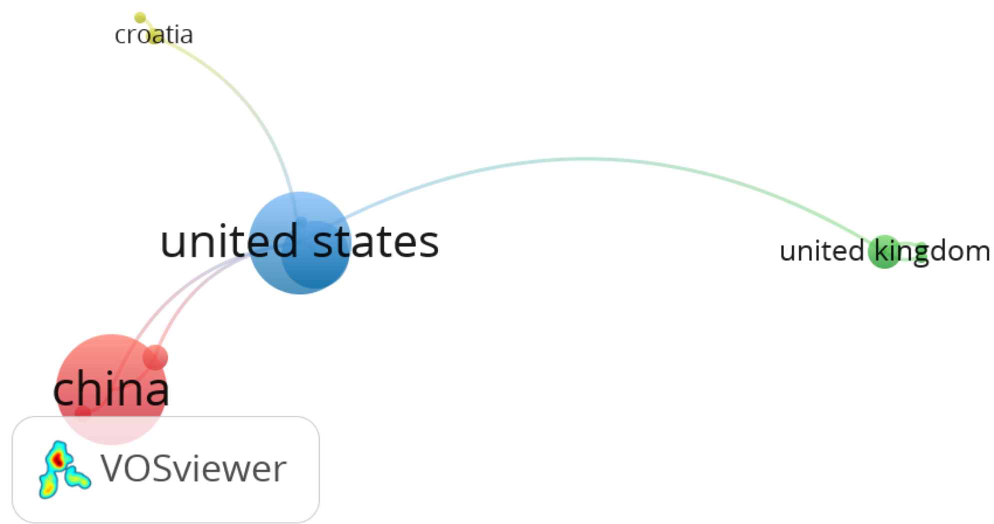
Textbox 1. Systematic literature review protocol.

- Title: confirm a clear, descriptive title for our systematic review.
- Background: explain the rationale behind this research, why it is needed, and its significance. In addition, define the main research aim and research question of the systematic literature review.
- Eligibility criteria: define what was included and excluded from the review inclusive of the search time frame.
- Information sources: list the database searched for academic sources and the gray literature used.
- Search strategy: detail the search terms used and justification of their use.
- Study selection: describe the process for screening that involved reviewers and selecting studies.
- Data extraction and analysis: define the data extraction process and the outcome of data analysis.

This paper presents the findings of a thorough analysis and critical assessment of the pertinent literature using systematic database-searching approaches. The critical assessment in this systematic literature review involves evaluating study quality, relevance, biases, findings synthesis, and implications. This systematic review draws information from credible industry sources as well as published, peer-reviewed English language papers. To comprehend the development of this idea, we consulted trustworthy industry publications known as "gray paper" and the Scopus database covering the last 33 years (ie, 1991-2024). A total of 53 sources comprised of 50 peer-reviewed academic articles, 1 conference proceeding, and 2 reports were included in the research.

After completing a bibliographic analysis, a suitable keyword search strategy was chosen, such as "AI applications in nursing," and the search was restricted to the previously specified time frame. On the basis of the outputs of the Scopus database, which have been used in this study to build the various bibliometric maps, bibliometric networks were created using the VOSviewer program (Centre for Science and Technology Studies). After exporting the sophisticated Scopus-based search results to the VOSviewer program, a network visualization was created, as shown in Figure 1, to show how the authors in this area are related to each other through publications on this subject.

Figure 1. Most countries with research subject "artificial intelligence applications in nursing"—network visualization.



In the network visualization illustrated in Figure 1, the United States, followed by China and the United Kingdom, has been publishing journals about AI applications in the field of nursing, demonstrated by the weight of those countries' representation. Furthermore, the links between those circles indicate that the relatedness of the journals in terms of cocitation links also illustrates that other countries' journal publications relied on the US publications. A VOSviewer mapping was then done using "AI applications in nursing" as the keyword.

This has prompted the expansion of the search keywords; Table 1 represents a series of search strings focusing on various aspects of AI and its applications, particularly in nursing and related technologies. Alongside, the rationale for each search string is provided along with the number of academic journals found for each. Each search string is designed to capture specific facets of AI to ensure a comprehensive and inclusive exploration of relevant literature. By using these specific search strings, the research ensures a thorough and targeted review of the literature across different aspects and applications of AI, with a special focus on health care and nursing.

Table 1. Search string in the Scopus database (N=2870).

Search string	Justification	Results, n (%)
"AI" AND "technology" AND "in" AND "nursing"	This broad search term is used to gather general information and foundational literature on AI[a], which will provide a broad understanding and context for more specific searches.	164 (5.71)
"Artificial" AND "Intelligence" AND "applications" AND "in" AND "nursing"	This phrase search ensures that both terms are explicitly present, helping to find more specific and relevant documents that discuss AI in a detailed manner.	226 (7.87)
"AI" AND "use" AND "in" AND "nursing"	This search string targets the literature that explores the application and impact of AI technologies specifically within the field of nursing, ensuring relevance to health care.	162 (5.64)
"Nursing" AND "AI" AND "applications"	By including these terms, the search focuses on technological advancements and their practical uses in the nursing profession, broadening the scope beyond just AI.	105 (3.65)
"Evolution" AND "of" AND "Artificial" AND "Intelligence" AND "Approaches" AND "in" AND "nursing"	This search string aims to find the literature on the historical development and various methodologies within AI, providing context and background on how AI approaches have changed over time.	6 (0.20)
"Symbolic" AND "AI" AND "Approach"	Symbolic AI is a specific paradigm within AI research. This search will help identify works focused on this particular approach, which is crucial for understanding different AI methodologies.	620 (21.60)
"Data-Driven" AND "AI" AND "Approach" AND "in" AND "nursing"	Data-driven AI approaches, including ML[b] and neural networks, are fundamental to modern AI. This search focuses on the literature that discusses these data-centric methodologies.	2 (0.06)
"Artificial" AND "General" AND "Intelligence" AND "Approach" AND "in" AND "nursing"	AGI[c] represents a more advanced and comprehensive form of AI. This search will help identify research on AGI, exploring its potential and challenges.	10 (0.34)
"Artificial" AND "Intelligence" AND "application" AND "in" AND "nursing"	This search string is designed to find specific case studies and examples of how AI is being applied in nursing, providing practical insights and real-world applications.	230 (8.01)
"Rothman" AND "Index" AND "Use" AND "for" AND "Patient" AND "Acuity" AND "and" AND "Risk"	The Rothman Index is a specific tool used in health care. This search targets the literature on its use and effectiveness in assessing patient acuity and risk, relevant for AI applications in patient monitoring.	1 (0.03)
"Social" AND "robots" AND "use" AND "in" AND "nursing"	Social or companion robots are an emerging area within AI and robotics. This search aims to find the literature on their use, particularly in providing care and support in health care settings.	104 (3.62)
"TeleRobots"	Telerobots are used for remote operations, which can be highly relevant in health care for tasks such as remote surgery or patient care. This search focuses on this specific technology.	105 (3.65)
"Natural" AND "Language" AND "Processing" AND "in" AND "nursing"	NLP[d] is a key area within AI, crucial for developing systems that can understand and process the human language. This search targets the literature on NLP, which has substantial applications in health care communication and data analysis.	256 (8.91)
"Robotic" AND "Process" AND "Automation" AND "in" AND "nursing"	RPA[e] is a form of business process automation technology based on AI. This search string is aimed at finding the literature on how RPA can be applied in health care operations and administration.	13 (0.45)
"Machine" AND "Learning" AND "use" AND "in" AND "nursing"	ML is a core component of AI. This search aims to gather comprehensive literature on ML techniques and their applications across various domains, including health care.	208 (7.24)
"nurse" AND "role" AND "transformation"	AI is driving significant changes in health care by automating tasks, supporting decision-making, and transforming traditional nursing functions. The focus on "role" and "transformation" highlights how nurses' responsibilities are evolving due to AI integration, requiring new skills and altering patient care practices. These keywords enable a targeted exploration of the evolving landscape of nursing in the context of AI-driven health care.	658 (22.92)

[a]AI: artificial intelligence.

[b]ML: machine learning.

[c]AGI: artificial general intelligence.

[d]NLP: natural language processing.

[e]RPA: robotic process automation.

Data Collection and Analysis

Academic Search

The inclusion and exclusion criteria are presented in Textbox 2.

Textbox 2. Inclusion and exclusion criteria.

Inclusion criteria

- Empirical studies, conference proceedings, and reports
- Papers with clear research questions and objectives on the application of artificial intelligence in the field of nursing
- Time period: from 1991 to 2024 (ie, 33 years)
- Papers published in the English language

Exclusion criteria

- Conceptual papers, editorials, academic book sections, and literature reviews
- Industrial sectors other than health care
- Publications before 1991
- Other languages

Data Selection

Overview

Information from all the 401 references was compiled in a soft copy folder. These references were independently reviewed by the main author, who selected the final list of papers to be analyzed. The first author examined the articles' topics and content and used our criteria for inclusion and exclusion of material to eliminate papers whose research questions were not fully aligned with the scope of this review. The second author upheld the main author's decision to exclude the resource from the study. The inclusion criteria are the characteristics that must exist to be included in this study, while the exclusion criteria are those characteristics that disqualify a data source from inclusion in the paper, which leads to the identification of 50 relevant journals published between 1991 and 2024.

Gray Literature Search

With regard to searching supplementary sources, it was imperative to expand the search radius to include official newspapers and reliable industry sources because the topic of this study is a prominent issue in industry trends. These sources capture the expert opinions of subject matter experts and produce additional information from trustworthy sources such as the European Parliamentary Research Service and National Bureau of Economic Research. Pertinent supplemental sources were found as a result, and the study report examined them all.

The critical assessment of the included studies was conducted through a self-rating process by 2 authors. Each author independently reviewed and appraised the quality of the studies based on predefined criteria relevant to the study designs, including risk of bias, methodology, and relevance to the research question. The 2 authors then compared their ratings, and any discrepancies were discussed and resolved through a consensus. This self-rating approach was used to streamline the evaluation process while ensuring consistency in the appraisal.

Furthermore, Table 2 provides a summary of the selected data in the systematic review literature by type.

