Original Paper

A Data Strategy for the Secondary Use of Nursing Real-World Data: Feasibility Study of the NuDaK Framework

Renate Nantschev, RN, MSc; Elske Ammenwerth, PhD

Institute of Medical Informatics, UMIT TIROL – Private University for Health Sciences and Health Technology, Hall in Tirol, Tirol, Austria

Corresponding Author:

Renate Nantschev, RN, MSc
Institute of Medical Informatics, UMIT TIROL – Private University for Health Sciences and Health Technology
Eduard Wallnöfer Zentrum 1
Hall in Tirol, Tirol 6060
Austria
Phone: 43 681 10399763
Email: renate.nantschev@edu.umit-tirol.at

Abstract

Background: Routinely collected nursing real-world data (RWD) offer opportunities for nursing research to generate new knowledge and to improve quality of care. Furthermore, advanced analytics with nursing routine data opens the development of predictive models that support risk detection and decision support. However, using RWD in nursing remains challenging due to insufficient data standardization, variable data quality, and difficulty for organizations implementing a targeted data strategy for the secondary use of nursing data. To address this gap, we developed the NuDaK nursing data strategy framework as a practical guide to enable and optimize the secondary use of routinely collected nursing documentation data.

Objective: This study aimed to evaluate the completeness and applicability of the NuDaK framework for the secondary use of nursing routine data in a clinical nursing setting.

Methods: We conducted a formative feasibility study in collaboration with a large acute care hospital in Austria. The NuDaK framework was piloted on a surgical ward and evaluated across 4 implementation phases. Data were collected via phase-specific workshops and project meetings, continuous self-evaluation protocols by the project lead, and a final online focus group interview. Data were analyzed using deductive qualitative content analysis regarding the completeness and applicability of NuDaK. Evaluation results were aggregated and used to update NuDaK.

Results: The 4 core NuDaK aspects ("Purpose & Benefits," "Dataset & Nursing Documentation," "Data Modeling & Software Requirements," and "Data Integration & Data Quality") were confirmed as conceptually complete. Based on the evaluation results, NuDaK was updated by (1) adopting a circular structure with 4 temporal phases, (2) defining interdisciplinary stakeholder involvement across phases, (3) specifying professional roles and responsibilities across phases, and (4) integrating an iterative nursing information model spanning dataset development, specification, and refinement. Further NuDaK additions comprise a data quality concept with operationalized dimensions and validation procedures.

Conclusions: The NuDaK nursing strategy framework was refined through real-world piloting. It now provides organizations and researchers with a nursing-specific, end-to-end nursing data strategy to enable systematic RWD use. While its feasibility was demonstrated in a real-world nursing setting, further implementation studies should assess transferability across heterogeneous institutions with different documentation systems, governance structures, and levels of digital maturity.

JMIR Nursing 2026;9:e95546; doi: 10.2196/95546

Keywords: nursing informatics; nursing real-world data; secondary use of nursing data; data strategy framework; nursing outcome analytics

Introduction

Background

Digitization in nursing care has advanced considerably. In Germany, digital nursing and treatment documentation averaged 68.4% in 2023 [1]. Similar levels are reported in Austria and Switzerland, where 76% of nursing documentation in hospitals is electronic and 88% of facilities use IT-supported nursing documentation in at least half of their units [2]. Likewise, in the United States, over 96% of hospitals used an electronic health record (EHR) system in 2022 for clinical notes, most with integrated digital nursing documentation [3].

This digital transformation not only changes documentation processes but also involves the routine collection of vast amounts of real-world data (RWD).

RWD are defined as data routinely collected during care provision relating to patient health status or the delivery of health care from various sources, for example, EHRs and documentation systems, disease-specific medical research registries, and data from wearable or mobile devices [4].

RWD can support both nursing research and quality improvement. They offer several advantages because RWD are routinely collected from various health-related sources; they can provide comprehensive and timely representations of patient care and care delivery [5]. The use of RWD in nursing research can help generate evidence-based knowledge for practice, improve the quality of patient care, and further develop nursing practice by, for example, assessing the effectiveness of nursing interventions in achieving specific nursing outcomes based on routine data. In addition, RWD-based analyses can support personalized, patient-centered care while optimizing resource use [6,7]. Advanced analytics with nursing routine data also opens the development of predictive models that support applications such as risk detection and decision support [8-10].

Despite these potentials, the secondary use of nursing RWD remains limited. One major challenge is that non-standardized data entry limits data reuse. In many settings, standardized nursing terminologies are insufficient or inconsistently applied [2]. This lack of standardization makes it difficult to harmonize and integrate data for meaningful analysis [4,11]. Additionally, the widespread use of free-text documentation, while valued by clinicians for flexibility and completeness, makes data analysis difficult [2,12].

Another challenge is low data quality. Incomplete or incorrect documentation can lead to data gaps and analytical errors and reduce the reliability and validity of results [4].

Finally, many institutions lack a comprehensive data strategy that addresses all aspects relevant to nursing research and nursing outcome evaluation [2]. Without a coordinated approach, data are often collected in isolated application systems, using inconsistent formats and varying documentation standards and terminologies. This fragmentation limits data interoperability, complicates data linkage across settings, and restricts the ability to aggregate information for large-scale analyses [13].

To address these challenges, the literature describes several foundational components that should be addressed to enable the secondary use of nursing RWD. A key component is the standardization and structured documentation of nursing documentation. The consistent use of standardized nursing terminologies and structured data formats enhances the ability to characterize nursing care and analyze its effects on nursing outcomes. Structured and standardized documentation practices are essential for ensuring high data quality, interoperability, and meaningful integration with other clinical data sources [14,15]. Equally important is the precise definition of secondary use purposes and the identification of outcome indicators. Together, these elements provide the foundation for outcome-oriented and evidence-based nursing research [16,17]. Nursing-sensitive outcome indicators serve as key measures in this context. They capture outcomes directly influenced by nursing care, such as patient falls, pressure injuries, pain management, or infection rates, and provide a quantitative basis for evaluating care quality and effectiveness [18-26].

Another essential component for the secondary use of nursing data is the establishment of a defined dataset. A clearly specified dataset determines which data elements are relevant for describing nursing practice and nursing outcomes, ensuring that the collected information is complete, consistent, and comparable across institutions and care settings [27,28].

To ensure that these data are reliable and interpretable, data information models specify the structure and relationships among these datasets and define their meaning in a standardized form. These models bridge clinical and technical perspectives by translating nursing knowledge into formal data specifications, reducing ambiguity and enabling accurate data capture in electronic documentation systems [27,29].

