Generative AI at the Bedside: An Integrative Review of Applications and Implications in Clinical Nursing Practice

WATSON, Adrianna, BOND, Carmel <http://orcid.org/0000-0002-9945-8577>, AVEYARD, Helen, SMITH, Graeme and JACKSON, Debra

Citation:

WATSON, Adrianna, BOND, Carmel, AVEYARD, Helen, SMITH, Graeme and JACKSON, Debra (2025). Generative AI at the Bedside: An Integrative Review of Applications and Implications in Clinical Nursing Practice. Journal of Clinical Nursing. [Article]

Generative AI at the Bedside: An Integrative Review of Applications and Implications in Clinical Nursing Practice

Abstract

Aim

The aim of this integrative review is to critically appraise and synthesize empirical evidence on the clinical applications, outcomes, and implications of generative artificial intelligence in nursing practice.

Design

Integrative review following Whittemore and Knafl's five-stage framework.

Methods

Systematic searches were performed for peer reviewed articles and book chapters published between 1 January 2018 and 30 June 2025. Two reviewers independently screened titles/abstracts and full texts against predefined inclusion/exclusion criteria focused on generative artificial intelligence tools embedded in nursing clinical workflow (excluding nursing education-only applications). Data were extracted into a standardized matrix and appraised for quality using design-appropriate checklists. Guided by Whittemore and Knafl's integrative review framework, a constant comparative analysis was applied to derive the main themes and subthemes.

Data Sources

CINAHL, MEDLINE, and Embase.

Results

Included literature was a representative mix of single-group quality improvement pilots, mixed-method usability and feasibility studies, randomized controlled trials, qualitative descriptive and phenomenological studies, as well as preliminary and proof-of-concept observational research. Four overarching themes emerged: 1) *Workflow Integration & Efficiency,* 2) *AI-Augmented Clinical Reasoning,* 3) *Patient-Facing Communication & Education,* and 4) *Role Boundaries, Ethics & Trust.*

Conclusion

Generative artificial intelligence holds promise for enhancing nursing efficiency, supporting clinical decision making, and extending patient communication. However, consistent human validation, ethical boundary setting, and more rigorous, longitudinal outcome and equity evaluations are essential before widespread clinical adoption.

Implications for the Profession and Patient Care

Although generative artificial intelligence could reduce nurses' documentation workload and routine decision-making burden, these gains cannot be assumed. Safe and effective integration will require rigorous nurse training, robust governance, transparent labeling of AI-generated content, and ongoing evaluation of both clinical outcomes and equity impacts. Without these

safeguards, generative artificial intelligence risks introducing new errors and undermining patient safety and trust.

Impact 1.1: What problem did the study address?

This study addressed the fragmented and under-synthesized evidence base on how generative artificial intelligence (GenAI) is actually used in day-to-day clinical nursing practice. Specifically, it examined the benefits, risks, and implementation challenges GenAI tools pose for nursing care at the bedside.

Impact: 1.2: What were the main findings?

All GenAI interventions in nursing required ongoing human validation and clear ethical boundaries.

GenAI tools could demonstrate potential to reduce documentation and routine task burden, streamline workflows, and improve patient communication, but evidence of long-term patient outcomes, economic impacts, and equity was limited.

Safe and effective GenAI integration depends on robust nurse training, governance, transparent content labeling, and systematic evaluation of clinical and equity outcomes.

Impact: 1.3: Where and on whom will the research have an impact?

This research may impact frontline nurses, nurse leaders, and healthcare organizations by providing practical guidance for integrating generative AI safely and ethically at the bedside.

Findings can inform policymakers, educators, and researchers globally on critical gaps and priorities, supporting evidence-based decision-making and workforce development.

Ultimately, patients and health systems worldwide will likely benefit from safer, more effective, and more equitable AI-enabled nursing care.

Reporting Method: PRISMA 2020

Patient or Public Contribution: None

Trial or Protocol Registration: Not registered

What does this paper contribute to the wider global community?

- Synthesizes international evidence on generative AI applications in bedside nursing, highlighting both benefits and critical risks.
- Defines essential safeguards (e.g., human oversight, robust governance, equity evaluation) for safe, effective clinical integration.
- Sets a global research agenda by identifying key gaps in long-term outcomes, economics, and patient equity.

Introduction

Generative artificial intelligence (GenAI) refers to advanced technologies that can produce new content, such as text, images, and simulations, by analyzing and summarizing vast datasets, and these tools are now increasingly accessible in healthcare environments (Shen et al., 2025). The rapid advancement of GenAI may hold significant promise for transforming various aspects of healthcare delivery (Saben & Dubovi, 2024). For example, applications like ChatGPT©, which utilize large language models (LLM), are able to interpret a wide range of patient information communicated in natural language, including free-text clinical notes, symptom descriptions, past medical history, medication lists, laboratory results, and imaging indications. These applications can then generate relevant clinical responses. This capability could demonstrate potential for supporting tasks such as rapid diagnostic suggestions and imaging referral recommendations through intuitive, conversational interfaces (Rosen & Saban, 2023; Rosen & Saban, 2024).

Artificial intelligence (AI) is now increasingly prevalent in healthcare, and turning a blind eye is no longer viable (Douglas et al., 2020; Jackson, 2025; Russell et al., 2021; Von Gerich et al., 2022; Watson, 2024; Watson, 2025). The critical conversations in nursing have shifted from whether we should use emerging technologies to how we should apply them responsibly and what core values must be safeguarded along the way (Jackson, 2025; Johnson & Galatzan, 2025; Watson, 2024; Yip et al., 2025).

Artificial intelligence has been progressively integrated into nursing informatics and critical care practice in multiple forms. For example, early predictive models embedded in electronic health records (EHRs) have been used to flag patients at high risk of sepsis or falls, allowing nurses to intervene earlier in the clinical trajectory (Douglas et al., 2020; Russell et al., 2021; Von Gerich et al., 2022). Expert systems such as sepsis early warning scores and drug-interaction alerts have supported nurse decision-making by providing real-time prompts at the bedside (Watson, 2024). Natural language processing (NLP) is also increasingly applied in documentation workflows, such as ambient listening technologies that automate progress notes, reducing nurses' time spent on clerical tasks (Peterson, 2023; Watson, 2024). Similarly, machine learning algorithms embedded in clinical dashboards can detect subtle physiological trends and notify nurses of deteriorating perfusion before overt shock develops (Douglas et al., 2020; Stokes & Palmer, 2020). These tools highlight how AI is already shaping daily surveillance and monitoring responsibilities. In acute care settings, AI-powered chatbots have been piloted to answer routine patient questions and provide mental health support, freeing nurses for more complex care. However, risks have also been reported; some chatbots failed to provide appropriate responses in sensitive scenarios such as self-harm or abuse disclosures (Freitas et al., 2023; Sweeney et al., 2021). Within critical care pharmacology, nurses are beginning to encounter AI-assisted titration technologies. Closed-loop vasopressor systems and AI-driven monitoring dashboards are being trialed to adjust infusion rates in real-time, shifting the nurse's role toward supervisory oversight and rapid troubleshooting rather than manual titration (Joosten et al., 2021; Porcellato et al., 2025; Watson, 2025). Studies show these systems can improve stability but still require nurse judgment to override automation when patient safety is at risk (Almagharbeh, 2025; Vincelette et al., 2025).

