VIVA MEDIKA
Jurnal Kesehatan, Kebidanan, dan Keperawatan

Homepage: http://ejournal.uhb.ac.id/index.php/vm
P-ISSN: 1979-2026 E-ISSN: 2656-1034
DOI: 10.35960/vm.v19i2.2250

AI Agents for Nursing Task Augmentation: A Focused Literature Overview of Clinical, Operational, Educational, and Governance Implications

Sony Kartika Wibisono[1], Purwono Purwono[2,*], Muhammad Ahmad Baballe[3], Imam Ahmad Ashari[4], Annastasya Nabila Elsa Wulandari[5]
[1,4,5] Department of Informatics, Universitas Harapan Bangsa, Indonesia
[2] Peneliti Teknologi Teknik Indonesia, Yogyakarta, Indonesia
[3] Master of EngineeringAcademic,Nigerian Defence Academy, Kaduna, Nigeria
[1]sonykartikawibisono@uhb.ac.id, [2]purwono@ieee.org; [3]mbaballe@kanopoly.edu.ng,
[4]imamahmadashari@uhb.ac.id, [5]anstsya.new@gmail.com

ABSTRACT

The rapid development of artificial intelligence has accelerated the emergence of AI agents, defined as autonomous or semi-autonomous systems that integrate perception, contextual reasoning, decision-making, interaction, and action within defined workflows. Although AI in nursing has been widely reviewed, existing syntheses often combine predictive models, generative tools, decision-support systems, and agent-based architectures, leaving the specific contributions and implementation maturity of AI agents insufficiently differentiated. This focused literature overview examined peer-reviewed publications published between 2021 and 2025 using targeted database searches and a structured narrative synthesis. The review classified the evidence according to agent architecture, automated nursing tasks, implementation maturity, and reported clinical and operational implications. Three overlapping architectural categories were identified: LLM-driven agents, cognitive agents, and multi-agent systems. Applications were concentrated in clinical documentation and handover, predictive monitoring, medication safety, triage, and clinical decision support. The evidence suggests potential improvements in timeliness, documentation consistency, risk detection, workflow integration, and the reduction of repetitive administrative work. The maturity of the evidence varied considerably. Monitoring and sensor-enabled safety systems showed closer links to rsmitheal-world practice, whereas LLM-driven documentation and multi-agent triage systems were more frequently supported by conceptual, prototype, or simulation-based evidence. The synthesis therefore supports supervised task augmentation rather than the replacement of professional nursing judgment. Major implementation requirements include system reliability, bias mitigation, transparent accountability, human oversight, workforce competency, organizational readiness, and clinical governance. Nursing education should prepare practitioners to evaluate AI-generated outputs, recognize uncertainty and inappropriate recommendations, and apply appropriate escalation procedures. Future research should prioritize real-world, multisite, and longitudinal evaluations that measure patient safety, workload redistribution, verification burden, and clinical outcomes. This review provides a focused conceptual distinction between conventional AI decision-support tools and agent-based systems while integrating their clinical, operational, educational, and governance implications.

Keywords: Artificial Intelligence; AI Agents; Nursing Practice; Task Augmentation; Clinical Governance

*** Purwono Purwono.**
Tel.: 082113940427
Email:purwono@ieee.org

1. INTRODUCTION

The rapid advancement of artificial intelligence (AI) has increasingly influenced healthcare delivery, including nursing practice (Ramadan et al., 2024). AI technologies are widely used for large-scale data analysis and to support tasks traditionally performed by nurses, including clinical documentation, patient monitoring, risk prediction, and triage (Huang et al., 2022; Nawab, 2024). These applications range from conventional predictive models and rule-based decision-support tools to more interactive and autonomous computational systems.

Within this technological landscape, AI agents have emerged as a specific form of AI-enabled task support. In this review, an AI agent is defined as an autonomous or semi-autonomous computational entity capable of perceiving relevant inputs, maintaining or retrieving contextual information, reasoning or planning toward a defined objective, and producing an action or interaction within a workflow (Dong et al., 2024; Zheng et al., 2025; Acharya et al., 2025). Agent architectures commonly integrate functional components related to perception, planning, memory, and action rather than relying exclusively on the direct generation of predictions or textual responses.

This definition distinguishes AI agents from conventional AI decision-support systems. Conventional systems typically generate classifications, predictions, risk scores, or recommendations in response to predefined inputs, whereas the interpretation of these outputs and the subsequent actions remain entirely dependent on the clinician. Agent-based systems may additionally maintain goals, coordinate sequential tasks, interact with external tools or information sources, adapt their responses to contextual changes, and initiate predefined workflow actions. Similarly, a large language model used solely to generate text should not automatically be classified as an AI agent. Its operation becomes agent-like when it is integrated with components such as memory, planning, tool use, contextual feedback, or action-selection mechanisms.

Recent studies have reported the use of LLM-driven agents for clinical documentation and decision support (Zheng et al., 2025), cognitive assistants for adaptive monitoring and contextual support (Preum et al., 2021), and multi-agent architectures for coordinated clinical triage (Lu et al., 2024). Generative AI systems have also been evaluated for producing structured nursing documentation and diagnostic recommendations based on electronic nursing record data (Ju et al., 2025). Other AI-enabled systems support the early detection of patient deterioration, longitudinal risk monitoring, and medication error prevention in hospital and home-care settings (Ciampi et al., 2022).

These applications differ substantially in their degree of autonomy and evidence maturity. Predictive monitoring systems may continuously analyze physiological data but generally require nurses to interpret alerts and determine appropriate clinical actions. Generative documentation systems can produce text or diagnostic suggestions but remain dependent on professional verification. Multi-agent systems may coordinate multiple reasoning processes, although many have been evaluated using simulated cases or controlled datasets rather than in routine clinical practice. The term AI agen*t* should therefore not be applied uniformly to every AI-enabled nursing technology.

Existing reviews have examined AI in nursing from several complementary perspectives. Ng et al., (2022) synthesized the role of AI in clinical nursing care, whereas Martinez-Ortigosa et al., (2023) reviewed broader applications of AI in nursing care. Ventura-Silva et al., (2024) focused on the organization of nursing care, including workflow and managerial implications. Other reviews have examined AI in nursing education (Lifshits & Rosenberg, 2024) and generative AI across nursing education, practice, and research (Park et al., 2024). These studies provide important evidence regarding clinical applications, educational opportunities, workflow effects, and ethical concerns. They primarily organize AI according to clinical domains, technological functions, or implementation issues rather than according to agent architecture and agentic capabilities.

Zheng et al., (2025) provide the closest agent-focused synthesis by examining LLM-driven agents in nursing practice. Their review establishes the growing importance of language-model-based agents but concentrates primarily on LLM-driven architectures. Barrett & Jones, (2025) further distinguish task augmentation, automation, and hybridization within AI-integrated nursing work, offering a sociotechnical framework for understanding nurse–AI task allocation. Their work presents a conceptual framework rather than a synthesis of agent architectures and implementation evidence.

Despite the expansion of the AI-in-nursing literature, three conceptual and empirical gaps remain. One challenge is that the terms AI tool, generative AI, decision-support system, autonomous system, and AI agent are frequently used without consistent operational boundaries. This inconsistency can lead to conventional predictive models or standalone language-generation tools being classified as agents despite the absence of goal-directed planning, memory, interaction, or action capabilities.