Table 2. Systematic review analysis summary by type.

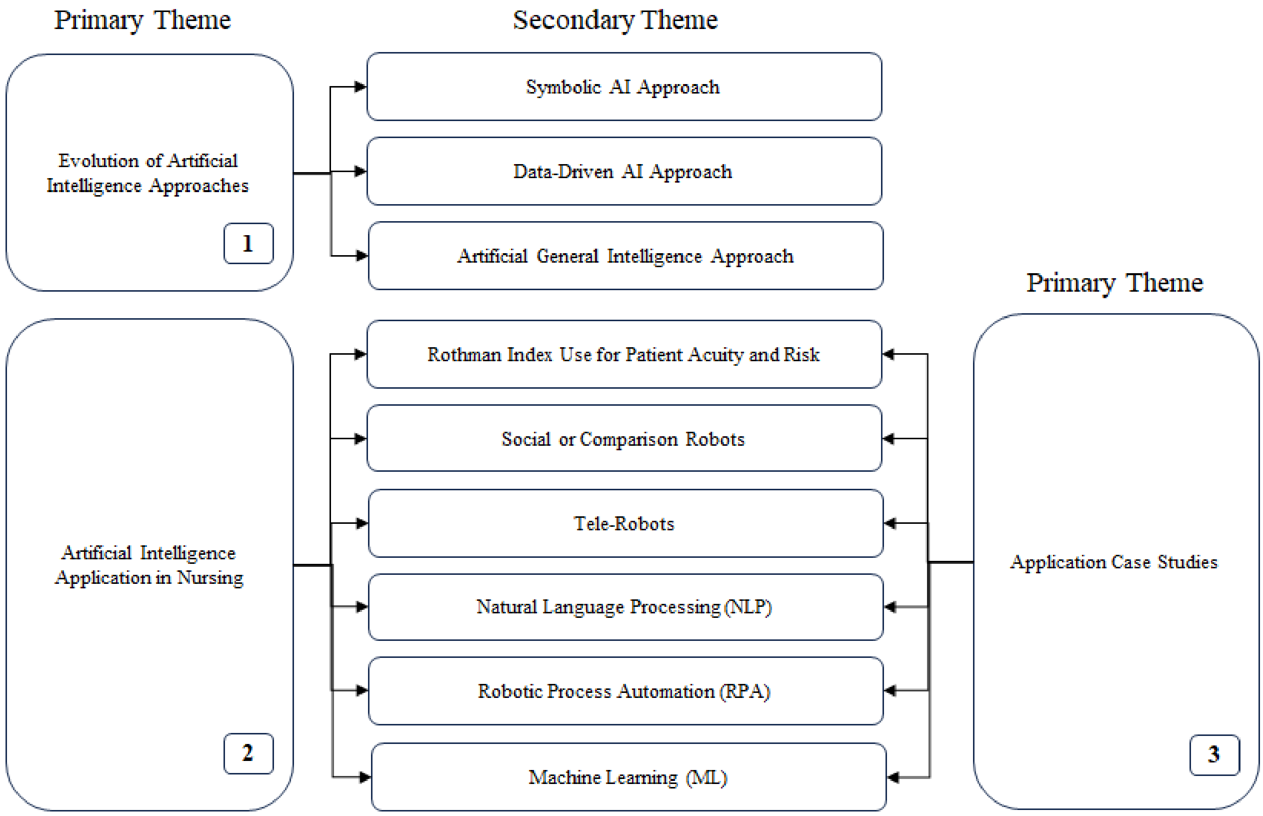
Reference type	Values, n (%)
Journal article	50 (94)
Conference proceeding	1 (2)
Report	2 (7)

Thematic Analysis

The study made use of secondary data sources; these resources offer details on the numerous ways AI is being used in nursing and how this is changing the role of nurses. The collected material was examined using a thematic analysis approach. The information gathered from the literature study was carefully reviewed by 2 independent reviewers and categorized in accordance with the primary themes and supporting themes that surfaced. Common patterns, trends, and important discoveries must be found to comprehend the connection between the findings. The primary themes that were identified from the literature are mapped in Figure 2.

This methodology was consistently applied throughout the paper and the results produced will be discussed in the subsequent section.

Figure 2. Primary and secondary themes in the systematic literature review. AI: artificial intelligence.



A thematic analysis was conducted explicitly designed to assess the impact of AI on the role of nurses. The use of the keyword "nurse role transformation" triggered an extensive review of the relevant literature, where we systematically screened for statements or data points related to the interaction between nurses and AI-based technologies. For example, we examined whether articles explicitly discussed shifts in task allocation, automation of routine functions, or changes in decision-making responsibilities. Furthermore, we used a matrix systematization process where each identified AI technology was mapped against the roles and responsibilities traditionally held by nurses, as well as any newly emerging roles due to the technology's integration. This allowed us to systematically capture how AI is transforming the scope of nursing practice, such as by enabling nurses to focus more on patient-centered care while AI systems manage data analysis or administrative tasks.

In the Results section, we expanded on this by introducing a subtheme explicitly titled "roles of nurses and role transformation." This subtheme synthesized case studies and literature findings that demonstrated specific examples of role shifts, such as how AI-assisted diagnostic tools are enabling nurses to participate more actively in clinical decision-making or how AI-driven administrative systems reduce the clerical burden on nurses. These shifts were categorized into functional changes (such as delegation of monitoring tasks to AI systems) and strategic changes (such as enhanced involvement in decision-making processes due to AI's real-time data processing capabilities).

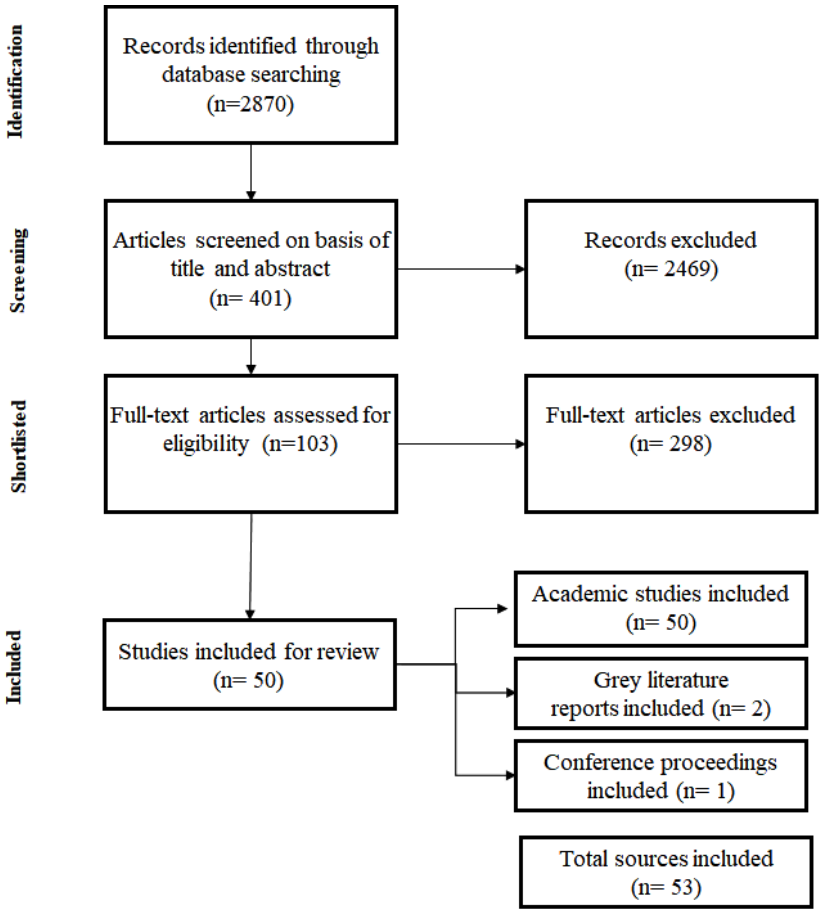
This systematic approach clarifies the origin of our conclusions, making it explicit that role transformation insights were derived from both the literature and our thematic analysis, supported by the matrix framework developed during the study.

Results

Results of Data Collection

The outcome of the journal searches yielded 2870 results, out of which 401 sources were shortlisted to be analyzed further, as demonstrated in Figure 3 and elaborated further in the subsequent section.

Figure 3. The systematic article selection process for this review.



Evolution of AI Approaches

Wave 1: Symbolic AI Approach

Expert systems and symbolic AI are 2 terms used to describe the initial wave of early AI approaches. In this case, human specialists develop exact rule-based processes, or "algorithms," that a computer may use to decide how to respond intelligently to a particular circumstance. A variation of this strategy called fuzzy logic allows for varying degrees of confidence in a scenario, which is helpful for capturing intuitive knowledge and enabling the algorithm to make wise judgments in the presence of numerous, uncertain, and interconnected factors. In contexts with rigorous rules and variables that are clear-cut and measurable, which do not vary significantly over time, symbolic AI performs well. These techniques may seem old, yet they are still used today [3].

Wave 2: Data-Driven AI Approach

The second wave of AI consists of more modern, "data-driven" methodologies that have advanced quickly over the past 2 decades and are primarily to blame for the present rebirth of AI. These do away with the first wave AI's reliance on human specialists by automating the learning of algorithms. Artificial neural networks (ANNs) are modeled after how the brain functions. The translation of inputs into signals that are then sent across a network of synthetic neurons to produce outputs that are seen as reactions to the inputs. ANNs can handle increasingly complicated issues by adding additional neurons and layers. An ANN with several layers is simply referred to as deep learning. ML is the process of changing a network so that its outputs are seen as useful or intelligent answers to its inputs. By using evolutionary concepts to produce slow improvements in huge populations of ANNs or by making gradual changes to individual ANNs, ML algorithms may automate this learning process [3].

Wave 3: Artificial General Intelligence Approach

The third wave of AI is a hypothetical term for potential future waves of AI. First and second wave approaches are referred to as "weak" or "narrow" AI in that they can act intelligently in just certain contexts and issue domains, whereas "strong" or "general" AI refers to algorithms that can act intelligently across a variety of contexts and problem domains. With existing technology, such artificial general intelligence is not feasible and would need paradigm-shifting development. Advanced evolutionary techniques, quantum computing, and brain emulation are a few possible strategies that have been considered. Although self-explanatory and contextual AI may have modest goals compared to other futuristic AI types, their potential influence and implementation challenges should not be understated [3].