Nursing scholarship also highlights the ethical tensions of this technological shift. Concerns include algorithmic bias, such as cardiovascular algorithms that under-diagnose women's myocardial infarction symptoms (Starke et al., 2021), and unsafe chatbot responses that encourage harmful patient decisions (Watson, 2024). Scholars emphasize that while GenAI and other AI systems offer opportunities for efficiency, they must be implemented in ways that protect nurses' relational roles, empathy, and professional judgment (Pepito et al., 2020; Watson, 2025).

Thus, as GenAI becomes more advanced and deeply embedded in clinical systems, nursing must move beyond hesitation and toward meaningful engagement. Our task is to harness these innovations in ways that enhance patient outcomes and support nurses to give best possible care (Jackson, 2025). As noted by Ronquillo et al. (2021), it is vital that nurses are not passive recipients of change, but rather active participants in guiding the adoption of AI within care environments.

GenAI is increasingly trialed as a partner in nursing workflows, challenging and reshaping the foundational practices and values of the profession (Alruwaili et al., 2025; Bienefeld et al., 2024; McDonald, 2024; Watson, 2024). However, as McCormack (2025) warns, the adoption of digital healthcare technologies can sometimes create an 'illusion of progress,' where technological innovation is mistaken for genuine improvements in care quality. Accordingly, critical reflection is key as GenAI becomes further integrated into clinical nursing practice.

Rationale and Gap

Given the rapid advancement and increasing integration of GenAI in various healthcare settings, it is essential to understand the scope of the empirical literature on its applications in clinical nursing practice. This integrative review is timely and necessary to ensure that technological innovation translates into meaningful improvements in both nursing practice and patient outcomes. The aim of this integrative review is to critically appraise and synthesize empirical evidence on the clinical applications, outcomes, and implications of generative artificial intelligence (GenAI) in nursing practice.

Review Question

What empirical evidence exists regarding the clinical applications, outcomes, and implications of generative artificial intelligence in nursing practice since 2018?

Methods

Design

This integrative review was guided by the five-stage methodological framework proposed by Whittemore and Knafl (2005), which supports a systematic approach to identifying, appraising, and synthesizing diverse forms of empirical evidence, including quantitative, qualitative, and mixed-methods studies. The integrative review methodology was purposefully selected over a scoping review to enable not only a comprehensive mapping of the literature but also an in-depth critical appraisal and thematic synthesis of findings. This approach facilitated the evaluation of study quality, comparison across methodological designs, and the extraction of nuanced themes and patterns related to GenAI implementation in clinical nursing practice. Integrating and interpreting heterogeneous evidence allowed the results of this review to move beyond cataloguing existing knowledge and into developing practice-relevant insights, identifying gaps, and offering evidence-based recommendations for clinical application, policy, and future research.

Search Strategy

Systematic searches were conducted across three databases: Cumulative Index to Nursing and Allied Health Literature (CINAHL), Medical Literature Analysis and Retrieval System Online (MEDLINE), and Embase. This comprehensive search strategy was designed by an academic health sciences librarian using the following search string:

(genAI OR "gen AI" OR "generative AI" OR "artificial intelligence, generative") AND nurs AND (patient OR "clinical decision making" OR "clinical practice" OR "clinical applications" OR documentation)*

Limitations were placed to retrieve literature exclusively published between 1 January 2018 through 30 June 2025, reflecting the release of generative GenAI technologies. The final search was performed on 30 June 2025. Searches were augmented through citation-tracking, reference-list checks, and a structured review of relevant grey literature, including book chapters containing primary empirical data.

Inclusion and Exclusion Criteria

Included studies met the following criteria: 1) articles were published in English from January 2018 onward, 2) articles focused explicitly on GenAI tools (e.g., GPT-based chatbots, large language models) used by nurses in direct clinical workflows, and 3) articles or book chapters were directly relevant to clinical practice, including workflow efficiency, clinical decision-making, patient care or communication.

Date range of publications were from 1 January 2018 onward because 2018 marks the first year in which transformer-based and other GenAI models were publicly released, ushering in a new class of AI technologies distinct from earlier predictive or rule-based systems. Professional and scholarly dialogue about applying these generative systems in healthcare, particularly within nursing contexts, also began in 2018. Aligning the search window with this technological and conceptual inflection point ensures that included studies evaluate genuinely GenAI interventions relevant to contemporary nursing practice.

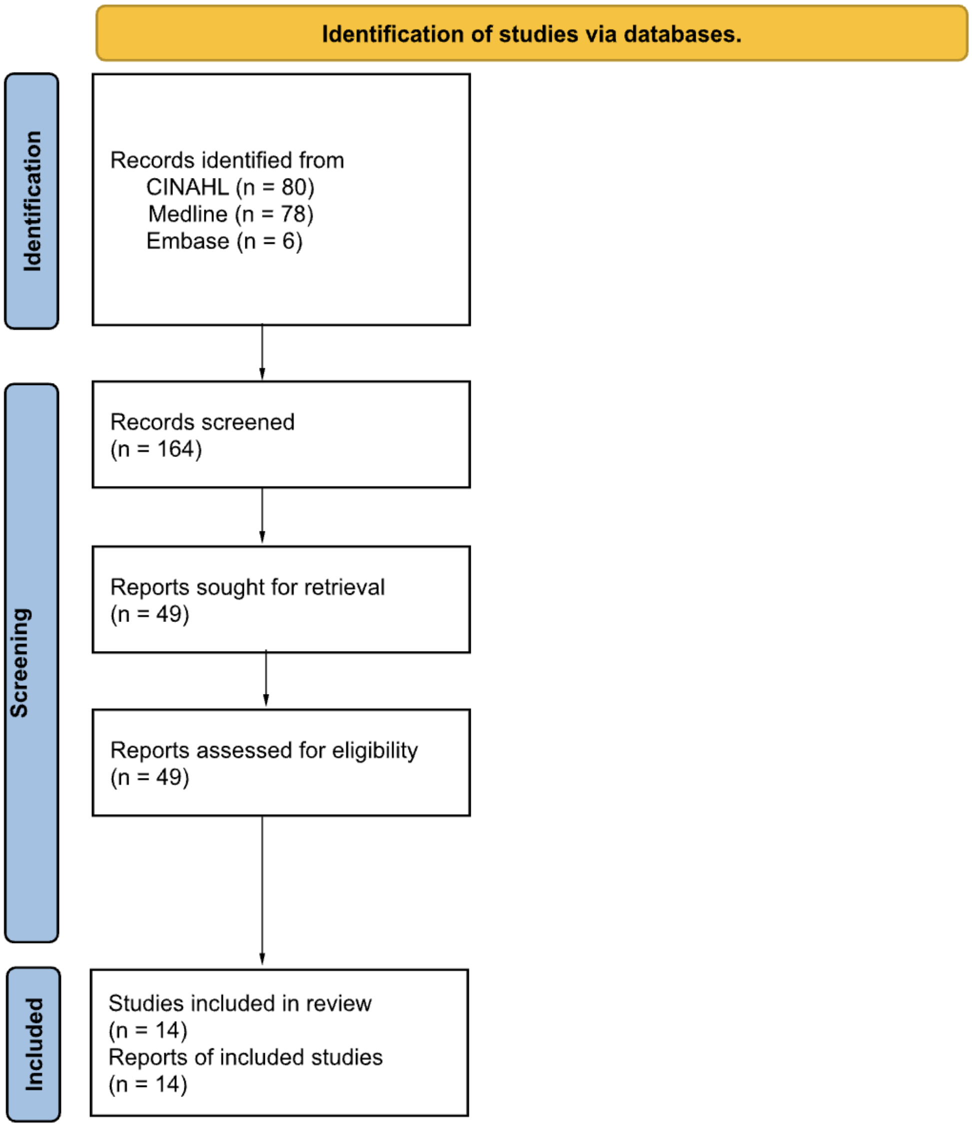
This integrative review also excluded studies whose primary aim was to use GenAI for pre-licensure, postgraduate, or continuing education purposes (e.g., virtual tutoring, simulation debriefing, examination item generation). This was integrated into the search design to better capture interventions that could be embedded in *direct clinical workflow* (e.g., documentation, triage, decision support, patient communication) as these applications have the most immediate implications for bedside practice, workload, and patient outcomes. Bracketing out education-specific research helped to maintain conceptual coherence in the synthesis and avoided conflating pedagogical innovations or concerns with tools designed to augment day-to-day clinical care.