A further critical component is data quality, which determines the reliability and validity of any secondary analysis. High-quality nursing data to be used for RWD analysis must be complete, correct, current, consistent, and plausible, which is supported by systematic data validation and transparent data reporting [30-33].

Finally, data integration is required to consolidate data from multiple sources into an analysis-ready structure (eg, via a data warehouse or data mart), which typically requires close interdisciplinary collaboration among IT professionals, data scientists, and nursing researchers to ensure clinical relevance and technical feasibility [16].

While these components represent core building blocks identified in the literature for enabling the secondary use of nursing RWD, existing frameworks for this task often address only selected aspects rather than providing an integrated, nursing-specific strategy.

The SPIRIT (Systematic Planning of Intelligent Reuse of Integrated Clinical Routine Data) framework offers a secondary use strategy for the entire organization,

comprehensive analyses to define system and user requirements, and the implementation of clinical data integration, management, and analysis solutions. However, it does not address specific nursing outcomes and nursing routine data [16].

The framework for mining and analyzing standardized nursing care plan data facilitates the extraction of valuable knowledge from nursing care plan data, demonstrating its applicability in improving nursing outcomes, supporting nursing staff, and enhancing administrative functions [9]. However, the handling of nonstandardized care plan data remains insufficiently addressed, particularly in terms of data harmonization and quality analysis, which are essential to ensuring the reliability and interoperability of the extracted information. The Health Data Quality Framework provides an automated approach for assessing and visualizing the quality of EHR data to support analytics such as predictive modeling and quality measures [33]. This framework primarily targets data quality and thus does not constitute an end-to-end strategy for the reuse of nursing routine data.

Finally, the Governance Framework developed for secondary use of routine health care data in nursing homes outlines governance, legal, and ethical principles for responsible reuse [34]. This framework is not intended to be a comprehensive nursing data strategy across the full life cycle.

Overall, most available frameworks focus on individual dimensions (eg, governance, data quality, or analytic reuse) without systematically combining all key components into a coherent, nursing-specific data strategy.

In summary, no existing framework comprehensively supports the secondary use of routinely collected nursing documentation data in a nursing-specific way. Current approaches do not adequately address specific nursing outcomes and the particular characteristics of nursing routine data. As a result, an important gap remains in how nursing RWD can be systematically leveraged for nursing-related purposes.

To address this gap, we developed the NuDaK nursing data strategy framework to enable and optimize the secondary use of routinely collected nursing documentation data.

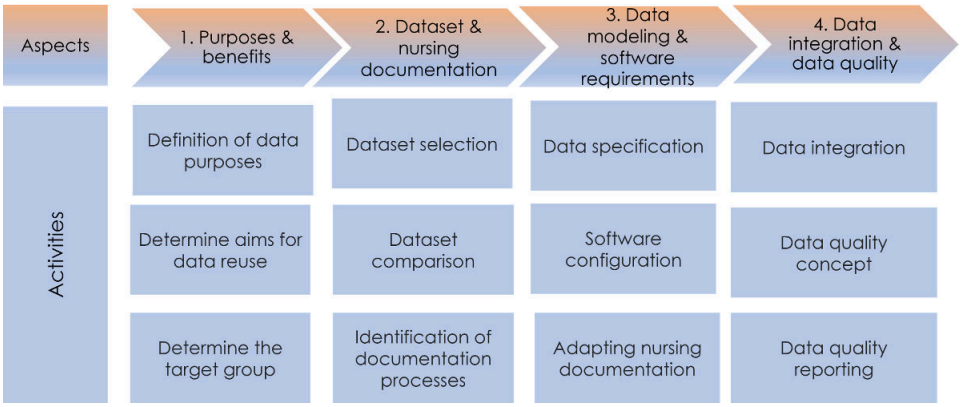
NuDaK does not aim to be universally exhaustive but focuses on addressing nursing data, nursing-related outcomes, and the specific characteristics of routine nursing data. The NuDaK framework serves as a practical guide for both organizations and research teams by structuring how clinical routine data can be prepared, governed, and analyzed for nursing purposes.

In organizational settings, within single institutions or across organizations, NuDaK supports assessing the effectiveness of nursing interventions on nursing outcomes, strengthening nursing quality management, enabling personalized patient-centered care while optimizing resource use, and informing workflow steering.

In research settings, NuDaK underpins robust evidence generation in nursing by improving data quality, standardization, and governance, thereby enabling advanced analytics such as data mining, AI-based predictive modeling, risk detection, and decision support.

The NuDaK framework encompasses 4 interrelated aspects that reflect the key components of a comprehensive data strategy (see Figure 1 for an overview) [35,36].

Figure 1. NuDaK nursing data strategy framework (2022) for the secondary use of routine nursing data.



1. Defining the purposes and benefits of using routinely collected data
2. Establishing a standardized nursing dataset and adapting nursing documentation to nursing workflows
3. Data modeling and specification of software requirements
4. Implementing processes for data integration and quality assurance [35,36]

Some information on NuDaK has been published in [35,36]. While NuDaK was built on the results of a narrative-integrative literature review with conceptual synthesis, a comprehensive description and evaluation have not yet been conducted. This is therefore the objective of this paper.

Objective

This study aimed to evaluate the completeness and applicability of the NuDaK data strategy framework in a real clinical nursing setting and, based on these findings, to refine NuDaK as a practical guide for organizations and researchers

to enable the secondary use of routinely collected nursing documentation data.

Methods

Study Design

This study is reported in accordance with the STARE-HI (Statement on Reporting of Evaluation Studies in Health Informatics) guideline to enhance the quality and transparency of reporting for quantitative and qualitative health informatics evaluation studies. In line with STARE-HI, the *Methods* section is structured to provide sufficient detail on the study design and methods used in this study [37].

We conducted a 1-year formative feasibility study (February 2023 to June 2024) to pilot and evaluate the NuDaK framework in a real clinical nursing setting. The evaluation focused on 2 primary dimensions: (1) applicability (how well it fits the health care environment) and (2) completeness (whether it covers all necessary aspects and activities comprehensively).

The study used 3 complementary evaluation methods: a formative qualitative external evaluation, a qualitative self-evaluation, and an online focus group interview at the end of the piloting phase.

Theoretical Background

Formative evaluations are particularly employed when a program, product, concept, or intervention is still in development or testing. They primarily serve an optimization function: the goal is to continuously improve the evaluation object during its development or application. Continuous feedback of evaluation results into practice is beneficial, enabling improvements to be implemented and tested through multiple feedback loops [38].

Qualitative methods play an important role in the context of formative evaluations. They enable differentiated capture of experiences, assessments, pros and cons, and concrete change requests from evaluation participants and users [38]. For evaluating the NuDaK framework, this approach is appropriate because the framework is still in the testing phase and requires continuous improvements through iterative feedback loops.