AI Applications in Nursing

There is still much to learn about the innovative and intricate challenges surrounding AI. For health care businesses to best serve patients and physicians, AI must be fully used [8]. In subsequent sections, we have discussed examples of AI applications in nursing.

Rothman Index Use for Patient Acuity and Risk

The level of acuity and risk of a patient are both reflected by the Rothman Index. The electronic medical record (EMR) data connected to 26 variables, including 11 graphically represented nurse evaluation measures, are used to determine scores. The introduction of the Rothman Index was accompanied by doubts regarding its accuracy and dependability in delivering results that could be put into practice. At first, there was not enough peer-reviewed research on the technology to persuade nurses and other professionals that the outcomes would improve patient care. The capacity of nurses to affect patient care is crucial, as evidenced by a recent study indicating that the Rothman Index's performance is positively impacted by nurses' evaluation data [9].

Social or Companion Robots

Social robots are made to react to human interactions in a way that makes them human. Sophia (Hanson Robotics) is an illustration of a social robot designed as a companion for older adults that shows the possibility of technological developments to enhance how robots operate. Robots are being developed by researchers all across the world to enhance therapeutic telemedicine applications, reduce suicide rates, and more. The role of nurses in providing care will evolve as robots learn to carry out nursing tasks such as ambulation support, vital sign assessment, drug administration, and infectious disease procedures. According to research, nonnursing chores and activities take up between 8% and 16% of nursing time. With robot assistance, nurses will be able to reclaim this time and devote it to patients more. Does this imply that nursing is doomed to extinction? Absolutely not; in fact, the exact reverse is happening. Robots created and used for patient care and older adults' assistance are being developed by nurses. Nurses can receive assistance from the robots at the bedside or in the community [9].

Telerobots

Telerobots can support health care professionals who are at "high risk for infection due to routine patient contact, handling of contaminated materials, and challenges associated with safely removing protective gear." Furthermore, telepresence robots support nurse-led treatments for the promotion of healthy lifestyles and the management of chronic illnesses by combining an initial in-home visit to launch the health care program with subsequent remote telehealth visits made at the patient's home. Data are gathered on participant health outcomes as well as the robot intervention's usefulness and pleasure [9].

NLP Approaches

A language is a system of rules or a collection of symbols that are integrated and used to express ideas or disseminate information. NLP serves users who lack the time to learn new languages or become proficient in their current ones because not all users have a strong background in machine-specific language. In actuality, NLP is a branch of linguistics and AI whose goal is to enable computers to comprehend assertions and words spoken in human languages. It was developed to make the user's job easier and to fulfill their desire to speak to a computer in natural language. It can be divided into 2 categories: natural language generation and natural language understanding, which progresses the task of understanding and producing the text [10]. In nursing, NLP assists in nursing practice and decision-making. Using NLP, it is possible to analyze nursing records, spot patterns and trends in patient care, and gain knowledge that will enable nurses to give patients more individualized and effective treatment [11].

Robotic Process Automation

Robotic process automation (RPA) is a method that uses robotics as a set of techniques for the operation and use of automata (ie, robots) in the execution of multiple tasks in place of humans as the standard, method, or system. RPA results in the automatic execution of administrative, scientific, or industrial tasks. RPA tools are a set of methods intended to enhance productivity by automating and minimizing the number of repetitive jobs. The inclusion of AI algorithms and techniques to the use of RPA enhances the accuracy of the execution of automated procedures [12].

ML Algorithms

The study of algorithms and statistical models that computer systems use to carry out a particular task without being explicitly taught is known as ML. There are several daily-use programs that incorporate learning algorithms. One of the reasons a web-based search engine like Google works so well every time it is used to search the internet is because of a learning algorithm that has mastered the art of ranking websites. These algorithms are used for several different tasks, including data mining, image processing, and predictive analytics. The major benefit of ML is that once an algorithm understands how to use data, it can carry out its task autonomously [13].

The study of ML considers how to automatically generate reliable predictions from complicated data. It is strongly tied to contemporary statistics, and in fact, statisticians have contributed many of the most brilliant concepts in ML (eg, the lasso, trees, and forests). However, the ML community has been more focused on the single objective of maximizing predictive performance, in contrast to statisticians who have frequently concentrated on model inference—that is, knowing the parameters of their models (eg, testing on individual coefficients in a regression). "Out-of-sample" tests, which assess how well a model trained on 1 dataset would predict fresh data and serve as the benchmark for the whole ML discipline [14].

Application Case Studies

Overview

Relevant case studies and examples from the literature were used to highlight the involvement of nurses in the identified AI applications or tool deployments and to assess whether there is a change or anticipated elimination to the role of nurses in patient care. These case studies focused on certain types of AI applications that are tied to certain activities carried out by nurses as part of their core patient care role. The following examples were selected due to their applicability, importance, and contribution to the comprehension of the study's subject.

Rothman Index Use

The Specialized Workforce for Acute Transport (SWAT) team of nurses trained in critical care, advanced cardiovascular life support, and trauma care at Yale New Haven Hospital is a real-world example of using the Rothman Index technique. When signs point to a patient's condition deteriorating, they immediately receive alerts on their mobile phones. The SWAT team looks through the EMR, evaluates the patient as needed, and works together with clinical nurses and other medical personnel on pertinent areas of treatment. SWAT nurses identify as "a second pair of eyes" in their own description. The index's information came from widely available nursing literature. Given that the index is updated in real time from the EMR, timely submission of nurse assessment data is essential for the computation and value of the index score [9].

Social or Companion Robots

Sophia is an illustration of a social robot designed as a companion for older adults that shows the possibility of technological developments to enhance how robots operate. Sophia had a refurbishment in 2018 that included movement features, and she is currently the first robot to be granted citizenship in a nation (ie, Saudi Arabia) [9]. The next generation of social robots with cutting-edge AI is the LOVOT robot. The social robot LOVOT was well received by most patients with dementia. LOVOT exhibited beneficial impacts, improved communication, and promoted social engagement. Although LOVOT had no appreciable benefits on social well-being, it provided individuals with a break from daily living. Following their interactions with LOVOT, some residents experienced emotional overstimulation. The social robot was embraced by medical specialists and nurses, who saw LOVOT as a new tool for working with patients with dementia as a supporting tool and not as a replacement of the care provider's role [15].

A major factor in concentrating on the older adults is AI. It can, for instance, strengthen the bonds between older adults and their relatives or care teams. Furthermore, an AI chatbot can converse with the older adult without any difficulties and may remind them of important dates, such as medication intake and medical exams. Many of the AI smartphone apps available now have the ability to screen wellness data in a less intrusive manner, including daily activity, food, and, shockingly, older adults lifestyle choices. In certain situations, it could be helpful to anticipate and, thus, prevent any potential hypertension or irregular heart rate. In essence, robotic "pets" are also helping to improve patient attention while also assisting in the battle against emotions of loneliness. One such model, called Tombot, is a small, dog-like device designed to relieve anxiety in patients with dementia. Its head movements, appearance, and swinging tail are remarkably similar to those of the genuine dog, giving owners the impression that they have their own pet to truly concentrate on. The care of the older adults is one of the problems that low-income nations are experiencing. The global shift of older populations has worsened the shortage of trained people in the older adults health care context. Given that the number of older people worldwide is predicted to almost treble in the next 3 decades, there may be a greater need for older adult care [16].

Telepresence Robots

Health care professionals who are at "high risk for infections due to routine patient contact, handling of contaminated materials, and challenges associated with safely removing protective gear" are the focus of Tele-Robotic Intelligent Nursing Assistant, a remote-controlled robot, at Duke University Pratt School of Engineering and School of Nursing. Noting that Tele-Robotic Intelligent Nursing Assistant is 20 times slower than a nurse, it presently completes around 60% of the preset nursing tasks in the nursing simulation laboratory where it is being evaluated. Results from individuals getting telehealth coaching from home reveal that patients and clinicians alike find satisfaction in the mix of live face-to-face interventions and robotic telehealth visits. Designing meaningful treatments that can take use of new technology requires the nurse to have a key part in the development and execution of telehealth robots [9].

Inpatient rooms at the nonprofit, tertiary, 958-bed Cedars-Sinai Hospital in Los Angeles, California, are equipped with Alexa robots created by Amazon to serve as virtual nurse aides. To support patients with their daily routines, Alexa fulfills the monotonous activities performed by nurses. She also assists in answering medical queries and reminds patients to take their pills on time [17].

NLP Use

The most commonly used AI functions in studies of AI-related nursing activities were profiling and prediction, followed by assessment and evaluation. Virtual reality teaching interventions and learning successes were beneficial to nurses because they provided a safe learning environment with the possibility of multiple tries, overcoming challenges, the ability to consolidate knowledge, and professional efficacy [18]. Furthermore, the use of chatbots improves student learning compared to traditional teaching techniques [19], acting as supporting tools to nursing educators rather than eliminating the entirety of their role.

NLP is used by triage nurses to register and categorize patients based on their speech. When conducting triage activities, RMIS-AI is quicker than using the manual input approach, which decreases the time it takes to register patients and classify them. To address the existing level of subpar sensitivity and accuracy provided by nurses, technological augmentation is necessary [20].

Primary care nurses are faced with increasing demands from patients who have wounds from a variety of sources. Both nurses and patients can benefit from a chatbot that provides information properly verified on the basis of evidence. By providing instructions on the suggested wound dressing techniques for each type of wound, BOTCURATIVO, a chatbot, seeks to assist nonspecialists in the management of wounds. A reasonable degree of content validity was attained by the script that was created and implemented into the chatbot prototype. The chatbot's usability was seen as being good, which increased the device's credibility. Noting that regardless of their specialty, the nurse will always undertake wound management tasks [21].