Book-chapter contributions located during screening were treated as grey literature. Chapters that presented novel primary empirical data not duplicated in a peer-reviewed journal were retained and subjected to the same quality-appraisal rubric as the articles in the review; purely narrative, conceptual, or duplicative chapters were excluded.

Screening and Selection

After removing duplicates, two independent reviewers (XXX & XXX) conducted initial title and abstract screenings based on predefined inclusion/exclusion criteria. Conflicts at the title/abstract stage were resolved by consensus or by a third reviewer. Full-text screening followed the same independent, dual-reviewer approach, with conflicts again resolved through consensus discussion. Reasons for exclusion at the full-text level were documented systematically. Screening reproducibility was assessed at both stages. Dual independent reviewers achieved 88% agreement at the title/abstract stage and 92% agreement at full-text screening, with discrepancies resolved by a third reviewer. These agreement rates demonstrate sufficient reliability. The screening and selection process adhered to the PRISMA 2020 checklist, and outcomes are summarized in a PRISMA flow diagram (See Figure 1).

Figure 1. *PRISMA 2020 Flow Diagram*



Critical Appraisal / Quality Assessment

Given the methodological diversity of the included studies, the research team used design-specific quality appraisal tools recommended by the Joanna Briggs Institute (JBI, 2017) *Critical Appraisal Checklists*. Quantitative studies (RCTs, pre/post, cross-sectional) were evaluated for internal validity, sampling, measurement reliability, and bias. Qualitative studies were assessed on credibility, transferability, confirmability, and dependability. Mixed-method studies underwent dual appraisal, assessing both qualitative and quantitative rigor.

To enhance transparency, appraisal outcomes are reported in aggregate. JBI checklist scores across included studies ranged from 6 to 11 out of 13 possible points (mean = 8.7), indicating overall moderate-to-high methodological quality. As recommended by Whittemore & Knafl (2005), appraisal results informed interpretation of findings rather than exclusion of studies.

Ensuring Rigour

Methodological rigour and trustworthiness were enhanced by maintaining a comprehensive audit trail documenting each decision point in the review process. Dual-independent reviewer processes for screening, selection, quality appraisal, and data extraction were strictly followed. Regular team discussions ensured consensus of inclusion criteria and quality appraisal in order to help minimize bias, maximize rigor, and maintain transparency in decisions. Reflexivity was ensured by acknowledging and discussing researcher assumptions regarding GenAI's potential clinical benefits and challenges throughout analysis meetings. The standardized data extraction matrix was piloted on three studies by the review team, refined collaboratively, and then applied to the full set of included studies. This step ensured inter-reviewer consistency and helped to minimize the risk of interpretive bias during data abstraction.

Data Extraction and Synthesis

Data from included articles were systematically extracted into a standardized matrix. documenting author(s), year, study aims/objectives, settings, participants, design, data collection, analytical methods, and significant findings (see Table 1).

Table 1. *Data Matrix*

Author / Country	Research Objectives	Sample/ Setting	Study Design	Data Collection	Analysis	Key Findings
Alruwaili et al., 2025 *Saudi Arabia*	Explore NICU nurses' experiences with AI	33 NICU nurses, 4 units.	Interpretive phenomenology.	Interviews, focus groups.	Thematic analysis.	AI supports decisions, changes workflow, requires infrastructure/training, supplements nursing expertise.
Bienefeld et al., 2024 *Switzerland*	Compare data scientists' and clinicians' views on AI in ICUs.	19 data scientists, 61 ICU clinicians, 6 ICUs.	Multimethod.	Observations, survey, interviews.	Consensus, grounded theory, stats.	AI welcomed for monitoring, documentation; nurses want control for patient care.
Chen et al., 2024 *Taiwan*	Improve documentation with ChatGPT© tool.	ICU & ward nurses, Taiwan.	Development/implementation.	Needs assessment, feedback.	Descriptive/qual.	Time cut by 66%, improved accuracy/workflow, nurse oversight crucial to success.
Chen et al., 2025 *Hong Kong*	Compare AI chatbot vs. nurse hotline for anxiety/depression.	124 parents, Hong Kong.	Pilot RCT.	Pre/post questionnaires.	t-tests, regression.	Chatbot measured similar effects to nurse in mental health hotline 24/7 support, reduces anxiety/depression.