Another limitation is that existing syntheses rarely compare nursing applications according to both agent architecture and evidence maturity. Conceptual proposals, simulated systems, prototype evaluations, and real-world implementations are consequently discussed together, although they support different levels of inference regarding clinical effectiveness and operational sustainability. In addition, the implications of agent-based systems extend beyond task efficiency. Increasing system autonomy affects nursing competency requirements, professional accountability, human oversight, workflow design, and clinical governance.

The scientific contribution of this focused literature overview is fourfold. It establishes an operational distinction between AI agents and conventional AI decision-support systems; organizes agent-based nursing technologies into overlapping LLM-driven, cognitive, and multi-agent architectures; differentiates conceptual and prototype evidence from real-world implementation evidence; and integrates clinical, operational, educational, competency, and governance implications. This positioning treats agent-based automation not as an entirely new and fully established paradigm but as an emerging analytical category whose clinical significance requires further empirical validation.

Accordingly, this focused literature overview examines the use of agent-based systems for nursing task augmentation across documentation, clinical handover, monitoring, risk detection, medication safety, triage, and clinical decision support. The review aims to establish an operational distinction between AI agents and conventional AI decision-support systems, classify the principal agent architectures and nursing tasks addressed, distinguish conceptual, prototype-based, simulation-based, and real-world implementation evidence, and synthesize the clinical, operational, educational, competency, and governance implications of AI agent-based nursing task augmentation. A focused literature overview was selected because the evidence remains emerging and heterogeneous and includes conceptual publications, review studies, technical prototypes, simulated evaluations, and early implementation reports.

2. METHODS

2.1 Review Design

This study employed a focused literature overview supported by a structured narrative synthesis to examine recent developments in agent-based automation in nursing practice. This approach was selected because the literature remains emerging and heterogeneous, encompassing conceptual papers, review studies, technical prototypes, simulation-based evaluations, and early implementation reports. A narrative approach is appropriate when the primary objective is to interpret, compare, and critically synthesize an evolving body of literature rather than estimate a pooled intervention effect (Gregory & Denniss, 2018). This review was not designed as a systematic review because it did not aim to exhaustively retrieve every eligible publication or quantitatively aggregate study outcome. It also differs from a scoping review because it does not seek to comprehensively map the entire evidence base through a predefined systematic screening framework. Instead, the review focuses on publications that are directly relevant to the conceptual boundaries, architectures, applications, and implications of AI agents in nursing.

2.2 Literature Search Strategy

Relevant publications were identified through targeted searches of the Scopus database and the ScienceDirect platform. The search focused on peer-reviewed journal articles and conference proceedings published between January 2021 and December 2025. This period was selected to capture recent developments in large language models, generative AI, cognitive assistants, multi-agent architectures, and autonomous clinical support systems.

Search terms represented both agent architectures and nursing application domains. The principal terms included *"AI agent," "autonomous agent," "agent-based artificial intelligence," "agentic AI," "LLM-driven agent," "large language model agent," "multi-agent system," "cognitive assistant," "generative AI," "nursing," "nursing practice," "clinical documentation," "clinical handover," "patient monitoring," "triage," "medication safety,"* and *"clinical decision support."*

The terms were combined using Boolean operators and adapted to the search functions of each database. A representative search strategy was:

("AI agent" OR "autonomous agent" OR "agentic AI" OR "LLM-driven agent*" OR "large language model agent*" OR "multi-agent system*" OR "cognitive assistant*") AND ("nurs*" OR "nursing practice") AND ("documentation" OR "handover" OR "monitoring" OR "triage" OR "medication safety" OR "clinical decision support")

Titles and abstracts were assessed for relevance to the review objectives. Full-text articles were subsequently examined when the agent architecture, nursing relevance, or implementation context could not be determined from the abstract. The search was intended to identify contemporary and conceptually relevant literature rather than establish an exhaustive systematic corpus. Accordingly, a PRISMA-style quantitative screening process was not applied.

2.3 Eligibility Criteria and Study Selection

Publications were considered eligible when they met the following criteria: they were published between 2021 and 2025, written in English, published as peer-reviewed journal articles or conference proceedings, directly relevant to nursing practice or nursing-related healthcare activities, and addressed autonomous or semi-autonomous AI capabilities involving perception, contextual reasoning, planning, interaction, coordination, decision-making, or action initiation.

Eligible publications addressed one or more nursing-related domains, including clinical documentation, nursing diagnosis, clinical handover, patient monitoring, risk prediction, medication safety, triage, workflow coordination, and clinical decision support. Review articles and conceptual studies were retained when they contributed directly to the definition, classification, ethical analysis, or governance of AI agents in nursing.

Publications were excluded if they lacked a clear relationship to nursing practice, were not peer-reviewed, fell outside the specified publication period, or discussed AI only at a general level without contributing to the review objectives. Studies that focused exclusively on conventional statistical analysis, machine learning, or predictive modeling without demonstrating agent-like reasoning, interaction, coordination, or action capabilities were not classified as direct evidence of AI agent implementation.

Conventional predictive and decision-support studies were retained as contextual evidence when they illustrated nursing tasks that could potentially be incorporated into agent-based workflows. These systems were analytically distinguished from AI agents unless they demonstrated autonomous or semi-autonomous goal-directed capabilities.

2.4 Evidence Classification and Data Synthesis

Information from each publication was organized according to publication year, study design, healthcare setting, AI system or agent type, nursing task, level of implementation, reported outcomes, and implementation challenges. To distinguish technological concepts from practical implementation, the publications were classified into three evidence categories: conceptual or review-based evidence, prototype- or simulation-based evidence, and real-world implementation evidence.

Conceptual or review-based evidence included theoretical frameworks, architectural discussions, policy analyses, and literature reviews. Prototype- or simulation-based evidence included systems evaluated using virtual patients, simulated clinical cases, experimental datasets, or controlled technical environments. Real-world implementation evidence included systems deployed or evaluated in actual clinical, organizational, community, or home-care settings.

The findings were organized into four analytical domains: types of AI agents, categories of automated nursing tasks, reported clinical and operational impacts, and implementation challenges. A structured narrative synthesis was conducted by comparing the architectural characteristics, degree of autonomy, nursing tasks, implementation context, evidence maturity, and reported implications of the identified systems. The synthesis emphasized similarities, differences, implementation gaps, and future research priorities rather than quantitative aggregation.

No formal quality appraisal or risk-of-bias assessment was undertaken because the review included heterogeneous publication types and was designed to provide a focused conceptual synthesis rather than estimate comparative intervention effectiveness. This methodological decision limits the ability to determine the relative strength of individual studies and is acknowledged as a limitation of the review.

3. RESULTS

The findings of this focused literature overview were synthesized into four analytical domains: types of AI agents identified in nursing practice, categories of automated nursing tasks, reported clinical and operational impacts, and implementation challenges. Given the absence of quantitative screening data, the results are presented as a structured thematic synthesis.

3.1 Types of AI Agents Identified

The reviewed literature supports three principal analytical categories of AI agents relevant to nursing: LLM-driven agents, cognitive agents, and multi-agent systems. These categories describe overlapping architectural properties rather than mutually exclusive system classes. They can be differentiated according to two complementary dimensions: the system's primary reasoning architecture and its coordination structure.