Setting

The study was conducted in collaboration with a large acute care hospital in Austria and the software company that provides the electronic nursing documentation. Together with the stakeholders, we selected a surgical ward to pilot the NuDaK framework.

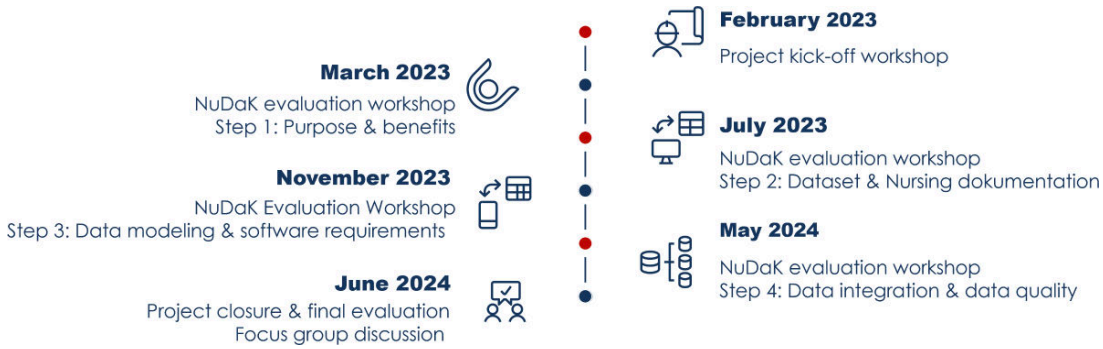
The stakeholders involved in the project were systematically selected through a stakeholder analysis.

In total, 15 stakeholders were identified. Ten active stakeholders participated directly in project activities (director of nursing, experts from the unit of nursing development, ward manager, a practical nurse at the piloting station, IT expert at the hospital, staff member of the software company, and expert in data security and quality management). Five passive stakeholders (executive management, IT business management, and IT requirements and implementation management) supported strategic decision-making and were regularly informed about progress but did not participate in the planned project activities.

Study Flow

The pilot implementation followed a detailed project plan aligned with the 4 aspects of the NuDaK framework. Overall project management was led by the principal researcher, who also served as the project leader. On the clinical side, a designated nurse representative served as the primary point of contact and coordinated all on-site activities. In addition, the software vendor assigned a dedicated project lead to support implementation and technical coordination. The study was conducted from February 2023 to June 2024 and comprised 6 phases: a kick-off phase (February 2023) for initial training and framework introduction; 4 piloting steps corresponding to NuDaK implementation steps 1 to 4, with final evaluation workshops; and project closure (June 2024) for final evaluation. Figure 2 outlines the evaluation timeline within the piloting steps.

Figure 2. NuDaK framework evaluation activities within the piloting steps.



Outcome Measure

This study evaluated the NuDaK framework using 2 primary outcome measures: completeness and applicability.

Completeness addresses whether the framework covers all necessary aspects and activities comprehensively, ensuring there are no gaps in the intended scope. Applicability examines how well the framework fits the real health care

environment, including its clarity, usability, and suitability for end users in practice.

The evaluation addressed these predefined criteria using the following guiding question, as presented in Table 1.

Table 1. Evaluation questions on completeness and applicability of the NuDaK framework.

Main evaluation question	Subquestions
Completeness evaluation	
Q1. Is the set of aspects covered in the NuDaK framework complete and sufficient?	• Are additional aspects or themes required to improve comprehensiveness? • Do the activities follow a coherent and meaningful sequence?
Q2. Are the activities assigned to each NuDaK aspect logically connected?	
Q3. Is each activity within the NuDaK framework feasible to perform in practice?	• Are temporal dependencies between activities clearly defined and logically structured? • Does the framework link resource needs to specific aspects, activities, or methods? • Is the time required for each step realistic and proportionate to available resources?
Q4. Is each NuDaK framework activity feasible in practice, given the required resources and time?	
Applicability evaluation	
Q5. To what extent is the NuDaK framework applicable in practice?	• Are the described aspects and activities understandable and logically structured? • Can users easily interpret the purpose of each activity? • Are the expected outcomes of each activity clear? • Are ambiguities or overlaps present between framework components? • Do the recommended methods adequately support data reuse processes in the nursing context? • Do responsibilities and competencies correspond to the requirements of each activity?
Q6. How clear and comprehensible are the NuDaK framework's aspects and activities for end users?	
Q7. Are the methods and tools recommended within the NuDaK framework suitable for achieving the intended objectives?	
Q8. Are the designated expert roles appropriately aligned with the respective NuDaK framework activities?	

Data Collection

Overview

We combined 3 complementary data collection methods (workshops, internal self-evaluation, and focus group interviews) to address all evaluation questions regarding completeness and applicability of the NuDaK framework. This multimethod approach enabled triangulation of findings and captured both external perspectives (workshops and focus group) and internal methodological reflections (self-evaluation).

Workshops (External Formative Evaluation)

Four workshops were conducted with actively involved project participants to evaluate the applicability of the methods applied in each step and to identify practical barriers, ambiguities, or missing components in the NuDaK framework.

Each workshop included 6 to 10 participants, comprising the director of nursing, ward manager, nursing staff of the piloting station, nursing representatives, IT experts, a staff member of the software company, and an expert in data security and quality management.

Participants were selected based on their active involvement in the piloting steps and their direct experience with NuDaK framework activities.

The workshop questions were adapted to align with the overarching evaluation questions from Table 1, with context-specific modifications for each piloting step.

For example:

- Is the BPMN (Business Process Model and Notation) modeling method suitable for visualizing real documentation processes on the ward?
- Are nursing staff able to use this visualization to assess processes on the ward for completeness and accuracy?
- Are the activities assigned to each NuDaK aspect logically connected and do they follow a coherent sequence?
- Is each activity feasible to perform in practice, given the available resources and time?
- Are the described aspects and activities understandable to end users?

All workshops were audio-recorded and supplemented with written workshop documentation (participant notes, process diagrams, and feedback sheets). Audio files were stored securely and transcribed for qualitative analysis.

Internal Self-Evaluation

Throughout the project, the researcher conducted a structured self-evaluation to assess whether the methods used achieved the goals of each specific aspect and activity of the NuDaK framework. This involved a continuous review process using the project implementation guidelines as a reference framework.

The project guidelines helped identify whether the planned NuDaK activities could be implemented effectively and whether they were applicable and suitable for achieving the goals of each specific aspect and activity. They also helped determine whether any activities were missing, whether they were applied in the correct sequence, and whether the overall

timeline was appropriate. Furthermore, the self-evaluation considered whether the active stakeholders involved in the project had sufficient competencies for their assigned tasks. The researcher maintained continuous self-evaluation notes throughout the study. These notes were stored and served as supplementary data for qualitative analysis.