Chenard et al., 2024 *USA*	Test ChatGPT© as safe/effective for post-op ortho triage.	3 nurses, ChatGPT©, 3 surgeons.	Preliminary.	Nurse/AI responses rated by surgeons.	ANOVA, t-tests.	AI and nurse triage responses judged as similar in safety and completion of responses.
Elhilali et al., 2025 *Switzerland*	Evaluate GenAI-powered triage (STT+LLM) in ED.	Simulated triage cases, experts.	Proof of concept.	Audio, AI processing.	Accuracy (WER, ESI), comparison.	Noted potential for high triage accuracy; may reduce nurse documentation and workload.
English et al., 2024 *USA*	Assess large language model (LLM) usefulness for care teams, prompt optimization.	166 users, 9 clinics, 2023–24.	Prospective QI.	LLM draft replies, user surveys.	ANOVA, JMP Pro.	Nurses found AI helpful for efficiency, empathy, and tone in communication.
Soddu et al., 2025 *Italy*	Assess ChatGPT-4© on pressure injuries in infants.	Expert evaluation in simulation.	Cross-sectional.	AI responses, expert review.	Mixed-methods.	AI may be a safe and effective adjunct for clinical support.
Shah-Mohammadi & Finkelst	Evaluate ChatGPT-4o© for extracting	MIMIC-III dataset.	Experimental.	Automated data extraction.	Accuracy, reliability stats.	AI can extract key info for clinical informatics.

ein, 2025 *USA*	symptoms/signs from nursing notes.					
Saban & Dubovi, 2024 *Israel*	Explore ChatGPT© as clinical support for nurses.	30 ER RNs, 38 students.	Cross-sectional, scenario-based.	Online Qs, AI responses.	Comparative, descriptive.	Highlights need for careful AI integration in nursing support.
Lee et al., 2024 *Korea*	Test generative AI for reducing repetitive new nurse documentation tasks.	Univ. hospital, 30 nurses, 50K EHR entries.	Mixed-methods, pilot.	EHR data, user surveys.	Descriptive stats, qual feedback.	AI reduced documentation, improved usability, facilitated patient care.
Lee & Shin, 2025 *Korea*	Develop/evaluate chatbot for coronary artery disease recovery Q&A.	5 experts, 27 post-PCI patients.	Mixed-methods usability/feasibility.	Usage logs, surveys, interviews.	Descriptive, qualitative.	Chatbot supported education and post-discharge support, reducing clinician workload.
Huang et al., 2024 *USA*	Assess LLMs (e.g., GPT-4) for patient-friendly	156 MTurk respondents; randomized to GPT-4	Blinded, randomized survey.	Survey on interpretability, satisfaction.	Statistical survey analysis.	AI instructions rated higher for clarity and patient satisfaction compared to non-

	ED discharge instructions, patient perspective.	vs. standard instructions.				GPT discharge instructions.
Garcia et al., 2024 *USA*	Evaluate LLM use for drafting inbox replies to patient messages.	Academic medical center; attendings, APCs, nurses, pharmacists.	Prospective, single-group QI.	EHR log analysis, clinician surveys.	Pre-post, paired stats.	AI reduced clinician burden; improved communication; oversight and ongoing risk evaluation needed.

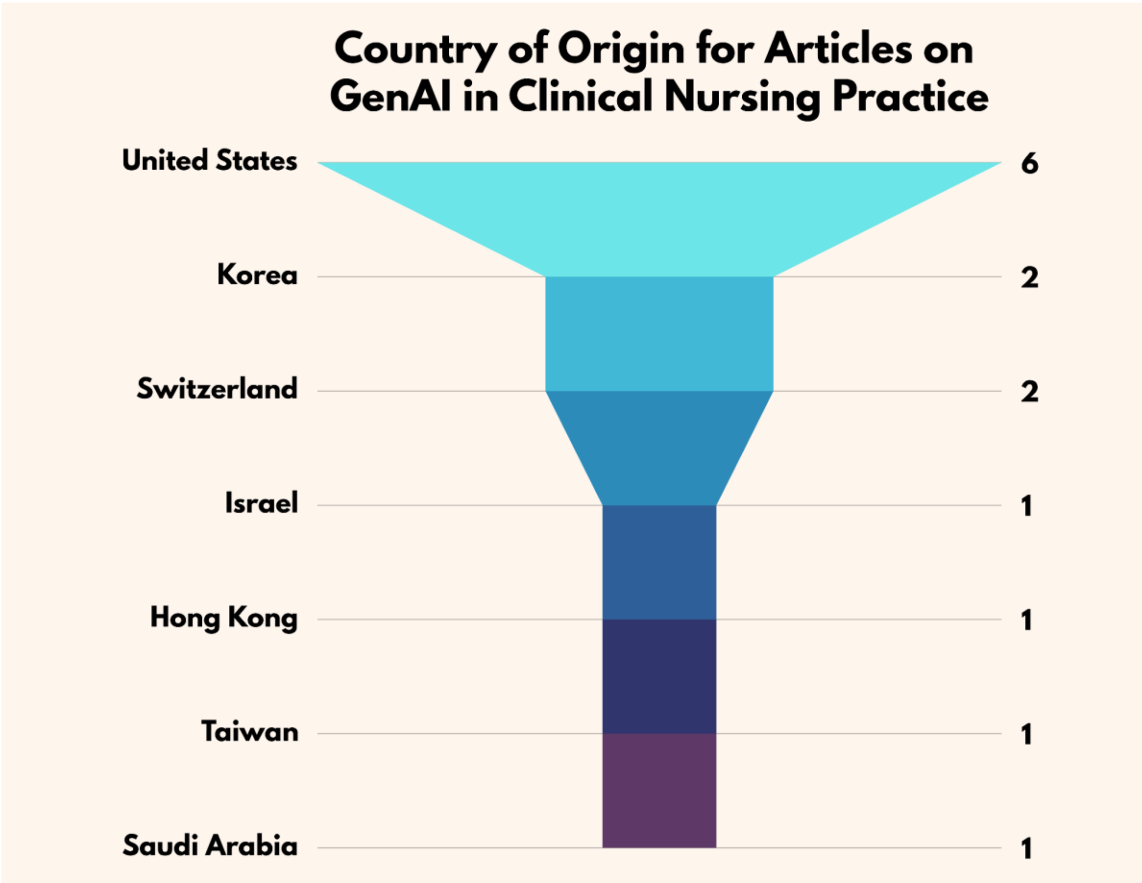
Data synthesis was guided by the constant comparative analytic method, a hallmark of integrative reviews, facilitating iterative identification and refinement of emergent themes across the studies. Each extracted study was repeatedly compared with others to identify commonalities, contrasts, and patterns. Initial thematic codes were independently generated by team members, followed by collaborative discussions to achieve interpretive consensus. Through this iterative comparison and categorization, a coherent thematic structure reflecting GenAI use in clinical nursing practice emerged, ensuring analytical rigour and interpretative depth.

Results

Study Characteristics

Articles chosen for this review totaled 14 academic articles (n = 10) and book chapters (n = 4). The articles in this review encompassed a wide range of research designs, including single-group quality improvement pilots, mixed-methods usability and feasibility studies, randomized controlled trials, cross-sectional surveys, and qualitative inquiries such as interpretive phenomenology. Those studies were conducted in varied clinical settings (e.g., emergency departments, intensive care units (ICUs), neonatal intensive care units (NICUs), and cardiology), and primarily assessed workflow efficiency, clinical reasoning enhancement, patient communication, and ethical or professional implications of GenAI. The included studies originated from a diverse range of international contexts, reflecting a global interest in GenAI applications within clinical nursing practice. Specifically, articles originated primarily in the United States (n = 6), followed by Korea (n = 2), Switzerland (n = 2), and one article each from Israel, Hong Kong, Taiwan, and Saudi Arabia. See Figure 2.