Most people lack the medical knowledge necessary to investigate or understand the severity of their conditions or symptoms. In this regard, NLP is essential to health care. These chatbots gather health information from patients, analyze it, and recommend actions to patients based on more pertinent knowledge of their physical conditions. Health care chatbots similar to NOVA—a virtual nursing assistant driven by AI—are helpful in the medical field because they help patients and point them in the direction of the right resources. When consumers or patients look up answers to inquiries they have about their health on the internet, chatbots are more helpful. A user of this program may text requests for health care and may receive pertinent health advice in return. A chatbot can provide medical information, including illness symptoms and treatment options. Patients receive professional guidance in real time, and their personal and medical data are kept in a database for future research. The number of AI-powered health care apps has significantly increased recently. Consequently, there are shorter wait times in offices, which saves money and energy. Patients may be helping in their own place and at their own speed while learning medical knowledge. User input is received by the system via text or speech data. The system interprets the input data. The virtual nursing help system may be accessed by the user who can also send an inquiry to it. The output that the system produces is a list of user symptoms and suggested diagnoses. In the area of virtual nursing help, the suggested system serves as the user's personal assistant. The created bots are useful for keeping track of patient information. The technology can also help numerous people at once [22].

In MobiGuide, the role of nurses in creating, providing, and assessing eHealth-based services was examined with an emphasis on atrial fibrillation home monitoring. To obtain suggestions, warnings, and reminders about drugs and measures that they should conduct, patients were given smartphones and electrocardiogram sensors. This mobile decision support system was regularly updated by a backend system. Health care professionals are supplied patient data so they may view it and take appropriate action. With their participation in the design of the caregiver interface, responsibility for the enrollment phase (ie, including patient training), daily data checks, triage of patient concerns, and patient interviews about their experiences with the system, nurses play a key role in such settings [23].

The Smart Wearable Physiological Signal Measurement Integration System is used in home care, nursing homes, and other health care settings to continuously monitor their patients' vital signs, which enables nurses to see early warning indicators of deterioration and take quick action to stop unfavorable outcomes. When a patient exits a specified area or when vital signs suggest an emergency, the system may notify the care providers, improving patient safety in nursing homes and home care settings [24].

The development of indoor positioning technology has made it feasible to track the movement of mobile medical equipment within a hospital ward, including patient monitors and electrocardiography devices. Nurses can quickly detect and locate a gadget with the help of an item tracking system, particularly while they are getting ready for a medical procedure or shift change. Given that nurses typically have a heavy workload, it would be ideal to give them access to a well-liked mobile app with an intuitive search interface that they can use on a regular basis. To help with this, DBOS, a dialogue-based object query system, offers voice and text inquiry services to nurses while mimicking a genuine discussion with users through the chatbot interface of the Line messaging app [25].

The accurate assessment of pain in the neonatal intensive care units is essential due to the high prevalence of painful experiences. Video-based assessment of neonatal pain could be reliably used, as confirmed by the high intrarater and interrater reliability between direct observation and the video-based assessment, as well as the AI method-based performance evaluation, even with various disturbances in real-world neonatal intensive care units. Video-based assessment is viable for neonatal pain assessment in a clinical setting, and the extent of neonatal pain can be evaluated remotely in real time, which can better identify and treat it and thus improve the neonatal pain condition. Video-Based Neonatal Pain Assessment can reduce the stressful surroundings of a clinical setting, the contextual noise, and other elements that could shift the focus of the trainees from the rating. There has been an increasing interest in using ML methods for understanding human behavioral responses to pain based on the analysis of facial expressions, crying sounds, and body movement. Several automated methods have been introduced to automatically assess infants' pain based on behavioral or physiological pain indicators analysis. Using AI-based neonatal pain assessment, the nursing staff can also use these recordings to judge the pain level by observing the painful procedure video in the nurse station and taking timely intervention measures, which could greatly reduce the bedside observation time and improve work efficiency [26].

Robotic Process Automation

Patients with diabetes mellitus face a 15% to 25% lifetime risk of developing diabetic foot ulcers (DFUs). Monitoring and assessing DFUs for complications and healing progress is essential, and this was traditionally performed using manual measurements. A past study compared conventional measurement methods with an AI-powered mobile application for wound imaging, the CARES4WOUNDS (C4W) system [27]. The length and breadth of the wound were the major characteristics measured. C4W measures had good intra- and interrater reliability compared to standard wound measuring. The C4W was a helpful tool for keeping track of DFU wound healing, yet it did not eliminate the role of wound care nurses.

Machine Learning

In 1 in 8 to 10 cases where primary care physicians and nurse practitioners used AI, they made better diagnoses, suggesting the potential for raising the standard of dermatologic treatment. The diagnoses showed improvements of 10% and 12% for primary care physicians and nurse practitioners, respectively, indicating a significant positive impact [28].

For long-term patient care, Vitalerter has developed a program that combines advanced biosensors and deep learning to provide contactless and continuous vital sign monitoring, as well as cloud-based early warning protection services. Some of the standout features of these systems are accurate body movement analysis, continuous heart and respiratory rate monitoring, and contactless detection of patients moving out of bed. In the event of an adverse event, the system will automatically sound an alarm to remind nurses to take immediate action and lower the risk of falls, pressure sores, and septicemia [29].

Converting pediatric nursing diagnoses into a digital format and adding them to a case base to evaluate how well the prototype handled these cases allowed for case comparison, retrieval, adaptation, and indexing. Therefore, this study offers a computational tool for the health sector that makes use of case-based reasoning, an AI method. While case-based reasoning is merely another paradigm for problem solving, what sets it apart from other AI approaches is how it differs from them. Rather than relying just on a general understanding of the issue or creating connections between problem descriptions and conclusions, this paradigm can use specific information from past experiences or real problem situations. It is acknowledged that using nursing care systematization necessitates that nurses develop a variety of abilities and adhere to theoretical support to enhance decision-making. Decisions should then be discussed with the patient whenever feasible. The application of these records or technology in various clinical health situations, in which observations about the care needs of patients accompany the decision-making process about the care provided, assists in the subsequent evaluation of the outcomes that are obtained with professional intervention. In this way, it is known that the nursing care systematization collaborates to provide safe, logical, and effective nursing care. Organizing the administration of nursing care and assisting nurses in making decisions is also predicated on ensuring patient safety at various care levels [30].

Table 3 summarizes the thematic analysis and links the respective case studies outlined in this research paper.

On the basis of the themes outlined from the literature review, 81% (13/16) of the AI applications within the nursing fields are in the proof-of-concept phase, with 19% (3/16) of those deployed demonstrating a positive impact on the nursing role within the patient’s journey with the United States leading the way in such research and developments. Furthermore, applications that would enhance or streamline the nurses’ role seem to be focused on the treatment stage, followed by 25% on posttreatment care (ie, recovery). Noting that the applications cover various aspects of the nursing activities from diagnosis, treatment, wound management, education, and training to triaging.

Multimedia Appendix 1 [1-53] provides a clear summary of our systematic review by analyzing 37 sources in terms of key findings, methodology, sample size, potential biases, and validity. This is to ensure the robustness and reliability of the conclusions drawn from the systematic review.

Table 3. Outcome of the thematic analysis.

AI[a] application in nursing and country	Name	Nursing involvement	Status
Rothman Index			
United States (Yale New Haven Hospital)	Rothman Index	Treatment	Operationally deployed
Social or companion robots			
Saudi Arabia	LOVOT	Posttreatment care (ie, recovery)	Operationally deployed
United States	Tombot	Posttreatment care (ie, recovery)	POC[b]
Telepresence robots			
United States (Duke University Pratt School of Engineering and School of Nursing)	TRINA[c]	Treatment	POC
United States (Cedars-Sinai Hospital)	Alexa	Treatment	Operationally deployed
NLP[d]			
United States	RMIS-AI	Triage	POC
United States	BOTCURATIVO	Nurse education and training	POC
United States	NOVA-a virtual nursing assistant	Diagnosis	POC
United States	MobiGuide	Posttreatment care (ie, recovery)	POC
United States	Smart Wearable Physiological Signal Measurement Integration System	Posttreatment care (ie, recovery)	POC
United States	DBOS[e], a dialogue-based object query system	Treatment	POC
United States	VB-AI[f] NPA	Treatment	POC
RPA[g]			
Singapore	CARES4WOUNDS system, Tet-suyu	Wound management	POC
Machine learning			
United States	Artificial intelligence aid	Diagnosis	POC
United States	Vitalerter vital sign monitoring	Treatment	POC
United States	CBR[h]	Treatment	POC

[a]AI: artificial intelligence.
[b]POC: proof of concept.
[c]TRINA: Tele-Robotic Intelligent Nursing Assistant.
[d]NLP: natural language processing.
[e]DBOS: dialogue-based object query system.
[f]VB-AI: video-based artificial intelligence.
[g]RPA: robotic process automation.
[h]CBR: case-based reasoning.

Roles of Nurses and Role Transformation

The integration of AI technologies into health care has significantly transformed the roles of nurses, shifting their focus from routine tasks to more advanced and patient-centered care [6]. AI systems automate many traditional nursing responsibilities, such as monitoring patient vitals, data entry, and medication management, allowing nurses to prioritize clinical decision-making, patient education, and emotional support. This role transformation not only enhances the efficiency of health care delivery but also enables nurses to engage more deeply in patient care by using AI as a collaborative tool [31]. AI-driven systems support clinical decision-making, triaging, and diagnostic processes, leading to improved patient outcomes and job satisfaction among nurses [32]. Table 4 provides an overview of how AI can transform nursing roles across various functions.

Table 4. Traditional nurse role versus artificial intelligence (AI)–driven role transformation.