LLM-driven agents use a large language model as the central component for interpreting clinical language, generating structured outputs, maintaining dialogue, formulating recommendations, and coordinating interactions with external information or tools. Their agentic characteristics extend beyond standalone text generation when the language model is integrated with goal representation, memory, planning, contextual feedback, tool use, or action-selection mechanisms (Dong et al., 2024; Zheng et al., 2025). In nursing contexts, reported applications include documentation generation, nursing diagnosis recommendations, clinical handover, patient education, and decision support (Ju et al., 2025; Zheng et al., 2025).

Cognitive agents emphasize the integration of perception, contextual representation, memory, reasoning, and adaptive response. These agents are designed to interpret changes in patient or environmental conditions and modify their responses according to predefined goals or contextual developments (Preum et al., 2021). Their potential nursing applications include continuous monitoring, situational awareness, personalized assistance, and dynamic risk management. The degree of autonomy varies because many cognitive assistants continue to depend on professional interpretation and confirmation.

Multi-agent systems represent a coordination architecture in which multiple specialized agents exchange information, divide tasks, compare intermediate conclusions, or jointly produce consolidated outputs. Triageagent illustrates this architecture by coordinating several LLM-based agents to analyze clinical cases and generate structured triage decisions (Lu et al., 2024). Multi-agent coordination may support complex nursing activities that require parallel assessment, prioritization, or the integration of heterogeneous information.

The relationship among these categories is hierarchical and overlapping. An LLM-driven agent may also function as a cognitive agent when it incorporates memory, contextual reasoning, adaptation, and goal-directed action. Similarly, a multi-agent system may consist of several LLM-driven agents, cognitive agents, or a combination of both. Therefore, *LLM-driven*, *cognitive*, and *multi-agent* describe different architectural dimensions rather than three completely independent technologies.

Evidence maturity also differs across these categories. Current publications include conceptual frameworks, review-based evidence, technical prototypes, simulated evaluations, and limited real-world implementation. The identification of agentic capabilities should be distinguished from claims regarding clinical effectiveness. A system may demonstrate sophisticated reasoning or coordination in a controlled environment without establishing safety, sustainability, or effectiveness in routine nursing practice.

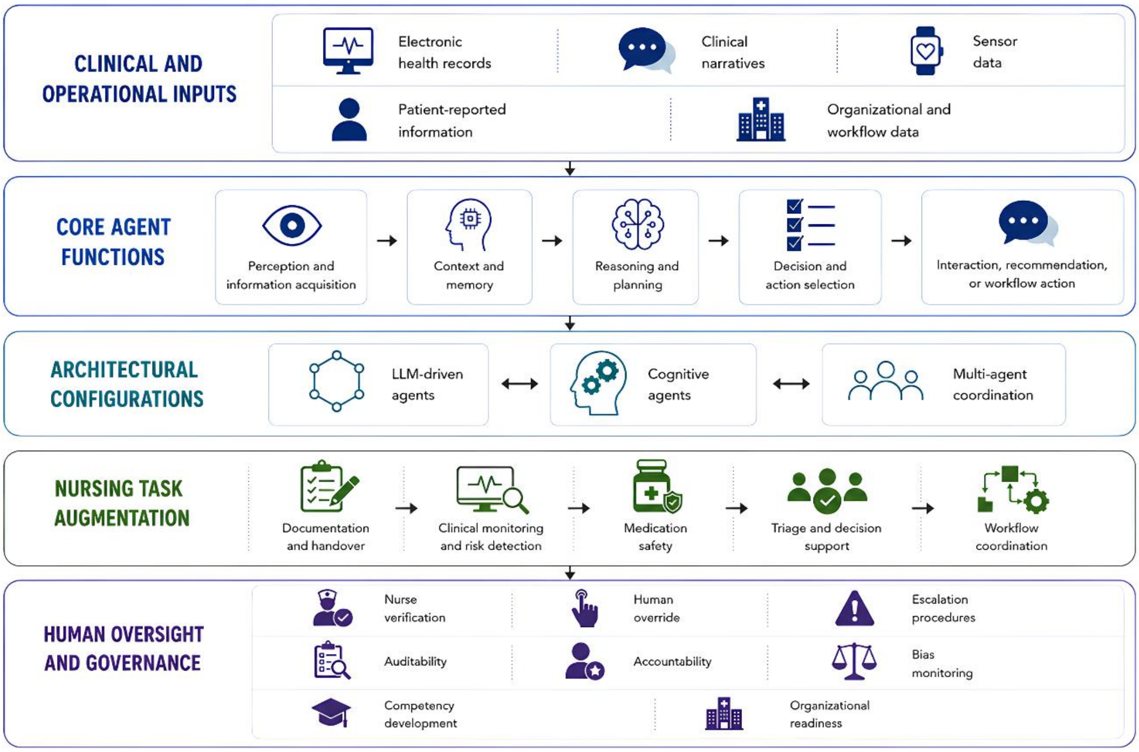

Figure 1. Conceptual Framework of AI Agents for Nursing Task Augmentation

Figure 1 illustrates that AI agent categories are interconnected architectural configurations operating through a common functional sequence. Clinical and operational data are perceived and interpreted using contextual information and memory, followed by reasoning, planning, and action selection. The resulting outputs may support documentation, monitoring, medication safety, triage, or workflow coordination. Human oversight and clinical governance encompass the entire process rather than functioning solely as a final validation step. Table 1 summarizes the three overlapping AI agent categories in nursing, including their architectural characteristics, coordination mechanisms, representative applications, and evidence maturity.

Table 1. Classification of AI Agents in Nursing Contexts

Agent category	Primary architectural basis	Defining capabilities	Coordination structure	Representative nursing applications	Evidence maturity	Sources
LLM-driven agents	Large language model as the primary language and reasoning component	Clinical language interpretation, contextual generation, dialogue, recommendation formulation, memory integration, and potential tool use	May operate as a single agent or within a coordinated multi-agent system	Nursing documentation, nursing diagnosis support, clinical handover, patient communication, and decision support	Mixed evidence, including reviews, technical prototypes, simulated evaluations, and early applied studies	(Dong et al., 2024); (Ju et al., 2025); (Zheng et al., 2025)
Cognitive agents	Integrated perception, contextual representation, memory, reasoning, and adaptive response	Continuous contextual assessment, goal-directed reasoning, adaptation, and situational support	Primarily single-agent, but may interact with external systems or other agents	Continuous patient monitoring, situational awareness, personalized assistance, and contextual risk management	Predominantly conceptual and prototype-oriented evidence in nursing-specific contexts	(Preum et al., 2021)
Multi-agent systems	Distributed architecture involving multiple specialized agents	Task decomposition, inter-agent communication, parallel reasoning, intermediate result comparison, and coordinated output generation	Multiple coordinated homogeneous o heterogeneous agents	Clinical triage, workflow coordination, complex case assessment, and prioritization	Primarily technical, prototype, and simulation-based evidence, with limited routine clinical evaluation	(Lu et al., 2024) ;(Zheng et al., 2025)

Note. The categories are not mutually exclusive. An LLM-driven agent may demonstrate cognitive-agent properties when it incorporates perception, memory, planning, adaptation, and action mechanisms. A multi-agent system may contain multiple LLM-driven agents, cognitive agents, or both. Evidence maturity refers to the predominant form of evidence reported in the reviewed literature and does not constitute a formal quality rating.

3.2 Automated Nursing Tasks

The reviewed literature identified four principal domains of AI-supported nursing tasks: documentation and clinical handover, clinical monitoring and risk detection, medication safety, and triage or clinical decision support. The degree of automation and agent autonomy differed substantially across these domains. Most systems supported partial task augmentation under human supervision rather than independently completing the entire nursing workflow.