Focus Group Interview (External Formative Evaluation)

The final online focus group interview (n=5) followed a semistructured guideline and was developed to (1) explore the conceptual understanding and perspectives on the secondary use of routine nursing data, (2) assess the applicability, comprehensibility, and practical relevance of the NuDaK framework, (3) explore challenges, barriers, and facilitating factors in implementing a data strategy for the secondary use of routine data, (4) critically reflect on the composition, roles, and contributions of the stakeholder group, and (5) explore the perceived institutional and organizational added value resulting from the project.

The focus group included 5 project participants (the director of nursing, an IT expert, a member of the software company, and 2 designated nursing representatives). The participants were selected based on their comprehensive involvement across all 4 piloting steps and their ability to provide synthesized perspectives on the framework's overall implementation.

The focus group interview was audio-recorded via the online platform and automatically transcribed.

Data Analysis

Data obtained from workshops and focus group interviews were audio-recorded and transcribed in summary form after each session.

We applied a deductive qualitative content analysis approach based on the 8 predefined evaluation questions (Q1-Q8) from Table 1. The coding process included initial coding based on the evaluation questions, and the analysis was supported by MAXQDA Analytics Pro 2022 (VERBI Software GmbH).

Findings from the internal self-evaluation were triangulated with external qualitative data from the workshops and focus group discussions to identify convergent and divergent assessments, identify refinement needs, and generate recommendations to optimize the NuDaK framework.

Ethical Considerations

This study was conducted in accordance with the standards laid out by the Research Committee for Scientific Ethical Questions. Research Committee for Scientific Ethical Questions is an independent body of the Tyrolean Private University UMIT that examines planned research projects that, due to legal provisions, do not fall under an ethics committee's jurisdiction.

Results

Overview

This section presents the results of the evaluation of the NuDaK framework (2022). Six deductive categories were developed to assess both its applicability and completeness. The categories include definitions and the corresponding evaluation questions, as described in the *Methods* section (see Table 2).

Table 2. Evaluation categories for completeness and applicability of the NuDaK framework.

Completeness	Applicability
Completeness of framework components (Q1)	Traceability of activities and comprehensibility for users (Q5 and Q6)
Structure, distinctness, and temporal logic of the framework components (Q2)	Suitability of methods and tools (Q7)
Feasibility of NuDaK framework activities in practice (Q3 and Q4)	Roles and responsibilities (Q8)

Completeness of the Framework Components (Q1)

The pilot evaluation indicates that the NuDaK framework's core aspects "Purpose & Benefits," "Dataset & Nursing Documentation," "Data Modelling & Software Requirements," and "Data Integration & Data Quality," cover the essential domains required to plan and implement the secondary use of nursing data. Accordingly, the aspect set was generally perceived as sufficient at the conceptual level.

However, the pilot also showed that several framework-level components require more explicit specification to make the framework complete for practical use, such as (1) explicitly embedding a temporal dimension (phases 1-4), (2) defining interdisciplinary stakeholder involvement throughout the phases, and (3) specifying professional roles with role profiles and responsibilities.

In addition, the pilot revealed targeted content gaps within individual aspects. Within "Data Modelling & Software Requirements," an explicit documentation and adaptation concept emerged as necessary to guide systematic changes in the documentation system. Within "Data Integration & Data Quality," additional activities were required, including a strategy workshop to clarify key contextual conditions (eg, technical infrastructure, legal constraints, and staffing) and the development of a data quality concept.

Finally, as a cross-cutting addition, the pilot highlighted the importance of a standalone, iterative information model that is explicitly integrated into the framework. Such a model supports both top-down and bottom-up refinement

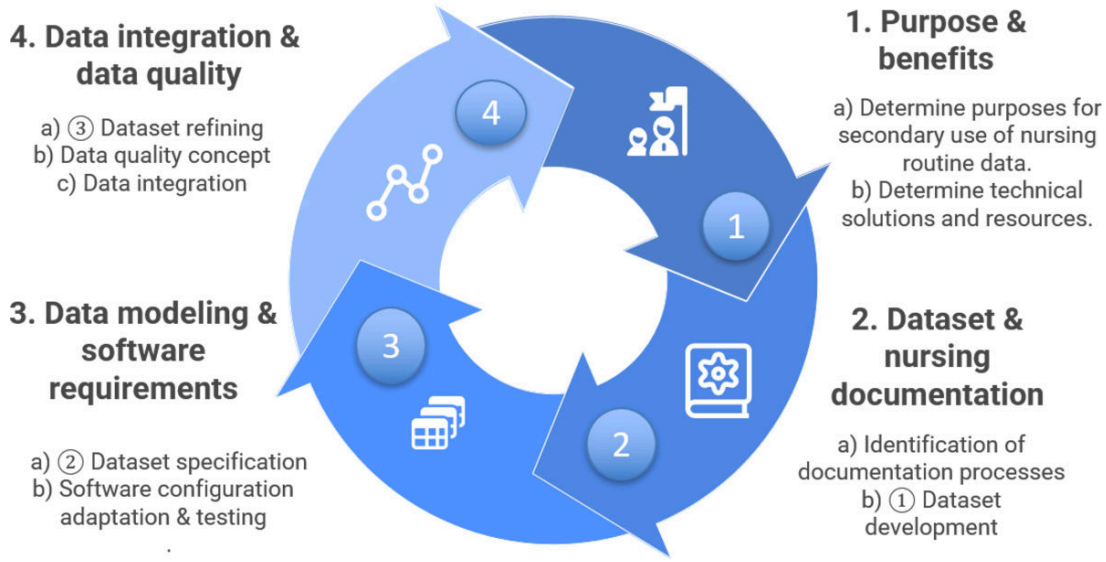
across dataset development, specification, and revision (see Multimedia Appendix 1). Early and continuous involvement of an interdisciplinary group was considered essential for defining shared objectives and balancing diverging interests. A combination of direct nursing practice, nursing management, nursing informatics/nursing development, and local IT was considered effective.

Structure, Distinctness, and Temporal Logic of the Framework Components (Q2)

The evaluation showed that the original linear and sequential representation of activities did not align well with practical requirements. Instead, it revealed a circular logic with recurring activities across multiple phases. Consequently, activities should not only be arranged in a single sequence but must also be distinguished by their temporal nature as either chronological key activities or iterative, cross-aspect activities.