Traditional nurse role	AI-driven role transformation	Example of AI technology involved	Impact on patient care	Academic reference
Monitoring patient vital signs	AI takes over continuous monitoring, alerting nurses only when intervention is needed	AI-based monitoring systems (eg, wearable sensors and IoT[a] devices)	Frees up nurses' time for more personalized, hands-on patient care and reduces error risk through automation	Ross et al [37]
Data entry and record keeping	AI automates data entry, streamlining the EHR[b] updating process	AI-enabled EHR systems with NLP[c]	Reduces administrative burden, allowing nurses to focus on direct patient care	Zou and Schiebinger [33]
Routine diagnostic procedures	Nurses assist in AI-driven diagnostics, focusing more on patient interaction and explaining results	AI diagnostic tools (eg, image analysis for radiology and pathology)	Enhances the role of nurses as educators, helping patients understand diagnoses and treatments	Ng et al [38]
Medication administration	AI systems manage medication scheduling, and dosing, with nurses overseeing AI-generated plans	Automated dispensing systems and AI-driven dose calculators	Reduces medication errors and ensures timely administration, allowing nurses to focus on patient observation	Shang [39]
Patient triage and assessment	AI aids in triaging by prioritizing patients based on real-time data, allowing nurses to focus on high-priority cases	AI-powered triage systems (eg, in emergency departments)	Increases efficiency in patient care and enhances the accuracy of triage decisions	Govindaraj et al [40]
Clinical decision support	Nurses collaborate with AI systems that provide real-time decision support based on predictive analytics and historical data	AI-based decision support systems (eg, IBM Watson and AI in ICU[d] for risk prediction)	Empowers nurses to contribute more significantly to clinical decision-making and patient care planning	El-Kareh and Sittig [41]
Health education and counseling	AI tools provide nurses with real-time personalized health data to tailor patient education more effectively	AI-driven patient education platforms (eg, AI chatbots and personalized health apps)	Enhances the nurse's ability to deliver personalized health education and counseling based on real-time insights	Li et al [42]
Supervision of junior staff	Nurses oversee AI-driven workflows and ensure that AI-generated protocols are followed, focusing more on clinical mentorship	AI systems for task delegation and workflow automation	Enhances leadership roles, allowing nurses to take on a supervisory role and focus on mentorship and training	Rony et al [43]
Wound care and management	AI tools help nurses monitor wound healing through image analysis and predictive algorithms	AI-based wound care imaging systems (eg, predictive models for healing times)	Improves the accuracy of wound assessment, reduces manual checks, and improves patient outcomes	Rippon et al [44]
Patient discharge planning	AI assists in generating discharge plans, predicting postdischarge risks, and automating referrals to follow-up care systems	AI-driven discharge planning tools	Optimizes discharge planning and postdischarge care, reducing the likelihood of readmissions	Jack et al [45]
Emotional support and communication	AI systems can handle administrative tasks, enabling nurses to spend more time on patient emotional support and communication	AI-powered administrative assistants (eg, scheduling systems, automated communication)	Allows nurses to prioritize emotional support and patient communication over routine tasks	Robert [9]

[a]IoT: Internet of Things.

[b]EHR: electronic health record.

[c]NLP: natural language processing.

[d]ICU: intensive care unit.

Recent studies support these findings, showing that AI systems can help optimize workflows, reduce administrative burdens, and allow nurses to contribute more meaningfully to clinical care. Health care AI tools, such as predictive analytics and automated documentation systems, have been shown to improve patient outcomes while minimizing the risk of human error in routine tasks [33,34]. Moreover, AI-based decision support tools in critical care environments enable nurses to make informed decisions quickly, positively impacting patient care quality [35]. These advancements are particularly evident in the transformation of nursing roles, as evidenced in a thematic analysis of health care AI implementations [36].

Critical Assessment of the Literature

The literature collectively covers a broad spectrum of AI applications, ranging from technical reviews and policy implications to specific domains such as health care and nursing. The mix of older foundational papers and recent studies provides both historical context and insights into current advancements. Practical and policy-oriented papers enhance the literature by addressing real-world applications and implications of AI. However, some biases were identified, particularly in policy reports like the one by Boucher [3], which reflect institutional viewpoints. The focus on health care and nursing in several papers could skew the overall perspective toward these fields. In addition, journals with lower impact factors might have less rigorous peer review processes, potentially affecting research quality.

The synthesis of findings indicates a strong direction toward integrating AI in various fields, particularly health care and nursing. There is a clear emphasis on the transformative potential of AI, along with discussions on challenges and ethical considerations. Comparative studies and reviews highlight the advantages and limitations of different AI approaches, suggesting the need for context-specific solutions. The quality and diversity of the studies imply that AI is a rapidly evolving field with significant interdisciplinary impacts. Practical guides and policy reports emphasize the need for continuous education and ethical considerations in AI deployment. The focus on health care underscores AI's potential to improve patient outcomes, though it also highlights the importance of rigorous evaluation and context-specific applications.

In conclusion to this section, the systematic literature review provides a comprehensive overview of AI, balancing theoretical foundations, recent advancements, practical applications, and policy implications. While some sources may carry biases or lack depth in certain areas, the collective insights offer valuable guidance for understanding AI's multifaceted impact as outlined in Multimedia Appendix 2 [1-53].

AI in Nursing From a Theory and Management Perspective

Numerous theoretical and managerial contexts have debated the use of AI in health systems in general and in tasks associated with the nurses' role in particular. As an overarching theoretical viewpoint, nurses will continue to provide direct patient care due to nuances in human behavior. The ability to incorporate new tools and technology will be required of nurses. As technology is being incorporated into nursing programs' curricula, the nursing profession is evolving. Thus, from a management viewpoint, nurses will continue to integrate the data produced by AI tools. They will need to have the skills to incorporate AI findings into evidence-based practice while combining that knowledge with nursing expertise [9].

Despite limitations in identifying numerous pieces of the literature that address the impact of deploying AI, given that many tools and techniques are in project or testing stages, our research contributes to current discussions on contextualized research.

Through this systematic literature review, we attempted to establish a foundation to identify the existing studies from the context emic perspective. By encouraging the research community to focus on "optimal allocation of effort between exploitation and exploration," looking at theoretical contributions from the periphery will progress management and organization science [46]. We encourage academics to perform empirical studies for the benefit of advancing literature in this arena. From a practical perspective, physicians, nurses, ML scientists, and hospital and clinical executive administrators when designing their clinical pathways could use this research when designing their treatment plans [47]. Furthermore, academics in the field of medicine, nursing, paramedics, hospital executive administration, patient access, and information technology will benefit from this systematic review as it allows them to build on the existing relevant literature.

Discussion

Principal Findings

The evolution of AI in nursing has transitioned from early symbolic AI, using rule-based algorithms and fuzzy logic, to modern data-driven approaches such as ML and ANNs and is now exploring hypothetical future waves such as artificial general intelligence. AI applications in nursing include the Rothman Index for patient acuity and risk assessment, social robots such as Sophia and LOVOT for older adults' companionship, telerobots for remote patient interaction, and NLP for enhancing decision-making and patient communication. RPA and ML are used to automate repetitive tasks and improve diagnostic accuracy, while AI-powered tools such as chatbot assistants and wearable monitoring systems assist in patient care and safety. Case studies demonstrate AI's role in supporting, rather than replacing, nursing functions, enhancing efficiency, and allowing nurses to focus more on direct patient care. The success or failure of the medical AI solution will depend on how closely system architects collaborate with real-world nurses in health care fields, as they are needed to work closely together to assess and evaluate which technologies will be prioritized for development [17].

Inadequate evaluation, careless supervision, a lack of fundamental nursing knowledge, a lack of service awareness, and unlawful activity by nursing personnel are all contributing factors to poor nursing care. Inadequate evaluation forces nursing staff members to advance their own skills, be able to analyze certain nursing conditions, and act quickly to take timely, scientifically sound action. Poor nursing is also largely caused by a lack of thorough inspection by nursing staff. Individuals who make nursing errors because of inadequate nursing knowledge should enhance their own skills and training. Ultimately, the nursing profession is a service sector. The essential spirit of service is required when treating patients. It is imperative that corresponding services are rendered completely in compliance with industry standards, and any illicit activities are forbidden. To address this, the use of AI alone will not mitigate those issues; instead, the health care facility's relevant departments must develop a mechanism to penalize slack investigation and prevent the recurrence of such unfavorable circumstances [48].

Methodological Approach Limitations

The limitations of our systematic literature review methodological approach include potential publication bias;

additionally, the quality and relevance of included studies can vary, impacting the overall reliability of the findings. The search strategy may also be limited by the databases and sources selected, potentially missing relevant literature elsewhere. Furthermore, the exclusion of non–English language studies might introduce language bias. Finally, the subjective nature of data extraction and thematic analysis can lead to inconsistencies and affect the validity of the conclusions.

Future Directions and Recommendations

Due to research that is still in the early stages of development and the considerable variation in AI types and situations, AI used in health care and nursing care is still a developing practice with minimal evidence. AI in nursing becomes a crucial component of health care delivery in general and nursing practice in particular. There is still a great deal of room for advancement with these systems in terms of ensuring not only the professional autonomy of nurses but also better access to sources of health information to maximize their use in multitasking, to cover the greatest number of factors that may affect the patient, environment, clinical practice, and various medical services. Additional research is required to determine how previous research findings using AI-based systems with virtual reality or simulated scenarios can be applied to real-world clinical nursing practice or to examine how these AI-based support systems may enhance patient safety and help nurses in specific clinical settings [28]. A blueprint of nurse involvement in the deployment of AI-based systems and applications can act as a guiding reference and is an area that is worth further research and exploration.

AI has become a game-changing technology that is transforming several industries, the health care industry most notably. It is essential for diagnosing uncommon genetic diseases, streamlining patient care in mental health clinics, supporting clinical judgment, and revolutionizing pathological research. However, the growing use of AI in health care also raises difficult moral, practical, and legal questions, especially in light of the General Data Protection Regulation framework in Europe. The significance of understanding data owner rights and developing moral guidelines for AI use in medical applications, particularly nursing, is another area for future research. Comprehending the ethical discussion around AI helps health care and nursing professionals create moral AI procedures for practice and assists in navigating the complex landscape of AI-driven health care regulations, ethical issues, and data protection [49]. AI presents several risks, particularly in the context of deep reinforcement learning–based mobile robot assistants. Ensuring safety in environments where humans and robots interact is crucial, especially when autonomous mobility robots rely on deep reinforcement learning for navigation and decision-making. This is particularly important in health care settings, where hospital patients using these AI-driven mobility assistants may face potential hazards that require careful evaluation and mitigation [50]. Therefore, there is an opportunity to further research the risks associated with the use of AI in nursing.