Documentation and clinical handover represented the most prominent application domain. Generative AI systems were used to transform patient information into structured nursing documentation, propose nursing diagnoses, and consolidate clinical data for shift handover (Ju et al., 2025; Tu et al., 2025). Ju et al., (2025) evaluated an AI-assisted electronic nursing record system using virtual patient scenarios and 40 nurses. The system reduced documentation time and improved perceived ease of use, although the generated content continued to require professional verification. Because the evaluation used virtual patient records rather than routine bedside documentation, the study is best classified as a simulation-based applied evaluation rather than a full clinical deployment.

Tu et al., (2025) reported a more implementation-oriented application in which generative AI was integrated into the nursing information systems of three hospitals in Taiwan. The system automatically consolidated clinical information and generated handover documentation, with the authors reporting a substantial reduction in documentation time. This study provides evidence of operational feasibility in an organizational setting, although its short report format provides limited information regarding output errors, nurse correction rates, long-term sustainability, and patient safety outcomes.

Clinical monitoring and risk detection constituted a second major task domain. AI technologies have been used to identify physiological deterioration, analyze longitudinal patient data, and support early risk recognition in critical care and oncology settings (Ng et al., 2022; O'Connor et al., 2024). Much of this evidence concerns predictive models and conventional clinical decision-support systems rather than autonomous AI agents. These systems generally generate risk scores or alerts, while nurses remain responsible for contextual interpretation, clinical escalation, and subsequent intervention. They should therefore be considered enabling components or potential perception modules within an agent-based workflow rather than direct evidence of complete agent implementation.

Medication safety represented a more explicit example of agent-based functionality. Ciampi et al., (2022) developed an intelligent home-treatment environment comprising several specialized components. A reinforcement-learning-based Tutor agent personalized medication reminders, a Checker agent identified medications, and an Advisor service assessed potential drug–drug interactions. The architecture demonstrated perception, adaptation, task specialization, and action initiation. Its evaluation was conducted in a laboratory-emulated domestic environment. The findings therefore establish technical feasibility but do not yet demonstrate effectiveness in routine home-care or nursing practice.

Triage and clinical decision support represented the most advanced domain in terms of multi-agent coordination. TriageAgent uses several specialized LLM-based agents to analyze clinical cases, exchange intermediate assessments, and generate Emergency Severity Index classifications (Lu et al., 2024). Its performance was evaluated using three clinical triage test sets and compared with other LLM-based approaches. Although the system demonstrated improved benchmark performance, it was not evaluated in routine clinical practice. Its findings therefore support technical and simulation-based feasibility rather than verified improvements in nursing workflow, patient safety, or real-world triage outcomes.

Across the four task domains, evidence of technical capability was stronger than evidence of sustained clinical effectiveness. Documentation and handover systems showed the clearest progression toward organizational implementation, whereas multi-agent triage remained primarily benchmark-based. Predictive monitoring studies provided clinically relevant capabilities but did not consistently meet the operational definition of an AI agent. Medication safety architectures demonstrated explicit agent components but remained at the prototype-evaluation stage. These differences indicate that task classification should be considered together with agent autonomy, human oversight, and evidence maturity.

Table 2. AI-Supported Nursing Tasks, Agentic Functions, and Evidence Maturity

Nursing task domain	Agentic or AI-supported functions	Predominant level of autonomy	Required human role	Predominant evidence maturity	Sources
Documentation and nursing diagnosis	Clinical-data synthesis, text generation, documentation structuring, and	Low to moderate	Nurses verify factual accuracy, clinical relevance, terminology, and	Simulation-based evaluation and early organizational implementation	(Ju et al., 2025); (Tu et al., 2025)

	diagnosis recommendation		final documentation		
Clinical handover	Data consolidation, summarization, and generation of structured handover reports	Moderate within a predefined documentation workflow	Nurses review generated handover content, correct omissions, and communicate critical contextual information	Early real-world organizational implementation	(Tu et al., 2025)
Clinical monitoring and risk detection	Continuous data analysis, trend identification, risk prediction, and alert generation	Low in most reported systems	Nurses interpret alerts, assess the patient, and initiate escalation or intervention	Primarily review-based evidence and conventional predictive-system studies	(Ng et al., 2022); (O'Connor et al., 2024)
Medication safety	Personalized reminders, medication identification, drug–drug interaction checking, and caregiver notification	Moderate within predefined functions	Patients, caregivers, and clinicians confirm medication use and respond to warnings	Technical prototype and laboratory-emulated evaluation	(Ciampi et al., 2022))
Triage and decision support	Case interpretation, retrieval-augmented reasoning, inter-agent discussion, and priority classification	Moderate to high at the technical reasoning level	Clinicians validate triage categories and retain final decision authority	Benchmark-based and simulation-based evaluation	(Lu et al., 2024)

Note. Level of autonomy refers to the extent to which a system can independently perform predefined analytical or workflow actions. It does not indicate authority to make unsupervised clinical decisions. Evidence maturity describes the predominant evaluation context and does not constitute a formal quality or risk-of-bias assessment.

Table 3. Comparative Characteristics of Key Publications Included in the Task-Level Synthesis

Study	Evidence context	Study design	Data source	Agent category	Principal nursing-related task	Evidence category	Main evidentiary limitation
(Preum et al., 2021)	International literature	Review	Healthcare cognitive-assistant literature	Cognitive assistants	Monitoring, contextual support, and decision assistance	Conceptual or review-based	Nursing-specific real-world evaluations were limited
(Ng et al., 2022)	International literature	Scoping review	Clinical nursing-care studies	Broad AI and predictive systems	Monitoring, prediction, and clinical decision support	Review-based	Included technologies were not consistently agent-based
(Ciampi et al., 2022)	Laboratory-emulated home-	System development and	Sensors, medication images, treatment	Specialized intelligent agents and services	Medication reminders, identification, and	Prototype or laboratory-based	No routine home-care or longitudinal

	treatment context	experimental evaluation	information, and simulated domestic activities		interaction checking		clinical implementation
(Martinez-Ortigosa et al., 2023)	International literature	Systematic review	Nursing-care applications across multiple settings	Broad AI applications	Documentation, prediction, monitoring, and care support	Review-based	Heterogeneity limited direct comparison of effectiveness
(Lu et al., 2024)	Clinical-triage benchmark context	Technical development and experimental comparison	Three clinical triage test sets	Heterogeneous LLM-based multi-agent system	Triage classification and prioritization	Benchmark or simulation-based	No evaluation within routine emergency-department workflow
(O'Connor et al., 2024)	International cancer-nursing literature	Systematic review	Cancer-nursing datasets and applications	Predominantly predictive and decision-support AI	Risk prediction, symptom monitoring, and care planning	Review-based	Most included applications were development or validation studies rather than deployed agents
(Ventura-Silva et al., 2024)	International literature	Scoping review	Organization of nursing-care literature	Broad AI applications	Workflow organization, management, and prioritization	Review-based	Agent-based architecture was not the primary analytical focus
(Ju et al., 2025)	Republic of Korea	Applied comparative evaluation involving 40 nurses	Virtual patient electronic nursing-record scenarios	Generative AI-assisted documentation system	Nursing diagnosis and documentation	Simulation-based applied evaluation	Virtual cases and professional ratings do not establish routine clinical effectiveness
(Tu et al., 2025)	Taiwan	Development and organizational implementation report	Nursing information systems across three hospitals	Generative AI-assisted handover system	Clinical handover documentation	Early real-world implementation	Limited reporting on error rates, editing burden, and long-term clinical outcomes