In response, the revised NuDaK framework adopts a circular design comprising 4 phases, each assigned to an overarching aspect (see Figure 3). This structure makes iterative loops explicit and distinguishes between phase-specific key activities and recurring, cross-aspect activities.

Figure 3. NuDaK nursing data strategy framework 2025, representing aspects and activities.



Moreover, the pilot revealed that some activities must be scheduled more precisely to avoid late rework and loss of focus. A strategy workshop to clarify technical infrastructure, legal constraints, and human resources should be anchored early and assigned to "Purpose & Benefits" (phase 1). Likewise, a strategy for data querying and downstream processing should ideally be available by the end of phase 2 or, at the latest, at the beginning of phase 3. Finally, participants indicated that, in phase 4, the purpose and objectives of secondary use often require renewed clarification to maintain the intended focus of the data analysis.

Feasibility in Practice in Terms of Resources and Time (Q3 and Q4)

Overall, the evaluation suggests that the NuDaK framework activities are feasible to perform in practice. Nevertheless, feasibility is conditional on (1) an adequate time budget, (2) clearly assigned responsibilities, (3) access to relevant expertise, and (4) sufficient technical infrastructure.

At the activity level, participants described workshops, process modeling, and iterative concept development (eg, defining data purposes, modeling documentation processes, and specifying data quality criteria) as realistic and implementable when key stakeholders can be convened. In contrast, activities linked to implementation in the documentation system were perceived as particularly resource-intensive, especially testing adaptations of new applications, coordinating change management, and training staff after adaptations (including practical materials such as handouts and explicit clarification of "who does what").

Participants repeatedly reported that more time would have been beneficial, indicating that the planned time requirements of individual steps were perceived as tight in the pilot setting. In addition, the pilot underscored that sustainable continuation of a data strategy requires building organizational know-how over time, for example, through collaborations and structured training pathways (eg, university partnerships).

Taken together, the findings support making resource requirements and time demands more explicit within the framework by indicating resource-critical activities (eg, documentation system adaptations, data harmonization, and training) and by providing more realistic time estimates or planning guidance for key steps.

Traceability of Activities and Comprehensibility for Users (Q5 and Q6)

The evaluation showed that the NuDaK framework is generally applicable in practice. Participants reported that

they could independently develop a data strategy based on the framework, provided that the data purpose and supporting roles are defined early within an interdisciplinary setting.

Comprehensibility and traceability increased throughout the pilot project. However, at the outset, the framework was perceived as insufficiently concrete. Participants specifically noted that additional start-up input would have been helpful, including concrete examples and use cases.

A conceptual hurdle regarding the meaning of secondary use and the logic of routine data became apparent. Participants described the question "How can we benefit from the data and analyses?" as difficult to answer, and clarification of what secondary use entails emerged only during discussion. In addition, nursing managers and nursing developers often framed the topic in documentation-oriented terms and immediately sought documentation changes to enable analyses, which can conflict with the logic of routine data. These observations indicate that early alignment and clarification, supported by examples, are required from the outset.

Suitability of Methods and Tools (Q7)

Overall, the results suggest that the methods and tools recommended in the NuDaK framework are suitable for achieving the intended objectives and enabling the secondary use of nursing data. Workshops were described as central platforms for strategic decision-making, and audio recording was highlighted as a helpful documentation instrument to secure outcomes.

At the same time, the pilot indicated that methods often need to be adapted in application and, in some cases, applied repeatedly to achieve the desired results. Within "Dataset & Nursing Documentation," developing scenarios for documentation processes, BPMN modeling, and evaluation with nursing practitioners were perceived as effective and goal-oriented. BPMN modeling was considered particularly useful because it is understandable to nursing staff and provides a transparent basis for identifying software-related adaptation needs and relevant data sources.

For "Data Integration & Data Quality," the 2-step approach to developing the data quality concept was considered appropriate: selecting suitable quality dimensions from the literature (eg, completeness, concordance, plausibility, and timeliness) and subsequently operationalizing and validating them in a workshop with IT and nursing experts using practical examples and testing criteria. Moreover, data preparation requires continuous feedback from nursing staff, and harmonizing concepts (data objects) into a uniform data format is indispensable for larger datasets. In summary, the recommended methods most effectively support the intended objectives when applied iteratively and complemented by concrete examples/use cases and appropriate documentation instruments (eg, audio recordings).

Roles and Responsibilities (Q8)

The pilot implementation clearly demonstrated that professional roles, competencies, and responsibilities must be explicitly embedded within the framework, as they are essential for carrying out the activities.

Nursing informatics expertise emerged as central. Implementing a data strategy requires strong competencies in nursing informatics, and nursing informatics specialists should assume the lead and coordinating role. In addition, competencies in data science/data mining and data warehousing, as well as knowledge of programming languages and databases, must be combined with nursing/medical domain knowledge to interpret data correctly and understand nursing processes.

The designation of an independent data administrator role was identified as advantageous to ensure that data are available in the required format and granularity, particularly in phase 3 of implementation. Furthermore, topic-specific nursing expertise in the early phases (eg, a "pain nurse") was considered essential to align activities with nursing requirements.

To support practical implementation, responsibilities should be conceptualized not only as individual roles but also as clearly defined interdisciplinary expert groups for each phase to enable joint strategic alignment, holistic requirements consideration, and the balancing of diverging interests.

NuDaK Framework 2025: A Data Strategy to Reuse Nursing RWD

The evaluation of the applicability and completeness of the NuDaK Framework 2022 provided essential insights for its further development and refinement. Building on these findings, the framework was systematically adapted. In the following section, we present a short version of the updated NuDaK Framework 2025. A comprehensive description of the NuDaK Framework 2025, including detailed activities, methods, and role descriptions, is provided in Multimedia Appendix 1.

Major changes to the updated NuDaK framework comprise a clear temporal structure (phases 1-4). Furthermore, it is now conceptualized as a combination of chronological key activities and iterative, cross-aspect activities. In addition, the framework was strengthened by defining interdisciplinary stakeholder involvement, specifying professional roles, responsibilities, and competencies, and highlighting key prerequisites for feasibility.

Alongside this, a standalone, iterative information model was integrated to continuously improve the dataset (see Multimedia Appendix 1).

The NuDaK framework 2025 is a practice-oriented data strategy that structures the secondary use of routinely collected nursing data into 4 temporal phases and 4 corresponding aspects: (1) Purpose & Benefits, (2) Dataset & Nursing Documentation, (3) Data Modeling & Software Requirements, and (4) Data Integration & Data Quality (Figure 3). In addition to these aspects, the framework specifies activities and associated methods. Across all phases and aspects, implementation is enabled by defined professional roles and responsibilities (eg, nursing

informatics specialist, IT professionals, data administrator, ward management, and domain experts) that support the implementation of NuDaK.