Figure 2. *International Origins of Articles*



This geographical diversity underscores the widespread recognition of GenAI's potential in nursing practice internationally, while also highlighting the need for future research to explicitly address cross-cultural differences, resource variability, and context-specific implementation challenges to enhance the generalizability and applicability of findings across diverse healthcare settings. See Table 2 for an overview of themes.

Table 2. *Overview of Main Themes and Subthemes*

Main Theme	Subthemes	Empirical Evidence	Main Points
Workflow Integration & Efficiency	1.1 Documentation time-savings	Implementation of the "A+ Nurse" documentation tool reduced average charting time from roughly 15 minutes to 5 minutes per patient (Chen et al., 2024), and a parallel generative-AI pilot demonstrated comparable efficiency gains among newly qualified nurses (Lee et al. 2024).	Automation demonstrably releases nurse time for direct care; time-motion metrics should be core outcomes.
	1.2 Messaging & inbox relief	LLM-draft replies trimmed composition time while preserving empathy (Garcia et al., 2024; English et al., 2024). AI wrote patient-message drafts so quickly that nurses and doctors spent less time typing while still sounding caring and empathetic.	Embedding AI in comms reduces invisible cognitive load. Need to track impact on burnout.
	1.3 Infrastructure, training & policy enablers	Seamless Epic/Nursing Information System integrations; 84.8 % NICU nurses requested formal AI training (Alruwaili et al., 2025). While the AI system plugged easily into	Realizing efficiency improvements is contingent on dependable information technology infrastructure, clear governance, and comprehensive workforce training.

		the hospital's Epic/NIS software, nearly 85 % of NICU nurses said they still needed formal training to feel prepared to use it.	
AI-Augmented Clinical Reasoning	2.1 Symptom / sign extraction	GPT-4o mined nurse notes with high accuracy (Shah-Mohammadi et al., 2025). (i.e., GPT-4o could read nurses' ICU notes and correctly pick out the important symptoms and signs most of the time).	AI can structure free-text, but must embed validation checkpoints.
	2.2 Triage decision support	The speech-to-text plus LLM triage pipeline achieved 90–100 % concordance with expert benchmarks for both Emergency Severity Index acuity levels and chief-complaint classifications (Elhilali et al., 2025), and an independent study found ChatGPT's© postoperative patient replies were rated as equally safe as those drafted by trained orthopaedic triage nurses (Chenard et al., 2024).	LLMs can be trained to match some levels of nurse judgement on protocolized tasks, but there is a prevalent need to watch for automation bias.
	2.3 Clinical Q&A /	Expert panel deemed ChatGPT-4© responses accurate for	Domain-specific Q&A shows promise as bedside aid; needs guardrails and

	pressure-injury coaching	infant pressure injuries (Soddu et al., 2025). ChatGPT© answering standardized clinical scenarios, the LLM's decision scores matched those of expert nurses, though performance varied with case complexity (Saban & Dubovi 2024).	liability clarity. LLMs can supplement nurse judgment on well-structured decisions, but their clinical use should remain as validated, nurse-supervised support rather than autonomous authority
Patient-Facing Communication & Education	3.1 Discharge & after-visit instructions	GPT-4 summaries were rated as more understandable by lay readers (Huang et al., 2024).	Generative AI improves readability; ethical to disclose AI authorship.
	3.2 Recovery-phase chatbots	In a mixed-methods usability evaluation, patients recovering from coronary artery disease after percutaneous coronary intervention reported high usability scores and sustained interaction rates with the AI-based recovery chatbot, underscoring its practical feasibility for post-discharge support (Lee & Shin 2025).	Well-designed chatbots can extend education without extra staffing; current need to evaluate long-term adherence.
	3.3 Mental health & crisis support	In a pilot randomised controlled trial, users of the AI-based chatbot showed statistically significant post-intervention decreases in both depression and	24/7 AI counselling can complement human lines, especially during patient surges.

		anxiety scores, effects were equivalent to those observed in the nurse-hotline comparison arm (Chen et al., 2025).	
Role Boundaries, Ethics & Trust	4.1 Role evolution & autonomy	Looking at AI use in Saudi NICUs, participants described a professional shift from routine task execution toward a consultative, analytic role in which they validated and contextualised AI-generated recommendations (Alruwaili et al., 2025).	Clinical unit or ward educators should embed informatics, prompt-crafting and critical-AI literacy training for bedside nurses.
	4.2 Safety, transparency & explainability	Across studies, 93.9 % of NICU nurses indicated that AI-generated recommendations must be confirmed by a clinician, while separate investigations underscore the necessity of explicitly labeling AI-authored content to maintain transparency and trust (Alruwaili, 2025; Garcia, 2024; Huang, 2024).	A “trust-but-verify” stance is universal; designers must surface rationale, versioning, and authorship tags.

Main Theme 1: *Workflow Integration & Efficiency*

Nurses play a central role in guiding how GenAI tools are implemented, ensuring these systems are meaningfully aligned with clinical needs and patient care priorities (Alruwaili et al., 2025; Bienefeld et al., 2024). Their expertise is essential for interpreting, validating, and contextualizing AI-generated content, as well as identifying potential workflow disruptions or unintended consequences (Chen et al., 2024; Lee et al., 2024; English et al., 2024). Without nurse leadership in both the adoption and ongoing evaluation of GenAI technologies, the risk remains that efficiency gains may be superficial - or offset by new burdens and safety concerns, rather than resulting in truly improved patient outcomes (Garcia et al., 2024; Alruwaili et al., 2025).

Subtheme 1.1: Documentation time-savings

While GenAI-powered tools such as the "A+ Nurse©" documentation assistant show substantial promise in reducing documentation workload, these efficiency gains must be interpreted with caution. Chen et al. (2024) reported that A+ Nurse© reduced average charting time from approximately 15 to 5 minutes per patient, with similar gains among newly qualified nurses in a separate GenAI pilot (Lee et al., 2024).

Automated systems cannot replicate the adaptability of experienced clinical judgment, contextual adaptation, and patient-specific insights that underpin high-quality nursing records (Chen et al., 2024; Lee et al., 2024; Alruwaili et al., 2025; Bienefeld et al., 2024). Nurses' expertise is crucial for validating, editing, and contextualizing GenAI-generated documentation, ensuring that essential clinical details are neither omitted nor misrepresented (Chen et al., 2024; Lee et al., 2024). As such, time savings should never come at the expense of thoroughness, individualized care, or safety; ongoing human oversight by nurses is essential to prevent errors, omissions, or the loss of critical information, safeguarding both patient outcomes and professional standards (Chen et al., 2024; Lee et al., 2024; Alruwaili et al., 2025; Bienefeld et al., 2024).