(Zheng et al., 2025)	International literature	Scoping review	LLM-driven agent studies in nursing practice	LLM-driven agents	Decision support, education, documentation, and patient interaction	Review-based	Underlying evidence included heterogeneous and emerging applications

Table 3 demonstrates that the evidence base is methodologically heterogeneous. Several publications are scoping or systematic reviews that describe broad AI applications rather than direct AI-agent implementation (Ng et al., 2022; Martinez-Ortigosa et al., 2023; O'Connor et al., 2024; Ventura-Silva et al., 2024). Technical studies provide more explicit evidence of agent architectures but are frequently evaluated through virtual cases, laboratory environments, or benchmark datasets (Ciampi et al., 2022; Lu et al., 2024; Ju et al., 2025). Tu et al., (2025) represents a comparatively implementation-oriented example because the system was integrated into nursing information systems across three hospitals. Evidence concerning sustained accuracy, correction burden, unintended workflow effects, and patient-safety outcomes remains limited.

This distribution indicates that the current literature should not be interpreted as a uniform body of real-world clinical evidence. Review-based publications establish the breadth and relevance of AI-supported nursing functions, while prototype and benchmark studies demonstrate technical feasibility. Organizational implementation reports provide preliminary evidence of workflow integration but do not independently establish long-term clinical effectiveness. The maturity of evidence must therefore be considered when interpreting reported benefits.

3.3 Clinical Impacts

The reviewed literature identified four principal areas of potential clinical impact: earlier risk recognition, more structured clinical information, improved decision prioritization, and medication safety support. The reported outcomes, were evaluated using substantially different study designs and levels of implementation maturity. Improvements in model performance, documentation structure, or task completion time should not be interpreted as equivalent evidence of improved patient outcomes.

AI-based monitoring and predictive systems showed potential to support earlier recognition of physiological deterioration and clinical risk in critical care, oncology, and other high-acuity contexts (Ng et al., 2022; O'Connor et al., 2024). These technologies can process longitudinal patient information and identify patterns that may not be immediately apparent through intermittent observation. Most reported applications functioned as predictive or decision-support systems rather than autonomous AI agents. Nurses remained responsible for validating alerts, conducting clinical assessments, interpreting contextual factors, and determining whether escalation or intervention was required.

The distinction between predictive performance and clinical effectiveness is particularly important. O'Connor et al., (2024) found that AI applications in cancer nursing were predominantly used to develop and validate predictive models, whereas the implementation of AI-enabled tools in real-world nursing environments remained limited. Evidence that an algorithm accurately predicts deterioration or other health outcomes does not independently establish that its use improves response time, reduces adverse events, or produces better patient outcomes.

Generative AI systems demonstrated potential to improve the structure and consistency of nursing documentation and diagnosis recommendations (Ju et al., 2025). Standardized outputs may support information retrieval, clinical communication, and documentation completeness. Generated documentation, may contain omissions, unsupported statements, inappropriate terminology, or clinically inaccurate recommendations. Professional verification therefore remains necessary before AI-generated content is incorporated into the patient record or used to guide care.

Multi-agent triage architectures demonstrated another form of clinical decision support. TriageAgent coordinated specialized LLM-based agents and achieved stronger classification performance than the comparison approaches across three clinical triage test sets (Lu et al., 2024). This evidence, was derived from benchmark evaluation rather than deployment in routine emergency department practice. The study therefore supports the technical feasibility of coordinated multi-agent reasoning but does not establish improvements in nurse performance, patient waiting time, clinical safety, or patient outcomes under real-world conditions.

Medication safety systems also demonstrated agent-like capabilities through continuous sensing, medication recognition, personalized reminders, interaction checking, and automated notification mechanisms (Ciampi et al., 2022). These functions may provide an additional safety layer for patients receiving treatment at home. The reported architecture, was evaluated in a laboratory-emulated domestic environment. Its ability to reduce medication errors, prevent adverse drug events, or support sustainable home-care practice requires further evaluation with patients, caregivers, and nursing professionals in real-world settings.

Across these application domains, the strongest evidence currently concerns technical feasibility, predictive performance, documentation support, and workflow-level outputs. Evidence concerning patient safety, morbidity, continuity of care, and other clinical outcomes remains comparatively limited. The findings therefore support supervised task augmentation rather than the autonomous replacement of nursing assessment and professional judgment. Human oversight is especially important when system outputs affect triage, medication management, escalation decisions, or permanent clinical documentation.

3.4 Operational Impacts

At the operational level, AI-supported nursing systems were associated with potential changes in documentation processes, information integration, task distribution, workflow prioritization, and managerial decision support. The most consistently reported opportunities concerned the reduction of repetitive data-processing activities and the consolidation of information that would otherwise require manual review across multiple records or systems (Martinez-Ortigosa et al., 2023;Ventura-Silva et al., 2024).

Documentation represents a particularly relevant operational domain because repetitive data entry and fragmented electronic health record processes can consume time that might otherwise be allocated to clinical assessment, communication, and direct patient care. Generative and agent-supported systems may reduce selected documentation tasks by drafting structured records, extracting relevant information, and consolidating data. Task-level time savings, should not automatically be interpreted as an equivalent reduction in overall nursing workload. Nurses may need to review generated outputs, correct inaccuracies, resolve missing information, document exceptions, and manage new system-generated notifications.

AI-assisted clinical handover illustrates both the potential and the limitations of workflow automation. Automated data consolidation may reduce the time required to retrieve and organize patient information and may improve the structural consistency of handover documents (Tu et al., 2025). Handover more than the transfer of structured data. Nurses also communicate changes in patient condition, unresolved concerns, contextual observations, care priorities, and information that may not be adequately represented in the electronic record. AI-generated handover content should function as a structured draft or supporting layer rather than a substitute for direct professional communication.

AI-supported analytics may also assist managerial activities by identifying workload patterns, supporting staffing decisions, and highlighting operational priorities (Ventura-Silva et al., 2024). Staffing allocation and task prioritization, involve contextual variables that may not be fully represented in routine datasets, including patient complexity, staff experience, team composition, relational care requirements, and rapidly changing clinical conditions. Managerial recommendations generated by AI should therefore remain subject to professional review and organizational accountability.

The operational implications of AI depend on the relationship between the technology and the characteristics of the nursing task. Routine, standardized, and predictable tasks are more amenable to automation, whereas complex tasks involving uncertainty, interpersonal communication, ethical judgment, and contextual interpretation are more appropriately supported through augmentation or hybrid nurse–AI workflows (Barrett & Jones, 2025). This distinction prevents organizations from assuming that every technically automatable task should be delegated entirely to an AI system.

AI implementation may redistribute work rather than simply eliminate it. Reduced manual data entry may be accompanied by new responsibilities involving output verification, exception handling, alert management, system monitoring, and incident reporting. Poorly integrated systems may also introduce workflow interruptions, duplicate documentation, alert fatigue, or additional cognitive demands. Operational evaluations should therefore measure net workload rather than isolated task-completion time.