In *Phase 1 (Purpose & Benefits*), organizations define the aims, expected benefits, and scope of secondary data use. Objectives are established through interprofessional consensus workshops and aligned with feasibility considerations (eg, data processing location, technical infrastructure, data science competencies, and regulatory constraints). Key deliverables are a purpose statement and an agreed analytical scope.

Phase 2 (Dataset & Nursing Documentation) focuses on understanding documentation practices and deriving a robust dataset. Documentation workflows are captured using scenario-based analysis and process modeling (eg, BPMN), followed by expert review. In parallel, the dataset is developed as the first stage ① of the *information model* using a top-down approach (a literature-informed theoretical dataset), refined against real documentation through use cases, and validated by domain experts.

In *Phase 3 (Data Modeling & Software Requirements),* the dataset is formally specified (information model stage ② [dataset specification]) by standardizing elements, defining attributes and permissible values, and adding missing fields as required. These specifications are translated into software requirements, followed by system configuration, user-centered adaptation, training support, and a structured testing period to ensure applicability and completeness of data capture.

Phase 4 (Data Integration & Data Quality) operationalizes information model stage ③ (dataset refinement) and establishes prerequisites for valid analyses. Iterative test queries are used to verify availability, formats, value ranges, and data types across sources. A data quality concept is developed by selecting core quality dimensions (eg, completeness, concordance, plausibility, and currency) and defining practical evaluation examples and validation procedures with clinical and technical experts. Finally, data integration consolidates validated data into a unified, analysis-ready structure while ensuring traceability, governance, and alignment with refined research questions. Further details are provided in Multimedia Appendix 1.

Discussion

Principal Results

This feasibility study evaluated the completeness and applicability of the NuDaK framework (2022) in a real-world nursing setting and used the findings to refine the framework into NuDaK 2025. Overall, the core set of aspects was confirmed as sufficient to cover key domains required for the secondary use of nursing routine data. However, NuDaK implementation showed that a linear representation does not reflect the iterative character of data strategy work in clinical environments.

Consequently, NuDaK 2025 now adopts a circular structure with 4 temporal phases and makes feedback loops explicit. In addition, the pilot indicated the need to (1) embed an explicit temporal dimension (phases 1-4), (2) formalize professional roles and responsibilities, and (3) introduce a standalone, iterative information model spanning dataset development, specification, and refinement across phases 2 to 4. Taken together, these refinements indicate that feasibility in this context depends not only on the content of the framework, but also on its ability to reflect the actual workflow logic of clinical implementation.

Results in Relation to Other Frameworks

NuDaK 2025 offers a nursing-specific, end-to-end strategy that connects purpose definition, dataset and documentation work, technical specification and software adaptation, and, finally, data integration and quality assurance. This distinguishes it from frameworks that address only selected parts of the reuse pathway.

The SPIRIT framework offers valuable guidance for organizational planning and system-level preparation, but it is not tailored to nursing purposes and does not operationalize the pathway from documentation to a reusable nursing dataset in the same level of detail [16].

Likewise, the Health Data Quality Framework is for assessing and visualizing EHR quality but does not provide an integrated life cycle strategy for the secondary use of nursing data that includes documentation processes and software implementation [33]. Nursing data mining approaches, including predictive modeling, focus more strongly on analytic reuse and less comprehensively on the upstream prerequisites, such as governance, integration, and quality management, that are needed before advanced analysis can take place [34].

Against this background, NuDaK extends existing approaches by combining strategic, organizational, technical, and methodological elements within one nursing-specific framework.

It covers the entire process from nursing documentation to a usable dataset for nursing analytics in a practical, step-by-step manner.

Methodological Implications

This exploratory study also shows methodological implications for the development and evaluation of complex frameworks in nursing informatics. First, it shows that formative feasibility work can reveal structural requirements that may not be fully visible at the conceptual design stage, especially when the framework is tested in real clinical practice. Second, the findings suggest that framework completeness should be interpreted not only as coverage of content domains but also as clarity of sequence, timing, roles, and feedback mechanisms. Third, the study illustrates the value of combining external qualitative evaluation with internal self-evaluation because these perspectives complement one another and help identify both practical barriers and areas for refinement.

Meaning and Generalizability

NuDaK 2025 provides a structured pathway for health care organizations and informatics specialists to establish the secondary use of routine nursing data. In particular, it can be used by nursing managers, hospital directors, or health informatics experts to plan, build, and implement the use of routine nursing data to improve nursing practice and support nursing research.

Its practical relevance lies in addressing common barriers that limit the reuse of nursing data. NuDaK particularly addresses insufficient standardization, variable data quality, and missing governance across the documentation-to-dataset pathway.

By aligning technical specification, software adaptation, and data quality assurance with nursing workflows, NuDaK can support robust analyses of large nursing datasets and thereby strengthen nursing quality management and evidence generation.

In the longer term, it may also contribute to establishing the data prerequisites for advanced analytics and AI-related applications. However, the implementation of such downstream applications requires additional technical, organizational, and methodological work beyond the scope of the present feasibility study. Given that AI applications in nursing rely on standardized, complete, and high-quality datasets, NuDaK may provide a relevant methodological foundation [10].

Overall, NuDaK 2025 should be understood as a strategy framework that structures the secondary use of routinely collected nursing data for quality improvement, research, and future analytic development.

Strengths and Limitations

A key strength of this feasibility study is that NuDaK was evaluated in a real clinical nursing setting and across multiple implementation phases, which allowed practical barriers and refinement needs to emerge from actual use rather than from theoretical reflection alone. The combination of workshops, focus group data, and internal self-evaluation also provided complementary perspectives on the same framework.

This study has several limitations. It was conducted in a single hospital ward, so transferability to other organizations with different workflows, documentation systems, or nursing terminologies is limited. In addition, the findings are based on qualitative perceptions and researcher self-evaluation rather than objective outcome measures, so they should be interpreted as exploratory. Future studies should therefore examine NuDaK in additional settings and complement qualitative findings with more objective implementation and outcome data.

Conclusions

Building on the need to better leverage routinely collected nursing documentation data, we developed the NuDaK framework as a practical guide for health organizations and research teams. Refined through a feasibility study in a real-world nursing setting, NuDaK now offers a nursing-specific, end-to-end data strategy across 4 phases, with explicit feedback loops, defined professional roles, and an iterative information model. The present feasibility study supports the conceptual refinement of the framework.