Subtheme 1.2: Messaging & inbox relief

GenAI systems designed to draft patient messages and manage clinical inboxes may show potential to reduce the time required for communication tasks; with the goal of still maintaining empathy and professional tone (Garcia et al., 2024; English et al., 2024). In both related studies, clinicians reported that GenAI-assisted messaging streamlined workflows. Clinicians also experienced a decrease in the "invisible" cognitive load of large message volumes, known contributors to communication-related burnout. However, these apparent gains must be interpreted with caution.

Automated responses may introduce errors, misinterpretations, or send inappropriate information to patients (Garcia et al., 2024). Nurses play a critical role in reviewing, editing, and contextualizing GenAI-generated content. English et al. (2024) and Garcia et al. (2024) observed that oversight is necessary to preserve clinical quality. Another priority in these scenarios is to ensure that sensitive or complex issues are addressed appropriately. There is also a need for clear protocols to escalate patient queries that exceed the capabilities of automated systems. Ultimately, these considerations may serve to reinforce the ongoing importance of nurse involvement in patient communication (Alruwaili et al., 2025; Garcia et al., 2024).

Subtheme 1.3: Infrastructure, training & policy enablers

Despite evidence that GenAI tools can be integrated with existing hospital infrastructure such as Epic and Nursing Information Systems, successful and safe implementation is far from guaranteed. Alruwaili et al. (2025) found that nearly 85% of NICU nurses expressed the need for formal GenAI training to feel competent and confident in using these technologies. This highlights that infrastructure alone is insufficient; sustained investment in comprehensive nurse education, explicit governance policies, and clear operational frameworks is critical to ensure both technical competency and ethical accountability (Alruwaili et al., 2025).

Main Theme 2: *AI-Augmented Clinical Reasoning*

The capacity of GenAI to augment clinical reasoning is emerging as one of its most promising, and most closely scrutinized, contributions to nursing practice. Across the reviewed literature, multiple studies illustrate how GenAI tools are being deployed to extract and synthesize clinical data, support complex triage decisions, and provide real-time guidance for clinical scenarios (Shah-Mohammadi & Finkelstein, 2025; Elhilali et al., 2025; Saban & Dubovi, 2024). These applications highlight GenAI's potential to enhance the speed, consistency, and accuracy of clinical assessments, particularly in structured or protocol-driven environments. At the same time, evidence consistently emphasizes the necessity for robust human oversight, careful system validation, and explicit safeguards to prevent automation bias and ensure that technology complements, rather than supplants, professional nursing judgment (Chenard et al., 2024; Soddu et al., 2025; Bienefeld et al., 2024).

Subtheme 2.1: Symptom / sign extraction

Recent advances in GenAI demonstrate strong potential to enhance the speed and accuracy of extracting clinically relevant information from complex datasets. For example, Shah-Mohammadi & Finkelstein (2025) found that ChatGPT-4o© was able to extract structured clinical data, such as symptoms and signs, from free-text nursing notes in a given dataset. This capability could facilitate real-time clinical analytics, supporting rapid identification of patient deterioration or evolving conditions. Similarly, studies by Chen et al. (2024) and Lee et al.

(2024) show that GenAI-powered tools can streamline the extraction and organization of documentation in everyday nursing workflows. However, the literature consistently emphasizes the need for regular validation of these systems by clinical experts, to safeguard against misclassification, ensure clinical relevance, and maintain the accuracy and integrity of extracted data (Shah-Mohammadi & Finkelstein, 2025; Lee et al., 2024).

Subtheme 2.2: Triage decision support

GenAI is also making significant inroads in supporting clinical triage and decision-making. Elhilali et al. (2025) reported that an integrated speech-to-text and LLM pipeline achieved 90–100% concordance with expert benchmarks for Emergency Severity Index (ESI) acuity classification and chief complaint coding in simulated emergency scenarios. In real-world applications, Chenard et al. (2024) found that ChatGPT©-generated replies for postoperative orthopedic triage were rated comparable to those drafted by experienced triage nurses. Yet, there is a persistent need for human oversight to avoid automation bias and to ensure that AI-supported decisions are contextually appropriate and tailored to each patient's unique presentation (Bienefeld et al., 2024; Elhilali et al., 2025; Chenard et al., 2024).

Subtheme 2.3: Clinical Q&A / pressure-injury coaching

GenAI-driven clinical Q&A tools are emerging as effective adjuncts to nursing expertise, providing rapid, evidence-informed guidance in complex clinical scenarios. Soddu et al. (2025) found that expert review affirmed the accuracy of ChatGPT-4©-generated clinical advice related to pressure injuries in infants, while Saban & Dubovi (2024) demonstrated that ChatGPT© scored comparably to expert ER nurses in standardized decision-making vignettes, though its performance varied with case complexity. These findings indicate that GenAI tools can supplement nursing judgment, particularly for well-defined clinical questions and education.

Still, the need for robust clinical oversight remains paramount, as the quality and appropriateness of GenAI outputs can fluctuate in more ambiguous or atypical cases (Saban & Dubovi, 2024; Soddu et al., 2025; Bienefeld et al., 2024).

Main Theme 3: *Patient-Facing Communication & Education*

As GenAI becomes more embedded in clinical practice, its impact on patient-facing communication and education is increasingly evident (Bhuyan et al., 2025; Park et al., 2024). The reviewed literature highlights how GenAI technologies can enhance the clarity, accessibility, and continuity of information provided to patients across the care continuum. For example, studies have shown that large language models can generate discharge and after-visit instructions that are clearer and more understandable than standard summaries, thereby improving patient comprehension and engagement (Huang et al., 2024). Additionally, GenAI-powered chatbots are emerging as valuable tools to support patients through recovery phases, offering sustained, personalized education and support after hospital discharge (Lee & Shin, 2025). Early evidence further suggests that GenAI interventions, such as AI-driven chatbots, have the capacity to deliver accessible and continuous mental health and crisis support, augmenting traditional services and expanding the reach of nursing care (Chen et al., 2025; Garcia et al., 2024).

Subtheme 3.1: Discharge & after-visit instructions

Recent studies demonstrate that large language models, such as GPT-4, can generate patient discharge instructions that are not only clearer and more readable than standard documents, but also rated higher for satisfaction by lay readers (Huang et al., 2024). This improvement in health communication suggests a meaningful opportunity to enhance patient understanding, engagement, and self-management after clinical encounters. Ethical and practical considerations emphasize the importance of transparent labeling of AI-generated materials

(Huang et al., 2024), and several studies recommend that patients be informed when content is produced or supported by GenAI to maintain trust and accountability (Garcia et al., 2024; Park et al., 2024). Continued nurse oversight is required to ensure instructions are contextually appropriate and culturally sensitive.