Future evaluations should include documentation time, verification and correction time, the frequency of system errors, alert burden, workflow interruptions, cognitive workload, nurse satisfaction, time spent in direct patient care, and effects on communication across professional teams. Longitudinal assessment is also required to determine whether initial efficiency gains remain sustainable after implementation, training, and workflow adaptation.

The available evidence suggests that AI agents may improve selected operational processes, particularly those involving information retrieval, documentation structuring, and predefined analytical tasks. Their organizational value depends on appropriate task–technology alignment, nurse participation in system design, effective workflow integration, clear role allocation, and governance mechanisms that preserve professional responsibility.

3.5 Implementation Challenges

Despite their potential benefits, the implementation of AI agents in nursing involves interconnected technical, ethical, professional, and organizational challenges. These challenges become more consequential as systems progress from generating recommendations to coordinating sequential tasks, interacting with external tools, and initiating predefined workflow actions.

Technical reliability represents a primary concern. AI-agent performance depends on the quality, completeness, representativeness, and timeliness of the data used during development and deployment. Inconsistent clinical documentation, missing data, changes in patient populations, and differences among healthcare settings may reduce model generalizability. LLM-driven agents may also produce unsupported, incomplete, or contextually inappropriate outputs. When an agent coordinates multiple reasoning steps or external tools, an error generated at one stage may influence subsequent actions and produce a cascading workflow failure (Dong et al., 2024; Zheng et al., 2025).

Evaluation should therefore extend beyond the accuracy of a single model output. Agent-based systems should be assessed at both the component and workflow levels, including data retrieval accuracy, reasoning consistency, action selection, tool interaction, failure recovery, and adherence to clinical constraints. Local validation is necessary before deployment because performance reported using benchmark datasets or simulated cases may not transfer directly to a specific healthcare organization or patient population.

Algorithmic bias represents a related concern. AI systems may reproduce or amplify disparities when training data inadequately represent particular demographic, linguistic, socioeconomic, or clinical groups. In nursing practice, biased outputs may affect risk classification, triage priority, symptom interpretation, patient education, and access to clinical resources. The ethical review literature identifies equity and algorithmic bias as major concerns, alongside patient autonomy, privacy, accountability, and the nurse–patient relationship (Alnawafleh et al., 2025).

Transparency is particularly important when an agent produces recommendations that affect clinical prioritization or patient safety. Nurses should receive sufficient information to understand the purpose of the system, the data used to generate its output, its known limitations, and the appropriate response when confidence is low or when the recommendation conflicts with clinical assessment. Complete technical explainability may not always be feasible, but operational transparency should remain essential.

Human-factor challenges include inappropriate trust, automation bias, reduced vigilance, alert fatigue, and potential deskilling. Nurses may accept AI-generated recommendations too readily when systems appear authoritative or when workload pressure limits independent verification. Poorly explained systems, in contrast, may be rejected even when they offer useful assistance. Successful implementation therefore requires calibrated trust, in which nurses understand both the capabilities and the limitations of the system.

Professional concerns also involve changes in role boundaries and responsibility. When an AI agent contributes to documentation, triage, monitoring, or medication-related decisions, responsibility for reviewing the output and acting on it must remain clearly defined. AI-generated recommendations should not create ambiguity regarding whether responsibility lies with the nurse, the healthcare organization, the software developer, or another clinical professional (Bodur et al., 2025).

Organizational readiness extends beyond the availability of digital infrastructure. Healthcare organizations require interoperable information systems, secure data governance, local validation procedures, staff training, technical support, workflow redesign, incident-reporting mechanisms, and leadership commitment. Nurses should participate in system selection, design, testing, and evaluation because implementation decisions directly affect nursing workload, patient communication, and professional accountability (Lora & Fo, 2025; Ramadan et al., 2024).

Empirical validation remains another central limitation. Explicitly agent-based applications, including TriageAgent and cognitive assistants, have predominantly been described through conceptual, technical, prototype-, or simulation-based studies (Lu et al., 2024; Preum et al., 2021). Multisite and longitudinal evaluations are needed to determine whether performance remains reliable across institutions, patient groups, workflow conditions, and periods of sustained use.

Collectively, these findings indicate that technical capability has developed more rapidly than implementation evidence, professional preparation, and governance infrastructure. Safe adoption requires coordinated attention to reliability, bias, transparency, human oversight, workforce competence, organizational readiness, and continuous post-deployment evaluation.

4. DISCUSSION

This focused literature overview extends the existing AI-in-nursing literature by examining AI agents as a distinct architectural and sociotechnical category rather than treating all AI-enabled systems as functionally equivalent. Previous reviews have primarily organized AI applications according to clinical domain, algorithmic function, educational use, organizational application, or ethical concern (Martinez-Ortigosa et al., 2023; Ng et al., 2022; Ventura-Silva et al., 2024). Although these reviews establish the relevance of AI to nursing, they do not consistently distinguish conventional predictive models, generative tools, clinical decision-support systems, and agent-based architectures according to their autonomy, reasoning, interaction, and action capabilities.

Zheng et al., (2025) provide the closest agent-focused synthesis by reviewing LLM-driven agents in nursing practice. Their analysis identified collaborative, augmentative, and interactive architectures and demonstrated that LLM-based agents may support clinical, home-based, and community nursing activities. The present review complements that work by incorporating cognitive agents, multi-agent systems, and AI-enabled systems that may function as components of future agent-based workflows. It also differentiates agent architecture from evidence maturity. A system may demonstrate planning, memory, tool use, or multi-agent coordination in a technical evaluation without establishing safety, effectiveness, or sustainability in routine nursing practice.

The reviewed evidence indicates that LLM-driven agents, cognitive agents, and multi-agent systems represent overlapping rather than mutually exclusive categories. LLM-driven agents are defined primarily by their central reasoning and language-generation technology, cognitive agents by their integration of perception, contextual reasoning, memory, and adaptation, and multi-agent systems by their coordination of several specialized agents. A multi-agent system may therefore contain LLM-driven or cognitive agents, whereas an LLM-driven agent may demonstrate cognitive properties when connected to memory, planning, external tools, and feedback mechanisms. This conceptual distinction addresses the inconsistent use of the term AI agent across the literature.

The maturity of the evidence differed substantially across nursing tasks. Among the reviewed studies, documentation and clinical handover showed relatively stronger evidence of organizational implementation than the other application domains. Generative systems have been evaluated for nursing diagnosis and documentation using virtual patient records, whereas AI-assisted handover has been incorporated into nursing information workflows (Ju et al., 2025; Tu et al., 2025). These applications demonstrate potential to reduce selected documentation activities and improve information structuring. Generated records, continue to require nurse verification because linguistic fluency does not guarantee factual accuracy, clinical relevance, or completeness.

Clinical monitoring and risk prediction were supported by a broader body of nursing literature, particularly in critical care and oncology contexts (Ng et al., 2022; O'Connor et al., 2024). Many of these applications are conventional predictive or decision-support systems rather than autonomous agents. Their contribution to agent-based nursing lies primarily in their potential role as perception or risk-assessment components within a broader workflow. This distinction is important because accurate risk prediction does not independently demonstrate that a system can interpret context, coordinate sequential actions, or improve clinical outcomes.