However, several issues remain only partially resolved. Further research is needed to evaluate the performance and transferability of the framework across heterogeneous institutions with different documentation systems, governance arrangements, vendors, and levels of digital maturity. Multisite studies are also needed to examine data harmonization and interoperability requirements across settings and software solutions. Future studies should also validate NuDaK through concrete nursing use cases and exemplar analyses while assessing its longitudinal sustainability and impact in routine practice.

Acknowledgments

The authors gratefully acknowledge all experts and practice partners for their valuable feedback and contributions to piloting and evaluating the NuDaK framework. We also thank the participating health care organization and the software company for their support in the feasibility study and technical implementation.

Generative AI Use

During the preparation and revision of this manuscript, the authors used ChatGPT (GPT-6.1 Sol; OpenAI) to support language editing, improve clarity and readability, and refine the structure and wording of selected passages. Napkin AI was used to support the graphical visualization of concepts. Perplexity AI and Elicit were used to assist with literature searches and the identification of relevant scientific publications. All AI-assisted outputs were critically reviewed and verified by the authors, who take full responsibility for the accuracy, originality, and integrity of the manuscript, including all references and citations.

Funding

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

Authors' Contributions

Conceptualization: RN (lead)
Data curation: RN
Formal analysis: RN (lead), EA (supporting)
Investigation: RN (lead), EA (supporting)

Methodology: RN (lead), EA (supporting)
Project administration: RN
Visualization: RN
Writing – original draft: RN
Writing – review & editing: RN (lead), EA (supporting).

Conflicts of Interest

None declared.

Multimedia Appendix 1

NuDaK nursing data strategy framework 2025.
[DOCX File (Microsoft Word File), 299 KB-Multimedia Appendix 1]

References

1. Hübner U, Liebe JD, Bicker L, et al. IT-Report Gesundheitswesen: Wie veränderte sich die Gesundheits-IT aus User-Perspektive? [Report in German]. Hochschule Osnabrück; 2024. URL: https://www.hs-osnabrueck.de/fileadmin/HSOS/Homepages/IT-Report_Gesundheitswesen/IT_Report_2024.pdf [Accessed 2026-11-09]
2. Nantschev R, Ammenwerth E. Challenges using electronic nursing routine data for outcome analyses: a mixed methods study. Int J Nurs Sci. Jan 2022;9(1):92-99. [doi: 10.1016/j.ijnss.2021.11.003] [Medline: 35079610]
3. National trends in hospital and physician adoption of electronic health records. Health IT Research & Analysis. 2026. URL: https://healthit.gov/data/quickstats/national-trends-hospital-and-physician-adoption-electronic-health-records/ [Accessed 2026-02-22]
4. Colman BD, Zhu Z, Qi Z, van der Walt A. From real world data to real world evidence to improve outcomes in neuro-ophthalmology. Eye (Lond). Aug 2024;38(12):2448-2456. [doi: 10.1038/s41433-024-03160-8] [Medline: 38844583]
5. Nahm ES, Zhu S, Seidl K, Chen L, Day J, Seong H. Real-world data for interdisciplinary health care research: a case example. ANS Adv Nurs Sci. 2023;46(4):349-362. [doi: 10.1097/ANS.0000000000000496] [Medline: 37102714]
6. McGrow K. Artificial intelligence in nursing: a journey from data to wisdom. Nursing. Apr 1, 2025;55(4):16-24. [doi: 10.1097/NSG.0000000000000165] [Medline: 40122866]
7. Westra BL, Christie B, Johnson SG, et al. Modeling flowsheet data to support secondary use. Comput Inform Nurs. Sep 2017;35(9):452-458. [doi: 10.1097/CIN.0000000000000350] [Medline: 28346243]
8. Raju D, Su X, Patrician PA, Loan LA, McCarthy MS. Exploring factors associated with pressure ulcers: a data mining approach. Int J Nurs Stud. Jan 2015;52(1):102-111. [doi: 10.1016/j.ijnurstu.2014.08.002] [Medline: 25192963]
9. Khokhar A, Lodhi MK, Yao Y, Ansari R, Keenan G, Wilkie DJ. Framework for mining and analysis of standardized nursing care plan data. West J Nurs Res. Jan 2017;39(1):20-41. [doi: 10.1177/0193945916672828] [Medline: 27756852]
10. Wei Q, Pan S, Liu X, Hong M, Nong C, Zhang W. The integration of AI in nursing: addressing current applications, challenges, and future directions. Front Med (Lausanne). 2025;12:1545420. [doi: 10.3389/fmed.2025.1545420] [Medline: 40007584]
11. Safran C. Update on data reuse in health care. Yearb Med Inform. Aug 2017;26(1):24-27. [doi: 10.15265/IY-2017-013] [Medline: 29063535]
12. Vuokko R, Mäkelä-Bengs P, Hyppönen H, Lindqvist M, Doupi P. Impacts of structuring the electronic health record: results of a systematic literature review from the perspective of secondary use of patient data. Int J Med Inform. Jan 2017;97:293-303. [doi: 10.1016/j.ijmedinf.2016.10.004] [Medline: 27919387]
13. Topaz M, Pruinelli L. Big data and nursing: implications for the future. Stud Health Technol Inform. 2017;232:165-171. [Medline: 28106594]
14. Macieira TGR, Chianca TCM, Smith MB, et al. Secondary use of standardized nursing care data for advancing nursing science and practice: a systematic review. J Am Med Inform Assoc. Nov 1, 2019;26(11):1401-1411. [doi: 10.1093/jamia/ocz086] [Medline: 31188439]
15. Liljamo P, Kinnunen UM, Saranto K. Assessing the relation of the coded nursing care and nursing intensity data: towards the exploitation of clinical data for administrative use and the design of nursing workload. Health Informatics J. Mar 2020;26(1):114-128. [doi: 10.1177/1460458218813613] [Medline: 30516092]
16. Hackl WO, Ammenwerth E. SPIRIT: Systematic Planning of Intelligent Reuse of Integrated Clinical Routine Data: a conceptual best-practice framework and procedure model. Methods Inf Med. 2016;55(2):114-124. [doi: 10.3414/ME15-01-0045] [Medline: 26769124]
17. Leiner F, Gaus W, Haux R, Knaup-Gregori P. Medical Data Management: A Practical Guide. Springer; 2003. ISBN: 9780387951591
18. Afaneh T, Abu-Moghli F, Ahmad M. Nursing-sensitive indicators: a concept analysis. Nurs Manag (Harrow). Jun 1, 2021;28(3):28-33. [doi: 10.7748/nm.2021.e1982] [Medline: 33829713]