Subtheme 3.2: Recovery-phase chatbots

GenAI-powered chatbots are emerging as highly usable, patient-centered tools for delivering ongoing education and self-care support during the recovery phase. For instance, Lee & Shin (2025) found that patients recovering from coronary artery interventions reported high satisfaction and frequent use of a chatbot designed to answer personalized recovery questions, particularly about medication, activity, and follow-up care. The scalability of such interventions allows education and support to be extended beyond traditional clinical boundaries without placing additional burdens on nursing staff (Bhuyan et al., 2025). However, effective deployment requires that chatbots are designed with input from clinicians and patients to ensure accuracy and relevance, and ongoing evaluation of engagement and long-term adherence is warranted (Lee & Shin, 2025; Garcia et al., 2024).

Subtheme 3.3: Mental health & crisis support

There is growing evidence that GenAI chatbots can offer effective and accessible mental health and crisis support, supplementing existing nurse-led services. In a randomized controlled pilot study, Chen et al. (2025) reported that AI chatbot use significantly reduced anxiety and depression symptoms among users, with effects comparable to those seen with nurse hotlines. These tools provide 24/7 availability and scalability during surges in demand, which is particularly valuable in settings with limited mental health resources (Bhuyan et al., 2025). Nonetheless, the literature stresses the need for ongoing human oversight, clear protocols for escalation and safety, and attention to ethical standards, such as transparency and equity, especially for vulnerable populations (Garcia et al., 2024; Park et al., 2024).

Main Theme 4: *Role Boundaries, Ethics & Trust*

The integration of GenAI into nursing practice is prompting a re-examination of professional boundaries, ethical responsibilities, and the core values that underpin safe, effective care. Two critical dimensions emerge: a shift in nurses' roles and autonomy, and the ongoing imperative to uphold safety, transparency, and explainability in AI-enabled workflows.

Subtheme 4.1: Role evolution & autonomy

Across multiple studies, GenAI is shown to be reshaping nursing roles—from routine, task-focused work to more analytic, consultative, and supervisory functions. For example, Alruwaili et al. (2025) found that NICU nurses increasingly described their role as validating and contextualizing AI-generated recommendations, positioning themselves as critical decision-makers. Similarly, Chen et al. (2024) and Lee et al. (2024) highlight that nurses act as essential validators of AI outputs, exercising clinical judgment to ensure recommendations are appropriate and safe. This trend is further supported by Shah-Mohammadi & Finkelstein (2025), who observed the indispensable role of nurses in confirming the accuracy of AI-driven symptom extraction, and Soddu et al. (2025), who reported that nurse oversight remained critical even with high-performing AI tools for clinical assessment.

Subtheme 4.2: Safety, transparency & explainability

A consistent theme across the literature is the need for robust clinical validation and clear communication of GenAI outputs. Alruwaili et al. (2025) reported that nearly all NICU nurses insisted on clinician confirmation of AI-generated recommendations, reflecting a recognition that AI can support, but not replace, human judgment in safeguarding patient safety. Transparent labeling of AI-generated content is also crucial: Garcia et al. (2024) found that clinicians valued explicit identification of AI-authored messages, while Huang et al. (2024) noted that clear disclosure increased patient trust and satisfaction. Additional studies (Shah-Mohammadi & Finkelstein, 2025; Chen et al., 2024) highlight the importance of explainability and traceability, with nurses needing to understand and interrogate the rationale behind AI recommendations to ensure safe and effective integration.

Discussion

GenAI's Transformative Potential and Workflow Efficiency

This integrative review, consistent with Bhuyan et al. (2025), highlights GenAI's transformative potential in nursing, particularly in potential for streamlining workflows and possibly even reducing documentation burden. Studies by Chia-Jung Chen (2024) and Lee et al. (2024) suggest that GenAI-assisted tools can reduce charting time by as much as 66%. Such a reduction could enable nurses to shift primary focus from administrative tasks to much needed clinical care and direct patient interactions. Efficiency gains can extend to patient messaging, where English et al. (2024) and Garcia et al. (2024) found that large language models may generate empathetic, timely draft replies, which could help to decrease communication-related workload and potentially mitigate burnout.

Decision Support, Clinical Informatics, and Patient Communication

Beyond workflow efficiency, GenAI has shown promise in supporting clinical reasoning, triage, and patient communication. In structured, protocol-driven scenarios, studies by Elhilali et al. (2025), Shah-Mohammadi & Finkelstein (2025), and Saban & Dubovi (2024) found GenAI systems can match clinician accuracy for symptom extraction and triage. In patient education, Lee & Shin (2025) and Huang et al. (2024) document high ratings for clarity and satisfaction with AI-generated discharge instructions and chatbots, pointing to scalable solutions for ongoing support. Chen et al. (2025) further demonstrated that an AI chatbot was as effective as a nurse hotline for reducing anxiety and depression, with the added benefit of 24/7 accessibility.

Quality of Evidence and Research Gaps

Although early outcomes from GenAI integration in nursing are promising, critical analysis reveals significant methodological and reporting limitations. Most studies remain confined to single-center settings, short-term pilots, or simulated environments, limiting their generalizability to real-world clinical practice (Chen et al., 2024; Elhilali et al., 2025; Lee et al., 2024). Given that generative AI in healthcare and nursing is an emerging field, much of the current literature consists of early-stage pilot studies and exploratory research, which contributes to the heterogeneity and limited rigor observed in the evidence base.

The evidence base is dominated by descriptive, observational, and quality improvement studies, with rigorous randomized controlled trials and robust longitudinal research still rare (Chen et al., 2025; Huang et al., 2024). As a result, it is difficult to draw firm conclusions about causality, long-term effectiveness, or potential harms. Patient-level outcome data, formal economic evaluations, and analyses of equity impacts are largely absent, raising concerns given the widespread claims of GenAI's efficiency and scalability (Bhuyan et al., 2025; Rogers & Baker, 2024; Garcia et al., 2024; Lee & Shin, 2025). The lack of attention to differential effects across populations and algorithmic bias is a particularly significant gap, as highlighted in recent reviews (Park et al., 2024; Ruksakulpiwat et al., 2024). Without systematic evaluation of these domains (Aveyard, 2023), there is a risk of exacerbating disparities or introducing new threats to patient safety and workforce well-being (Alruwaili et al., 2025; Bienefeld et al., 2024; McDonald, 2024; Watson, 2024). In sum, the current evidence base remains provisional, underscoring the urgent need for greater methodological rigor, longer-term follow-up, and focused attention to economic, legal, and equity-related outcomes (Aveyard & Bradbury-Jones, 2019; 2021).

Critical Safeguards for Ethical and Reliable GenAI Implementation

Across the literature, robust safeguards are emphasized as essential for ethical, safe, and reliable GenAI integration in nursing. Human oversight and clinical validation are consistently identified as necessary, especially in patient-facing or high-risk scenarios (Alruwaili et al., 2025; Bienefeld et al., 2024; Soddu et al., 2025). Oversight by nurses and clinicians is required to maintain communication quality, empathy, and accountability, even when efficiency gains are achieved (English et al., 2024; Garcia et al., 2024). Prerequisites for safe integration include comprehensive nurse training in AI literacy, explicit governance frameworks, transparent labeling of AI-generated content, and ongoing monitoring for equity and patient safety (Lee et al., 2024; Hoelscher, 2024; Bhuyan et al., 2025; Rogers & Baker, 2024; Huang et al., 2024; Park et al., 2024; Ruksakulpiwat et al., 2024). Validation checkpoints, error correction, and feedback systems must be built into AI workflows to support accountability and rapid response to emerging risks (Chenard et al., 2024; Shah-Mohammadi & Finkelstein, 2025). These safeguards are the minimum requirements for GenAI to enhance, rather than undermine, person-centered nursing care. See Table 3.