Medication safety and triage applications demonstrated more explicit agent-like architectures but remained less mature in terms of routine implementation. An intelligent home-treatment environment integrated specialized functions for medication reminders, identification, interaction checking, and notifications, although its evaluation was conducted in an emulated environment (Ciampi et al., 2022). TriageAgent implemented distributed reasoning among specialized LLM-based agents and outperformed comparison methods across three triage test sets, but it was evaluated as a technical benchmark rather than a deployed emergency-care intervention (Lu et al., 2024). TriageAgent therefore provides evidence of computational feasibility rather than direct evidence of improved nurse performance or patient outcomes in routine triage.

Across these domains, technical performance was demonstrated more frequently than clinical effectiveness. Improved classification accuracy, faster document generation, or more structured outputs are relevant intermediate outcomes, but they should not be treated as equivalent to improved patient safety, continuity of care, or clinical recovery. Real-world evaluations must determine whether AI-assisted processes reduce errors, improve escalation, support communication, and generate net benefits after verification time, correction burden, alert management, and workflow disruption are considered.

The findings support a supervised augmentation model rather than the autonomous replacement of nursing judgment. Barrett & Jones, (2025) distinguish augmentation, automation, and hybridization as different forms of nurse–AI interaction. Routine and predictable tasks may be more suitable for automation, whereas complex tasks involving uncertainty, communication, ethical judgment, and patient preferences require augmentation or hybrid nurse–AI workflows. This framework is particularly relevant to nursing because many clinical tasks combine structured information processing with relational and contextual judgment.

The potential operational value of AI agents therefore depends on task–technology fit. Documentation drafting, information retrieval, and predefined data consolidation may be partially automated, whereas final interpretation and communication should remain professionally supervised. Predictive monitoring may support earlier risk recognition, but nurses must assess the patient and determine the appropriate response. Multi-agent triage may structure case analysis, but final prioritization requires clinical accountability. The relevant question is not whether a task can technically be automated, but whether automation is clinically appropriate, ethically defensible, and operationally beneficial.

Implementation may also redistribute work rather than simply reduce it. Automated generation can decrease manual data entry while creating responsibilities related to output verification, exception management, alert review, system monitoring, and incident reporting. Evaluations that measure only generation time may therefore overestimate operational benefit. Future implementation studies should assess net workload, including verification time, correction frequency, false alerts, workflow interruptions, cognitive burden, and the proportion of time available for direct patient care.

Nursing education and competency development are central to safe implementation. Nurses require more than basic digital literacy. They must understand the functional differences among predictive AI, generative AI, and agent-based systems; recognize unsupported or biased outputs; evaluate uncertainty; protect patient information; and apply escalation or override procedures. Reviews of AI in nursing education identify opportunities for simulation, individualized learning, and knowledge development while also reporting limitations related to usability, technical reliability, feedback, and realistic clinical experience (Lifshits & Rosenberg, 2024; Labrague et al., 2025; Berdida et al., 2026).

Competency development should occur at foundational, supervisory, and leadership levels. Foundational competency includes AI concepts, privacy, bias, and ethical limitations. Supervisory competency includes output verification, exception handling, human override, and incident reporting. Leadership competency includes procurement assessment, local validation, workflow redesign, performance monitoring, and policy development. These competency levels represent a conceptual synthesis from the reviewed literature rather than a formally validated professional standard.

Clinical governance is equally important because increasing system autonomy can create uncertainty regarding authority and accountability. Governance frameworks should define permitted tasks, required human confirmation, escalation procedures, audit requirements, and conditions for suspending system use. Organizations should also specify the responsibilities of nurses, managers, clinical leaders, software developers, vendors, and data-governance personnel. Ethical evidence indicates that autonomy, privacy, accountability, equity, algorithmic bias, and the nurse–patient relationship are major concerns in AI-integrated nursing practice.

The need for nursing-specific governance is reinforced by policy research showing that general AI legislation does not consistently define nursing roles, protections, or professional accountability. Although the regulatory findings are specific to the United States, the broader issue is relevant across healthcare systems: institutional AI policies should not obscure the nursing scope of practice or transfer organizational and technical risks to individual nurses (Galatzan et al., 2025).

Future research should prioritize multisite, longitudinal, and real-world implementation studies. Evaluations should report not only model accuracy or task-completion time but also patient safety outcomes, error and correction rates, workload redistribution, alert burden, nurse acceptance, equity across patient groups, and effects on professional communication. Comparative studies should examine when single-agent, multi-agent, conventional decision-support, and hybrid nurse–AI approaches are most appropriate for specific nursing tasks.

The central contribution of this review is therefore not the assertion that AI agents already constitute a fully established new paradigm in nursing. Rather, the review positions agent-based systems as an emerging analytical category characterized by contextual reasoning, interaction, coordination, and action capabilities. Their value will depend on the quality of empirical validation, appropriate task allocation, nurse competency, organizational readiness, and governance mechanisms that preserve human oversight and professional accountability.

CONCLUSION

This focused literature overview indicates that AI agents represent an emerging approach to nursing task augmentation rather than a mature replacement for professional nursing practice. Their distinguishing characteristics include contextual perception, reasoning, memory, interaction, coordination, and action capabilities that extend beyond the fixed input–output functions of conventional predictive or decision-support systems.

Current applications are concentrated in documentation and clinical handover, monitoring and risk detection, medication safety, triage, and clinical decision support. The maturity of the evidence, remains uneven. Documentation and handover systems show preliminary progress toward organizational implementation, whereas multi-agent triage and several cognitive or medication-support architectures remain primarily supported by benchmark, simulation-based, or prototype evidence. Predictive monitoring contributes clinically relevant capabilities but should not be classified as agent-based implementation unless it demonstrates contextual reasoning, coordination, or action capabilities.

The available evidence supports supervised augmentation and hybrid nurse–AI workflows. Technical accuracy, faster output generation, and structured information processing are valuable intermediate outcomes, but they do not independently establish improved patient safety or clinical effectiveness. Nurses must retain the authority to verify, modify, reject, or escalate AI-generated outputs, particularly when these systems influence documentation, triage, medication management, or clinical prioritization.

Safe adoption requires reliable systems, representative data, transparent operational boundaries, human oversight, nursing competency development, organizational readiness, and clinical governance. Nursing education should prepare practitioners to critically evaluate AI-generated outputs, recognize uncertainty and bias, protect patient information, and apply appropriate override and escalation procedures.

Future research should prioritize real-world, multisite, and longitudinal evaluations that measure patient safety outcomes, error and correction rates, verification burden, workload redistribution, workflow disruption, equity, and professional accountability. The contribution of AI agents to nursing will depend not only on computational capability but also on appropriate task allocation, empirical validation, patient-centered implementation, and alignment with professional standards.

ACKNOWLEDGEMENTS

The authors would like to thank Peneliti Teknologi Teknik Indonesia and Universitas Harapan Bangsa for their support during the preparation of this manuscript.

DECLARATION OF GENERATIVE AI

During the preparation of this manuscript, the authors used ChatGPT, developed by OpenAI, for English-language editing and grammatical correction. The authors also used Scopus AI, provided by Elsevier, to support the discovery, exploration, and preliminary review of relevant scientific literature. The authors independently verified the cited sources and critically reviewed and revised all AI-assisted outputs. The authors take full responsibility for the accuracy, integrity, interpretation, citations, and final content of the published article.