19. Blume KS, Dietermann K, Kirchner-Heklau U, et al. Staffing levels and nursing-sensitive patient outcomes: umbrella review and qualitative study. Health Serv Res. Oct 2021;56(5):885-907. [doi: 10.1111/1475-6773.13647] [Medline: 33723857]
20. Heslop L, Lu S, Xu X. Nursing-sensitive indicators: a concept analysis. J Adv Nurs. Nov 2014;70(11):2469-2482. [doi: 10.1111/jan.12503] [Medline: 25113388]
21. Koch D, Kutz A, Volken T, et al. Derivation and validation of a prediction model to establish nursing-sensitive quality benchmarks in medical inpatients: a secondary data analysis of a prospective cohort study. Swiss Med Wkly. 2022;152:w30152. [doi: 10.4414/smw.2022.w30152] [Medline: 35315266]
22. Liu Y, Avant KC, Aungsuroch Y, Zhang XY, Jiang P. Patient outcomes in the field of nursing: a concept analysis. Int J Nurs Sci. Mar 2014;1(1):69-74. [doi: 10.1016/j.ijnss.2014.02.006]
23. Oner B, Zengul FD, Oner N, Ivankova NV, Karadag A, Patrician PA. Nursing-sensitive indicators for nursing care: a systematic review (1997-2017). Nurs Open. May 2021;8(3):1005-1022. [doi: 10.1002/nop2.654] [Medline: 34482649]
24. Stalpers D, van der Linden D, Kaljouw MJ, Schuurmans MJ. Using publicly reported nursing-sensitive screening indicators to measure hospital performance: the Netherlands experience in 2011. Nurs Res. 2016;65(5):362-370. [doi: 10.1097/NNR.0000000000000170] [Medline: 27579504]
25. Tuinman A, De Greef MHG, Finnema EJ, Roodbol PF. A systematic review of the association between nursing staff and nursing-sensitive outcomes in long-term institutional care. J Adv Nurs. Aug 2021;77(8):3303-3316. [doi: 10.1111/jan.14840] [Medline: 33764569]
26. Twigg DE, Kutzer Y, Jacob E, Seaman K. A quantitative systematic review of the association between nurse skill mix and nursing-sensitive patient outcomes in the acute care setting. J Adv Nurs. Dec 2019;75(12):3404-3423. [doi: 10.1111/jan.14194] [Medline: 31483509]
27. Guidance on a data quality framework for health and social care. Health Information and Quality Authority (HIQA); 2018. URL: https://www.hiqa.ie/sites/default/files/2018-10/Guidance-for-a-data-quality-framework.pdf [Accessed 2026-11-03]
28. Westra BL, Johnson SG, Ali S, et al. Validation and refinement of a pain information model from EHR flowsheet data. Appl Clin Inform. Jan 2018;9(1):185-198. [doi: 10.1055/s-0038-1636508] [Medline: 29539649]
29. Berner S, Möller J, Müller S. Whitepaper information modelling: a fundamental tool for leaders in the digital change. foryouandyourcustomers; 2019. URL: https://foryouandyourcustomers.com/insights/whitepaper/information-modelling/ [Accessed 2025-10-13]
30. Weiskopf NG, Bakken S, Hripcsak G, Weng C. A data quality assessment guideline for electronic health record data reuse. EGEMS (Wash DC). Sep 4, 2017;5(1):14. [doi: 10.5334/egems.218] [Medline: 29881734]
31. Kahn MG, Callahan TJ, Barnard J, et al. A harmonized data quality assessment terminology and framework for the secondary use of electronic health record data. EGEMS (Wash DC). 2016;4(1):1244. [doi: 10.13063/2327-9214.1244] [Medline: 27713905]
32. Bian J, Lyu T, Loiacono A, et al. Assessing the practice of data quality evaluation in a national clinical data research network through a systematic scoping review in the era of real-world data. J Am Med Inform Assoc. Dec 9, 2020;27(12):1999-2010. [doi: 10.1093/jamia/ocaa245] [Medline: 33166397]
33. Johnson SG, Pruinelli L, Hoff A, et al. A framework for visualizing data quality for predictive models and clinical quality measures. AMIA Jt Summits Transl Sci Proc. 2019;2019:630-638. [Medline: 31259018]
34. Wieland-Jorna Y, Verheij RA, Francke AL, et al. Setting up a governance framework for secondary use of routine health data in nursing homes: development study using qualitative interviews. J Med Internet Res. Jan 25, 2023;25:e38929. [doi: 10.2196/38929] [Medline: 36696162]
35. Nantschev R. NuDaQ- framework: a strategy to reuse electronic routine nursing process data for outcome analyses. Presented at: Nursing-generated data: predictive analytics and eHealth strategies; Mar 16-18, 2023; Rome, Italy. URL: https://acendio.net/wp-content/uploads/2023/03/Proceedings-ACENDIO_14.03.2023.pdf [Accessed 2026-11-03]
36. Nantschev R. Vom datenfriedhof zur datennutzung: sekundärnutzung von pflegerischen routinedaten – chancen der digitalisierung [Article in German]. Österreichische Pflegezeitschrift. 2021. URL: https://oegkv.at/service/oepz/ [Accessed 2021-10-10]
37. Brender J, Talmon J, de Keizer N, Nykänen P, Rigby M, Ammenwerth E. STARE-HI - Statement on Reporting of Evaluation Studies in Health Informatics: explanation and elaboration. Appl Clin Inform. 2013;4(3):331-358. [doi: 10.4338/ACI-2013-04-RA-0024] [Medline: 24155788]
38. Döring N, Bortz J. Forschungsmethoden Und Evaluation in Den Sozial- Und Humanwissenschaften [Book in German]. 5th ed. Springer; 2016. [doi: 10.1007/978-3-642-41089-5]

Abbreviations

BPMN: Business Process Model and Notation

EHR: electronic health record
RWD: real-world data
SPIRIT: Systematic Planning of Intelligent Reuse of Integrated Clinical Routine Data
STARE-HI: Statement on Reporting of Evaluation Studies in Health Informatics

Edited by Elizabeth Borycki; peer-reviewed by Giovanni Gioiello, Kim De Groot; submitted 17.Mar.2026; final revised version received 14.Sep.2026; accepted 21.Oct.2026; published 16.Nov.2026

Please cite as:
Nantschev R, Ammenwerth E
A Data Strategy for the Secondary Use of Nursing Real-World Data: Feasibility Study of the NuDaK Framework
JMIR Nursing 2026;9:e95546
URL: https://nursing.jmir.org/2026/1/e95546
doi: 10.2196/95546

© Renate Nantschev, Elske Ammenwerth. Originally published in JMIR Nursing (https://nursing.jmir.org), 16.Nov.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Nursing, is properly cited. The complete bibliographic information, a link to the original publication on https://nursing.jmir.org/, as well as this copyright and license information must be included.