Conclusions

To fully benefit from AI technology, nurses will need to develop their ability to collaborate with data scientists. Although computer science and nursing are two separate fields, knowledge and skill transfer between the two is crucial as technology develops so that nurses may learn to interpret the data. In the future, nurses will play the role of coaches who will assist people in managing their health and achieving better results. The provision of touch and building connections with patients are the foundations of the nursing profession and their role in patient care, and they will never be fully replaced by AI tools or robots, especially when collecting medical information, such as heart monitoring, urinalysis, and range-of-motion analysis [9].

As emerging AI technologies take over some of the jobs that nurses already do, nursing will be impacted. Although technology will alter the way nurses spend their time providing patient care, nurses will still be required. The nurse will acquire new ways of thinking about and processing information; they will become information integrators, health coaches, and providers of human care, assisted by AI technologies rather than being replaced by them [9]. Current research and implementations demonstrate the effectiveness and promise of AI in nursing practice. However, they do not eliminate the need for field supervision and emotional support from humans [51]. Therefore, and to answer the research question, "Will AI change the role of nurses in patient care?" based on the outcome of this literature review, the answer is "yes" with a varying extent depending on the AI tool in use by the nursing professionals.

Interest in incorporating AI into nursing practice will not go away, although its technological potential is not well understood. It is highly recommended that academic institutions and professional associations implement suitable educational and training initiatives. It is imperative that nurses enhance their comprehension of fundamental AI and its integration into nursing practice [52].

Acknowledgments

The work in this study was supported, in part, by the Open Access Program from the American University of Sharjah, United Arab Emirates.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

IAK was responsible for writing and analyzing the study. MN reviewed the content for accuracy and clarity.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Systematic review analysis.
[XLSX File (Microsoft Excel File), 15 KB - nursing_v8i1e63335_app1.xlsx]

Multimedia Appendix 2
Critical assessment of the literature.
[XLSX File (Microsoft Excel File), 15 KB - nursing_v8i1e63335_app2.xlsx]

Multimedia Appendix 3
PRISMA Checklist.
[PDF File (Adobe PDF File), 287 KB - nursing_v8i1e63335_app3.pdf]

References

1. Tecuci G. Artificial intelligence. WIREs Comput Stat 2011 Dec 07;4(2):168-180. [doi: 10.1002/wics.200]
2. Liu J, Kong X, Xia F, Bai X, Wang L, Qing Q, et al. Artificial intelligence in the 21st century. IEEE Access 2018 Mar 26;6:34403-34421. [doi: 10.1109/access.2018.2819688]
3. Boucher PN. Artificial intelligence: how does it work, why does it matter, and what can we do about it? European Parliament. 2020 Jun 28. URL: https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)641547 [accessed 2025-02-04]
4. Mohammad SM. Artificial intelligence in information technology. SSRN Electron J 2020 Jun [FREE Full text] [doi: 10.2139/ssrn.3625444]
5. Topaz M, Murga L, Gaddis KM, McDonald MV, Bar-Bachar O, Goldberg Y, et al. Mining fall-related information in clinical notes: comparison of rule-based and novel word embedding-based machine learning approaches. J Biomed Inform 2019 Feb;90:103103 [FREE Full text] [doi: 10.1016/j.jbi.2019.103103] [Medline: 30639392]
6. 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 [FREE Full text] [doi: 10.2196/23939] [Medline: 34406963]
7. Yang Q, Steinfeld A, Zimmerman J. Unremarkable AI: fitting intelligent decision support into critical, clinical decision-making processes. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI '19; May 4-9, 2019; Glasgow, UK. [doi: 10.1145/3290605.3300468]
8. McGrow K. Artificial intelligence: essentials for nursing. Nursing 2019 Sep;49(9):46-49 [FREE Full text] [doi: 10.1097/01.NURSE.0000577716.57052.8d] [Medline: 31365455]
9. Robert N. How artificial intelligence is changing nursing. Nurs Manage 2019 Sep;50(9):30-39 [FREE Full text] [doi: 10.1097/01.NUMA.0000578988.56622.21] [Medline: 31425440]
10. Khurana D, Koli A, Khatter K, Singh S. Natural language processing: state of the art, current trends and challenges. Multimed Tools Appl 2023;82(3):3713-3744 [FREE Full text] [doi: 10.1007/s11042-022-13428-4] [Medline: 35855771]
11. Mitha S, Schwartz J, Hobensack M, Cato K, Woo K, Smaldone A, et al. Natural language processing of nursing notes: an integrative review. Comput Inform Nurs 2023 Jun 01;41(6):377-384. [doi: 10.1097/CIN.0000000000000967] [Medline: 36730744]
12. Ribeiro J, Lima R, Eckhardt T, Paiva S. Robotic process automation and artificial intelligence in industry 4.0 – a literature review. Procedia Comput Sci 2021;181:51-58. [doi: 10.1016/j.procs.2021.01.104]
13. Mahesh B. Machine learning algorithms - a review. Int J Sci Res 2020 Jan;9(1):381-386. [doi: 10.21275/ART20203995]
14. Taddy M. The technological elements of artificial intelligence. National Bureau of Economic Research. 2018 Feb. URL: https://www.nber.org/system/files/working_papers/w24301/w24301.pdf [accessed 2025-02-04]
15. Dinesen B, Hansen HK, Grønborg GB, Dyrvig AK, Leisted SD, Stenstrup H, et al. Use of a social robot (LOVOT) for persons with dementia: exploratory study. JMIR Rehabil Assist Technol 2022 Aug 01;9(3):e36505 [FREE Full text] [doi: 10.2196/36505] [Medline: 35916689]
16. Rong J, Ji X, Fang X, Jee MH. Research on Material Design of Medical Products for Elderly Families Based on Artificial Intelligence. Appl Bionics Biomech 2022;2022:7058477 Retracted in: Appl Bionics Biomech. 2023 Nov 1:2023:9763260. [doi: 10.1155/2023/9763260] [Medline: 37946829] [FREE Full text] [doi: 10.1155/2022/7058477] [Medline: 35087604]

17. Lee D, Yoon SN. Application of artificial intelligence-based technologies in the healthcare industry: opportunities and challenges. Int J Environ Res Public Health 2021 Jan 01;18(1):271 [FREE Full text] [doi: 10.3390/ijerph18010271] [Medline: 33401373]
18. Hwang GJ, Chang PY, Tseng WY, Chou CA, Wu CH, Tu YF. Research trends in artificial intelligence-associated nursing activities based on a review of academic studies published from 2001 to 2020. Comput Inform Nurs 2022 Dec 01;40(12):814-824. [doi: 10.1097/CIN.0000000000000897] [Medline: 36516032]
19. Chang CY, Hwang GJ, Gau ML. Promoting students' learning achievement and self - efficacy: a mobile chatbot approach for nursing training. Br J Educ Technol 2022 Jan;53(1):171-188. [doi: 10.1111/bjet.13158]
20. Cho A, Min IK, Hong S, Chung HS, Lee HS, Kim JH. Effect of applying a real-time medical record input assistance system with voice artificial intelligence on triage task performance in the emergency department: prospective interventional study. JMIR Med Inform 2022 Aug 31;10(8):e39892 [FREE Full text] [doi: 10.2196/39892] [Medline: 36044254]
21. da Silva Lima Roque G, Roque de Souza R, Araújo do Nascimento JW, de Campos Filho AS, de Melo Queiroz SR, Ramos Vieira Santos IC. Content validation and usability of a chatbot of guidelines for wound dressing. Int J Med Inform 2021 Jul;151:104473. [doi: 10.1016/j.ijmedinf.2021.104473] [Medline: 33964703]
22. Bidve V, Virkar A, Raut P, Velapurkar S. NOVA-a virtual nursing assistant. Indones J Electr Eng Comput Sci 2023 Apr;30(1):307-315. [doi: 10.11591/ijeecs.v30.i1.pp307-315]
23. Parimbelli E, Sacchi L, Budasu R, Napolitano C, Peleg M, Quaglini S. The role of nurses in e-health: the MobiGuide project experience. Stud Health Technol Inform 2016;225:153-157. [Medline: 27332181]
24. Wang WH, Hsu WS. Integrating artificial intelligence and wearable IoT system in long-term care environments. Sensors (Basel) 2023 Jun 26;23(13):5913 [FREE Full text] [doi: 10.3390/s23135913] [Medline: 37447763]
25. Chu ET, Huang ZZ. DBOS: a dialog-based object query system for hospital nurses. Sensors (Basel) 2020 Nov 19;20(22):6639 [FREE Full text] [doi: 10.3390/s20226639] [Medline: 33228178]
26. Chen X, Zhu H, Mei L, Shu Q, Cheng X, Luo F, et al. Video-based versus on-site neonatal pain assessment in neonatal intensive care units: the impact of video-based neonatal pain assessment in real-world scenario on pain diagnosis and its artificial intelligence application. Diagnostics (Basel) 2023 Aug 12;13(16):2661 [FREE Full text] [doi: 10.3390/diagnostics13162661] [Medline: 37627921]
27. Chan KS, Chan YM, Tan AH, Liang S, Cho YT, Hong Q, et al. Clinical validation of an artificial intelligence-enabled wound imaging mobile application in diabetic foot ulcers. Int Wound J 2022 Jan;19(1):114-124 [FREE Full text] [doi: 10.1111/iwj.13603] [Medline: 33942998]
28. Jain A, Way D, Gupta V, Gao Y, de Oliveira Marinho G, Hartford J, et al. Development and assessment of an artificial intelligence-based tool for skin condition diagnosis by primary care physicians and nurse practitioners in teledermatology practices. JAMA Netw Open 2021 Apr 01;4(4):e217249 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.7249] [Medline: 33909055]
29. Lu ZX, Qian P, Bi D, Ye ZW, He X, Zhao YH, et al. Application of AI and IoT in clinical medicine: summary and challenges. Curr Med Sci 2021 Dec;41(6):1134-1150 [FREE Full text] [doi: 10.1007/s11596-021-2486-z] [Medline: 34939144]
30. Alazzam MB, Tayyib N, Alshawwa SZ, Ahmed MK. Nursing care systematization with case-based reasoning and artificial intelligence. J Healthc Eng 2022 Mar 9;2022:1959371 [FREE Full text] [doi: 10.1155/2022/1959371] [Medline: 35310193]
31. 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 Jul 26;2023:1-12. [doi: 10.1155/2023/3219127]
32. Rony MK, Akter K, Debnath M, Rahman MM, Johra FT, Akter F, et al. Strengths, weaknesses, opportunities and threats (SWOT) analysis of artificial intelligence adoption in nursing care. J Med Surg Public Health 2024 Aug;3:100113. [doi: 10.1016/j.glmedi.2024.100113]
33. Zou J, Schiebinger L. Ensuring that biomedical AI benefits diverse populations. EBioMedicine 2021 May;67:103358 [FREE Full text] [doi: 10.1016/j.ebiom.2021.103358] [Medline: 33962897]
34. Almagharbeh WT. The impact of AI-based decision support systems on nursing workflows in critical care units. Int Nurs Rev 2024 Jul 08. [doi: 10.1111/inr.13011] [Medline: 38973347]
35. Seibert K, Domhoff D, Bruch D, Schulte-Althoff M, Fürstenau D, Biessmann F, et al. Application scenarios for artificial intelligence in nursing care: rapid review. J Med Internet Res 2021 Nov 29;23(11):e26522 [FREE Full text] [doi: 10.2196/26522] [Medline: 34847057]
36. Rony MK, Kayesh I, Bala SD, Akter F, Parvin MR. Artificial intelligence in future nursing care: exploring perspectives of nursing professionals - a descriptive qualitative study. Heliyon 2024 Feb 08;10(4):e25718 [FREE Full text] [doi: 10.1016/j.heliyon.2024.e25718] [Medline: 38370178]
37. Ross A, Freeman R, McGrow K, Kagan O. Implications of artificial intelligence for nurse managers. Nurs Manage 2024 Jul 01;55(7):14-23. [doi: 10.1097/nmg.0000000000000143] [Medline: 38951725]
38. Ng ZQ, Ling LY, Chew HS, 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]
39. Shang Z. A concept analysis on the use of artificial intelligence in nursing. Cureus 2021 May 05;13(5):e14857 [FREE Full text] [doi: 10.7759/cureus.14857] [Medline: 34113496]