REFERENCES

Acharya, D. B., Kuppan, K., & Divya, B. (2025). Agentic AI: Autonomous Intelligence for Complex Goals—A Comprehensive Survey. *IEEE Access*, *13*, 18912–18936. https://doi.org/10.1109/ACCESS.2025.3532853

Alnawafleh, K. A., Almagharbeh, W. T., Alfanash, H. A., Alasmari, A. A., Alharbi, A. A., Alamrani, M. H., Alkubati, S. A., Altayar, M. A., & Rezq, K. A. (2025). Exploring the ethical dimensions of

AI integration in nursing practice: A systematic review. *Journal of Nursing Regulation*, *16*(3), 228–237. https://doi.org/10.1016/j.jnr.2025.08.001

Barrett, J. J., & Jones, C. B. (2025). Task augmentation, automation, and hybridization in nursing: A conceptual framework for artificial intelligence-integrated care delivery. *Nursing Outlook*, *73*(5), 102524. https://doi.org/10.1016/j.outlook.2025.102524

Berdida, D. J. E., Grande, R. A. N., Serag, R. M., & Abd Elmaksoud, D. M. F. (2026). Nursing students' artificial intelligence (AI) literacy, AI self-efficacy and AI self-competency: A cross-sectional design and structural equation model analysis. *Nurse Education in Practice*, *90*, 104673. https://doi.org/10.1016/j.nepr.2025.104673

Bodur, G., Cakir, H., Turan, S., Seren, A. K. H., & Goktas, P. (2025). Artificial intelligence in nursing practice: a qualitative study of nurses' perspectives on opportunities, challenges, and ethical implications. *BMC Nursing*, *24*(1), 1263. https://doi.org/10.1186/s12912-025-03775-6

Ciampi, M., Coronato, A., Naeem, M., & Silvestri, S. (2022). An intelligent environment for preventing medication errors in home treatment. *Expert Systems with Applications*, *193*, 116434. https://doi.org/10.1016/j.eswa.2021.116434

Dong, X., Zhang, X., Bu, W., Zhang, D., & Cao, F. (2024). A Survey of LLM-based Agents: Theories, Technologies, Applications and Suggestions. *2024 3rd International Conference on Artificial Intelligence, Internet of Things and Cloud Computing Technology (AIoTC)*, 407–413. https://doi.org/10.1109/AIoTC63215.2024.10748304

Galatzan, B. J., Johnson, E. A., Turchioe, M. R., Baker, C., & Wieben, A. (2025). Regulating at the AI frontier: The collision of policy, regulation, and nursing practice. *Journal of Nursing Regulation*, *16*(3), 207–215. https://doi.org/10.1016/j.jnr.2025.08.012

Gregory, A. T., & Denniss, A. R. (2018). An Introduction to Writing Narrative and Systematic Reviews — Tasks, Tips and Traps for Aspiring Authors. *Heart, Lung and Circulation*, *27*(7), 893–898. https://doi.org/10.1016/j.hlc.2018.03.027

Huang, K., Jiao, Z., Cai, Y., & Zhong, Z. (2022). Artificial intelligence-based intelligent surveillance for reducing nurses' working hours in nurse–patient interaction: A two-wave study. *Journal of Nursing Management*, *30*(8), 3817–3826. https://doi.org/10.1111/jonm.13787

Ju, H., Park, M., Jeong, H., Lee, Y., Kim, H., Seong, M., & Lee, D. (2025). Generative AI-Based Nursing Diagnosis and Documentation Recommendation Using Virtual Patient Electronic Nursing Record Data. *Healthcare Informatics Research*, *31*(2), 156–165. https://doi.org/10.4258/hir.2025.31.2.156

Labrague, L. J., AL Sabei, S., & AL Yahyaei, A. (2025). Artificial intelligence in nursing education: A review of AI-based teaching pedagogies. *Teaching and Learning in Nursing*, *20*(3), 210–221. https://doi.org/10.1016/j.teln.2025.01.019

Lifshits, I., & Rosenberg, D. (2024). Artificial intelligence in nursing education: A scoping review. *Nurse Education in Practice*, *80*, 104148. https://doi.org/10.1016/j.nepr.2024.104148

Lora, L., & Fo, P. (2025). Nurses' perceptions of artificial intelligence (AI) integration into practice: An integrative review. *Journal of Perioperative Nursing*, *37*(3). https://doi.org/10.26550/2209-1092.1366

Lu, M., Ho, B., Ren, D., & Wang, X. (2024). TriageAgent: Towards Better Multi-Agents Collaborations for Large Language Model-Based Clinical Triage. *Findings of the Association for Computational Linguistics: EMNLP 2024*, 5747–5764. https://doi.org/10.18653/v1/2024.findings-emnlp.329

Martinez-Ortigosa, A., Martinez-Granados, A., Gil-Hernández, E., Rodriguez-Arrastia, M., Ropero-Padilla, C., & Roman, P. (2023). Applications of Artificial Intelligence in Nursing Care: A Systematic Review. *Journal of Nursing Management*, *2023*, 1–12. https://doi.org/10.1155/2023/3219127

Nawab, K. (2024). Artificial intelligence scribe: A new era in medical documentation. *Artificial Intelligence in Health*, *1*(4), 12. https://doi.org/10.36922/aih.3103

Ng, Z. Q. P., Ling, L. Y. J., Chew, H. S. J., & Lau, Y. (2022). The role of artificial intelligence in enhancing clinical nursing care: A scoping review. *Journal of Nursing Management*, *30*(8), 3654–3674. https://doi.org/10.1111/jonm.13425

O'Connor, S., Vercell, A., Wong, D., Yorke, J., Fallatah, F. A., Cave, L., & Anny Chen, L.-Y. (2024). The application and use of artificial intelligence in cancer nursing: A systematic review. *European Journal of Oncology Nursing*, *68*, 102510. https://doi.org/10.1016/j.ejon.2024.102510

Park, G. E., Kim, H., & Go, U. R. (2024). Generative artificial intelligence in nursing: A scoping review. *Collegian*, *31*(6), 428–436. https://doi.org/10.1016/j.colegn.2024.10.004

Preum, S. M., Munir, S., Ma, M., Yasar, M. S., Stone, D. J., Williams, R., Alemzadeh, H., & Stankovic, J. A. (2021). A Review of Cognitive Assistants for Healthcare. *ACM Computing Surveys*, *53*(6), 1–37. https://doi.org/10.1145/3419368

Ramadan, O. M. E., Alruwaili, M. M., Alruwaili, A. N., Elsehrawy, M. G., & Alanazi, S. (2024). Facilitators and barriers to AI adoption in nursing practice: a qualitative study of registered nurses' perspectives. *BMC Nursing*, *23*(1), 891. https://doi.org/10.1186/s12912-024-02571-y

Tu, Y.-H., Chang, T.-H., & Lo, Y.-S. (2025). *Generative AI-Assisted Nursing Handover: Enhancing Clinical Data Integration and Work Efficiency*. https://doi.org/10.3233/SHTI251283

Ventura-Silva, J., Martins, M. M., Trindade, L. de L., Faria, A. da C. A., Pereira, S., Zuge, S. S., & Ribeiro, O. M. P. L. (2024). Artificial Intelligence in the Organization of Nursing Care: A Scoping Review. *Nursing Reports*, *14*(4), 2733–2745. https://doi.org/10.3390/nursrep14040202

Zheng, X., Zou, H., Wu, L., Dong, P., Yuan, W., & Chen, Y. (2025). Large language model-driven agents in nursing practice: A scoping review. *International Journal of Nursing Sciences*, *12*(6), 532–540. https://doi.org/10.1016/j.ijnss.2025.10.007