40. Govindaraj M, D AK, Khan P, Krishnan R, Gnanasekaran C, Lawrence J. Revolutionizing healthcare: the transformative impact of artificial intelligence. In: Revolutionizing the Healthcare Sector with AI. Hershey, PA: IGI Global; 2024.
41. El-Kareh R, Sittig DF. Enhancing diagnosis through technology: decision support, artificial intelligence, and beyond. Crit Care Clin 2022 Jan;38(1):129-139 [FREE Full text] [doi: 10.1016/j.ccc.2021.08.004] [Medline: 34794627]
42. Li YH, Li YL, Wei MY, Li GY. Innovation and challenges of artificial intelligence technology in personalized healthcare. Sci Rep 2024 Aug 16;14(1):18994 [FREE Full text] [doi: 10.1038/s41598-024-70073-7] [Medline: 39152194]
43. Rony MK, Parvin MR, Ferdousi S. Advancing nursing practice with artificial intelligence: enhancing preparedness for the future. Nurs Open 2024 Jan;11(1):10.1002/nop2.2070 [FREE Full text] [doi: 10.1002/nop2.2070] [Medline: 38268252]
44. Rippon MG, Fleming L, Chen T, Rogers AA, Ousey K. Artificial intelligence in wound care: diagnosis, assessment and treatment of hard-to-heal wounds: a narrative review. J Wound Care 2024 Apr 02;33(4):229-242. [doi: 10.12968/jowc.2024.33.4.229] [Medline: 38573907]
45. Jack BW, Austad K, Renfro DR, Mitchell S. Re-engineering the hospital discharge to improve the transition from hospital to home: overview and a look to the future. J Healthc Manag Standard 2023;3(1):1-17. [doi: 10.4018/JHMS.328775]
46. March JG. Exploration and exploitation in organizational learning. Organ Sci 1991 Feb 01;2(1):71-87. [doi: 10.1287/orsc.2.1.71]
47. Blakemore A, Stephenson C. Psychological wellbeing practitioners: an opportunity for new ways of working in occupational health. Occup Health Work 2017;14(2):27-30 [FREE Full text]
48. Han J, Li D, Guo C, Wang J, Xue S. Construction and reliability and validity test of home care assessment scale for elderly patients with chronic diseases based on intelligent medical care. Mob Inf Syst 2022 Jul 13;2022(1):1-10. [doi: 10.1155/2022/7697036]
49. Mohammad Amini M, Jesus M, Fanaei Sheikholeslami D, Alves P, Hassanzadeh Benam A, Hariri F. Artificial intelligence ethics and challenges in healthcare applications: a comprehensive review in the context of the European GDPR mandate. Mach Learn Knowl Extr 2023 Aug 07;5(3):1023-1035. [doi: 10.3390/make5030053]
50. Namba T, Yamada Y. Risks of deep reinforcement learning applied to fall prevention assist by autonomous mobile robots in the hospital. Big Data Cogn Comput 2018 Jun 17;2(2):13. [doi: 10.3390/bdcc2020013]
51. 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 [FREE Full text] [doi: 10.1007/s00146-021-01230-z] [Medline: 34054228]
52. Abuzaid MM, Elshami W, Fadden SM. Integration of artificial intelligence into nursing practice. Health Technol (Berl) 2022;12(6):1109-1115 [FREE Full text] [doi: 10.1007/s12553-022-00697-0] [Medline: 36117522]
53. Qayyum MU, Sherani AMK, Khan M, Hussain HK. Revolutionizing healthcare: the transformative impact of artificial intelligence in medicine. Bull Inform 2024;1(2):71-83 [FREE Full text]

Abbreviations

AI: artificial intelligence
ANN: artificial neural network
C4W: CARES4WOUNDS
DFU: diabetic foot ulcer
EMR: electronic medical record
ML: machine learning
NLP: natural language processing
RPA: robotic process automation
SWAT: Specialized Workforce for Acute Transport

Edited by E Borycki, K Cato; submitted 17.06.24; peer-reviewed by K Seibert; comments to author 04.08.24; revised version received 07.08.24; accepted 09.09.24; published 19.02.25.

Please cite as:
Al Khatib I, Ndiaye M
Examining the Role of AI in Changing the Role of Nurses in Patient Care: Systematic Review
JMIR Nursing 2025;8:e63335
URL: https://nursing.jmir.org/2025/1/e63335
doi:10.2196/63335
PMID:39970436

Impact of Attached File Formats on the Performance of ChatGPT-4 on the Japanese National Nursing Examination: Evaluation Study

Kazuya Taira[1], RN, PHN, PhD; Takahiro Itaya[2], RN, MPH, DrPH; Shuntaro Yada[3,4], PhD; Kirara Hiyama[2], RN, MPH; Ayame Hanada[2], RN, PHN, BHS

[1]Human Health Sciences, Graduate School of Medicine, Kyoto University, 53, Shogoinkawara-cho, Sakyo-ku, Kyoto, Japan

[2]Department of Healthcare Epidemiology, Graduate School of Medicine and Public Health, Kyoto University, Kyoto, Japan

[3]Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, Japan

[4]Faculty of Library, Information and Media Science, University of Tsukuba, Tsukuba, Japan

Corresponding Author:
Kazuya Taira, RN, PHN, PhD
Human Health Sciences, Graduate School of Medicine, Kyoto University, 53, Shogoinkawara-cho, Sakyo-ku, Kyoto, Japan

Abstract

Abstract: This research letter discusses the impact of different file formats on ChatGPT-4's performance on the Japanese National Nursing Examination, highlighting the need for standardized reporting protocols to enhance the integration of artificial intelligence in nursing education and practice.

(JMIR Nursing 2025;8:e67197) doi:10.2196/67197

KEYWORDS

nursing examination; machine learning; ML; artificial intelligence; AI; large language models; ChatGPT; generative AI

Introduction

Numerous generative artificial intelligences (AIs), exemplified by all versions of ChatGPT [1] and Llama [2], have been developed using large language models and evaluated in health care, particularly in nursing education [3,4], successfully passing national nursing examinations in several countries [5,6]. Generative AIs are evolving to handle multimodal information, including text and images [1]. However, previous evaluations have not assessed the impact of file formats [5,6].

Prompts, particularly long ones, can affect response accuracy owing to potential context loss or exceeded token limits [7-9]. In this study, we hypothesized that the file format attached to prompts could affect the results of nursing research that uses generative AI and aimed to evaluate its impact on ChatGPT-4's performance on the Japanese National Nursing Examination. The findings of this study would be useful for improving the quality of reports on future nursing research that uses generative AI.

Methods

Ethics Approval

This study did not require ethical approval or informed consent, as the data analyzed were obtained from a published database from the Ministry of Health, Labour and Welfare.

Generative AI Model

We used the original, unmodified GPT-4 (gpt-4 - 1106-preview, accessed March 2024) without additional training, tuning, or data. ChatGPT, launched by OpenAI in 2022, with GPT-4 released in March 2023, is currently widely used.

Input Data

The dataset included all 50 basic knowledge questions from the 2023 Japanese National Nursing Examination, along with 190 general questions. The passing standard for these basic knowledge questions was approximately 80%. ChatGPT-3.5 has consistently failed to meet this standard [4], leading us to consider whether performance might vary based on file format. Questions were prepared in TEXT (.txt), DOCX (.docx), PDF (.pdf), and IMAGE (.jpg) formats and in a format that directly described all questions in the prompt (PROMPT-ONLY format). Although other formats, including CSV, JSON, XML, and Markdown, could be used to present questions and choices, we excluded them to maintain consistency and focus on more common formats.

Prompt Engineering

The prompts for each file format are summarized in Textbox 1.

Textbox 1. Prompts provided to ChatGPT-4. The files (mentioned at the end of the prompt for TXT, DOCX, PDF, and JPG formats) were made viewable via OpenAI's application programming interface (API) function: ASSISTANT (type = retrieval).

<Prompt for PROMPT-ONLY format>

You are an expert in the field of nursing. Answer the given questions briefly and numerically. {Question number}. {Question}. Options: (1) {Option 1}, (2) {Option 2}, (3) {Option 3}, (4) {Option 4}

Example: *1. Which vessel sends blood from the fetus to the placenta in the fetal circulation? Options: (1) Common carotid artery, (2) Pulmonary artery, (3) Umbilical artery, and (4) Umbilical vein.*

<Prompt for TXT, DOCX, PDF, and JPG formats>

You are an expert in the field of nursing. Answer briefly and numerically all questions given by the file.

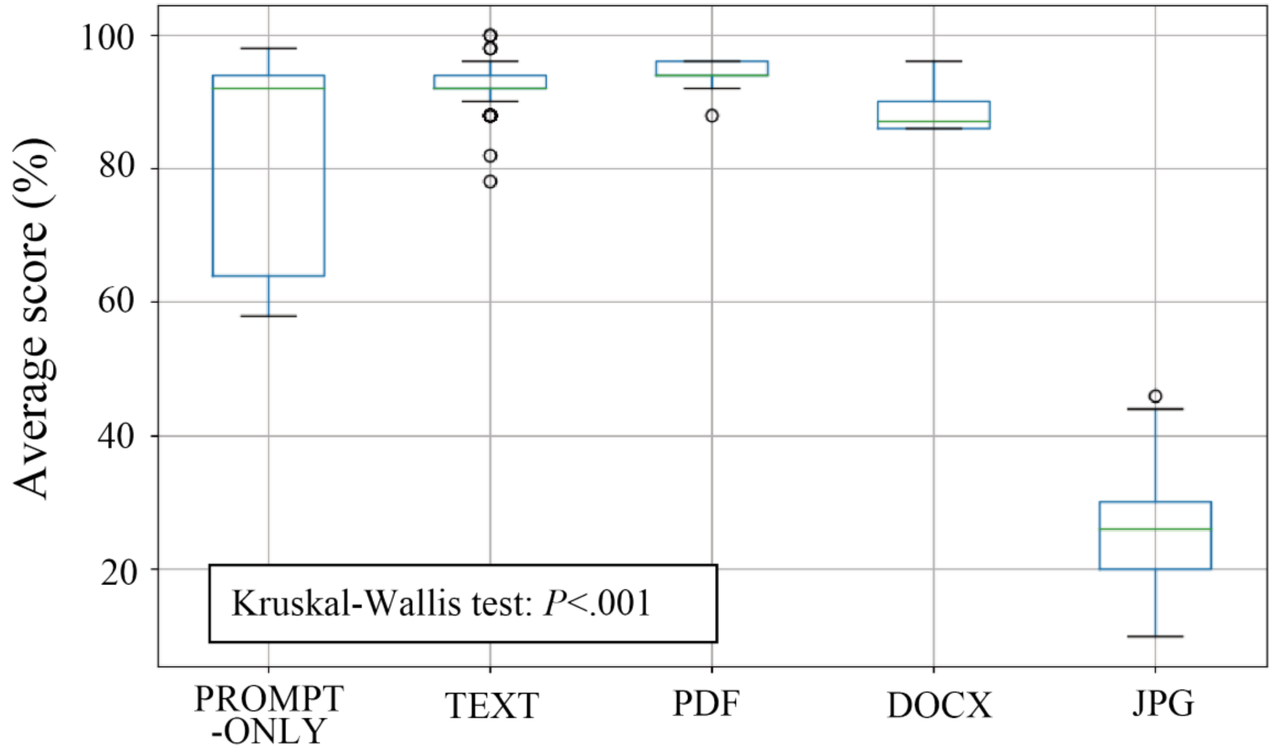
Data Analyses

Prompts for all formats were processed for 100 iterations each; the median and IQR of the percentage of correct answers were calculated. Differences among the percentages of correct answers by the attached file format were compared using the Kruskal-Wallis test and Dann-Bonferroni test. Statistical analyses were performed using Python (version 3.11.4) with the *pandas* (version 1.5.3) and *matplotlib* (version 3.7.1) libraries.

Results

The median percentages of correct answers were 92% (IQR 64% - 94%), 92% (IQR 92% - 94%), 94% (IQR 94% - 96%), 87% (IQR 86% - 90%), and 26% (IQR 20% - 30%) for PROMPT-ONLY, TEXT, PDF, DOCX, and JPG formats, respectively. The differences between the attached formats were statistically significant in all pairs ($P<.01$) except for the PROMPT-ONLY versus TEXT and PROMPT-ONLY versus DOCX pairs (Figure 1).

Figure 1. Performance evaluation of ChatGPT-4 on the Japanese National Nursing Examination by the attached file format. Outliers, shown as dots, are values below Q1 – 1.5 or above Q3 + 1.5 in the IQR.



		PROMPT-ONLY	TEXT	PDF	DOCX	JPG
Median (IQR)		92% (64%-94%)	92% (92%-94%)	94% (94%-96%)	87% (86%-90%)	26% (20%-30%)
Dann-Bonferroni test	PROMPT-ONLY		*P*>.99	*P*<.001	*P*=.58	*P*<.001
	TEXT			*P*=.003	*P*<.001	*P*<.001
	PDF				*P*<.001	*P*<.001
	DOCX					*P*<.001

Discussion

ChatGPT-4's performance on the Japanese National Nursing Examination varied significantly across file formats. The best performance was observed with PROMPT-ONLY, TEXT, and PDF formats (median scores >92%), followed by DOCX (87%), and the worst performance was with JPG (26%). The PROMPT-ONLY format exhibited a larger IQR and more variability than TEXT, PDF, and DOCX formats. JPG's poor performance highlights a significant limitation of generative AI, which excels at processing digital text but struggles with interpreting text from images. This "visual comprehension" gap has critical implications for AI applications involving nondigital text sources. The variability in PROMPT-ONLY performance may reflect reduced accuracy with longer prompts [7,8].

Therefore, to prepare for a future where generative AI is integrated into nursing practice and education [10], it is crucial to understand the interaction between humans and generative AI, including the impact of input file formats. Additionally, it is essential to report the following aspects in a standardized manner:

- Name and version of the generative AI model
- Presence of additional training, tuning, or knowledge transfer
- Prompt design and attached file formats
- Response generation parameters, including the number of iterations, temperature settings, and maximum token count
- Execution environment (if applicable)

However, as we only examined ChatGPT-4's performance on the Japanese National Nursing Examination and the impact of major file formats, investigations on other formats and AI models are warranted. Particularly, evaluating the performance of AI that specializes in image processing and image formats other than JPG and expanding the evaluations to include national nursing examinations in other countries and clinical questions in practice will be important in future research.

Acknowledgments

This study was supported by the Japan Society for the Promotion of Science (JSPS KAKENHI 22K17549). The funder played no role in the study design, data collection, analysis, interpretation, or writing of the report. We would like to thank Editage for the English-language editing. During the preparation of this work, the authors used DeepL and ChatGPT to improve the language and readability. The article was completely structured by author-oriented content; these artificial intelligence (AI) tools were only used to correct English expressions and check for grammar. Therefore, these AIs did not affect the results or interpretations. After using these tools, the authors reviewed and edited the content as necessary and take full responsibility for the content of the published article.

Conflicts of Interest

None declared.

References

1. OpenAI, Achiam J, Adler S, et al. GPT-4 technical report. arXiv. Preprint posted online on Mar 15, 2023. [doi: 10.48550/arXiv.2303.08774]
2. Topaz M, Peltonen LM, Michalowski M, et al. The ChatGPT effect: nursing education and generative artificial intelligence. J Nurs Educ 2024 Feb 5:1-4. [doi: 10.3928/01484834-20240126-01] [Medline: 38302101]
3. Touvron H, Lavril T, Izacard G, et al. LLaMA: open and efficient foundation language models. arXiv. Preprint posted online on Feb 27, 2023. [doi: 10.48550/arXiv.2302.1397]
4. Jin HK, Lee HE, Kim E. Performance of ChatGPT-3.5 and GPT-4 in national licensing examinations for medicine, pharmacy, dentistry, and nursing: a systematic review and meta-analysis. BMC Med Educ 2024 Sep 16;24(1):1013. [doi: 10.1186/s12909-024-05944-8] [Medline: 39285377]
5. 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]
6. Su MC, Lin LE, Lin LH, Chen YC. Assessing question characteristic influences on ChatGPT's performance and response-explanation consistency: insights from Taiwan's Nursing Licensing Exam. Int J Nurs Stud 2024 May;153:104717. [doi: 10.1016/j.ijnurstu.2024.104717] [Medline: 38401366]
7. Ratnayake H, Wang C. A prompting framework to enhance language model output. Presented at: AI 2023: Advances in Artificial Intelligence: 36th Australasian Joint Conference on Artificial Intelligence; Nov 28 to Dec 1, 2023; Brisbane, Australia. [doi: 10.1007/978-981-99-8391-9_6]
8. Levy M, Jacoby A, Goldberg Y. Same task, more tokens: the impact of input length on the reasoning performance of large language models. arXiv. Preprint posted online on Feb 19, 2024. [doi: 10.18653/v1/2024.acl-long.818]
9. Zhang ZY, Verma A, Doshi-Velez F, Low BKH. Understanding the relationship between prompts and response uncertainty in large language models. arXiv. Preprint posted online on Jul 20, 2024. [doi: 10.48550/arXiv.2407.14845]
10. Goldberg CB, Adams L, Blumenthal D, et al. To do no harm — and the most good — with AI in health care. NEJM AI 2024 Feb 22;1(3). [doi: 10.1056/AIp2400036]

Abbreviations

AI: artificial intelligence

Edited by E Borycki; submitted 15.10.24; peer-reviewed by G Sun, PH Liao, ST Arasteh; revised version received 23.12.24; accepted 26.12.24; published 22.01.25.

Please cite as:
Taira K, Itaya T, Yada S, Hiyama K, Hanada A
Impact of Attached File Formats on the Performance of ChatGPT-4 on the Japanese National Nursing Examination: Evaluation Study
JMIR Nursing 2025;8:e67197
URL: https://nursing.jmir.org/2025/1/e67197
doi:10.2196/67197

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org