Table 3. *Recommended Safeguards for Ethical GenAI Integration in Nursing Practice*

Ethical Safeguards	Description
Human Oversight	All GenAI outputs should be validated and contextualized by nurses/clinicians before use in care decisions.
Transparent Labeling	AI-generated documentation, discharge instructions, and patient messages should be explicitly identified as such.
Nurse Training & AI Literacy	Nurses should receive regular, structured education on GenAI tools, prompt crafting, ethics, and critical appraisal.

Governance & Policy Frameworks	Institutions must implement clear policies on accountability, escalation protocols, and safe use boundaries.
Equity Monitoring	Ongoing evaluation of algorithmic bias and impact across populations should be considered essential to prevent disparities.
Feedback & Error Correction	Systems should include mechanisms for reporting, correcting, and learning from AI-related errors.

The Risk of Depersonalization: Preserving Relational Care

A recurring concern is the potential for GenAI to depersonalize patient care, especially if AI systems replace or mediate key nurse-patient interactions. Strong professional resistance remains to using AI as an autonomous interface for patient engagement, with consensus that direct care should remain fundamentally human (Bienefeld et al., 2024). Nurses value GenAI's workflow benefits but insist on maintaining oversight to ensure empathy and relational care are not lost (English et al., 2024). Overreliance on AI risks eroding the human connection central to nursing, while nurse practitioners are uniquely positioned to balance technology with patient-centered, compassionate care (Rogers & Baker, 2024).

What This Integrative Review Adds: Functionalization and Conceptual Advances

Unlike prior scoping reviews that mapped the range of AI applications, this integrative review critically appraises study quality and synthesizes empirical findings to clarify how GenAI is being implemented, where it adds value, and what boundaries are required for safe, person-centered adoption (Watson, 2024; McDonald, 2024; Bhuyan et al., 2025; Hoelscher, 2024; Park et al., 2024; Ruksakulpiwat et al., 2024). The review moves beyond theoretical discussion to highlight GenAI's practical impact on workflow, decision support, communication, and mental health interventions (Chen et al., 2024; Lee et al., 2024; English et al., 2024; Garcia et al., 2024; Chen et al., 2025; Lee & Shin, 2025; Huang et al., 2024). By contextualizing these findings within nursing's core values and professional identity, this review frames relational care and clinical judgment as non-negotiable red lines for automation (Bienefeld et al., 2024; English et al., 2024; Watson, 2024). In light of ongoing calls for governance and transparency, the review identifies the need for economic evaluation, equity analysis, and long-term outcome studies to guide responsible implementation (Bhuyan et al., 2025; Garcia et al., 2024; Huang et al., 2024; Lee & Shin, 2025; McDonald, 2024; Park et al., 2024; Rogers & Baker, 2024; Ruksakulpiwat et al., 2024).

Strengths and Limitations

This integrative review is grounded in Whittemore and Knafl's (2005) recognized methodology, ensuring systematic processes for search, selection, data extraction, and synthesis (Aveyard & Bradbury-Jones, 2021). All steps, including review question development, inclusion/exclusion criteria, and search strategy, are transparently documented, supporting methodological rigor and reproducibility. Data extraction and critical appraisal were performed with standardized tools and dual independent reviewers, minimizing bias and enhancing validity. Thematic synthesis was iterative and collaborative, maintaining focus on comparability and quality appraisal. Adherence to contemporary reporting guidelines further strengthens the interpretive clarity and transparency of the review.

However, some limitations remain. The search was limited to English-language studies from three databases (CINAHL, MEDLINE, Embase), which may have led to omission of relevant evidence from other languages or sources. While grey literature (book chapters) was included, rapidly evolving GenAI innovations and unpublished studies may have been missed, introducing potential publication bias. Most included studies were short-term pilots or quality improvement projects with limited sample sizes and follow-up, restricting conclusions about long-term effectiveness and generalizability. Few studies reported patient-level outcomes, economic impacts, or equity analyses. Study quality and reporting were variable. To address these gaps, future reviews should expand database and language coverage, proactively seek unpublished or negative findings, and prioritize robust economic, equity, and longitudinal evaluations.

Nursing Implications

Nurse Clinicians

GenAI can ease documentation and routine task burdens, but nurses must remain actively involved in adoption, training, and daily validation of AI outputs (Alruwaili et al., 2025; Chen et al., 2024). Their judgment and feedback are essential for identifying workflow disruptions or threats to individualized care, and for ensuring patients are informed about GenAI's role in their care.

Nursing Leadership and Management

Clinical leaders must champion evidence-informed, transparent GenAI integration, but also be willing to review, adjust, or discontinue use if technology does not deliver safe, equitable improvements (Garcia et al., 2024; Alruwaili et al., 2025). Leadership should empower nurses to participate in procurement, workflow mapping, and establish rapid feedback mechanisms, while safeguarding relational care and patient safety as non-negotiable priorities.

Healthcare Organizations and Systems

Institutions should invest in information technology infrastructure, governance, and continuous professional development, as well as create reporting channels for issues and conduct equity audits to detect disparities (Park et al., 2024). Frontline nurse and patient input must guide all stages of selection and evaluation to ensure relevance and prevent unintended risks.

Nurse Educators and Professional Development Leaders

Education should combine GenAI training with deep emphasis on ethics, critical thinking, and relational care (Watson, 2024; Hoelscher, 2024). Simulation and hands-on practice should address not just technical skills but also transparency, consent, and effective communication.

Policymakers and Nurse Researchers

Research should target real-world outcomes, workflow, cost-effectiveness, and equity (Bhuyan et al., 2025). Policymakers should mandate open reporting, safety monitoring, and inclusion of frontline nursing perspectives. Funding should support multisite, longitudinal studies with transparent sharing of positive and negative findings.

Conclusion

This integrative review highlights GenAI as a transformative force at the bedside, offering meaningful opportunities to enhance nursing workflow efficiency, support clinical decision-making, and strengthen patient communication. However, fully realizing these promising identified benefits depends critically on thoughtful integration, robust clinician training, explicit governance frameworks, and sustained ethical vigilance. As GenAI continues to evolve, the nursing profession must remain actively engaged in shaping its trajectory to reflect core nursing values and priorities. Strategic investment, inclusive collaboration, and commitment to equity will be essential to ensuring that GenAI serves as a tool for nurse empowerment and improved clinical care for patients. With collective leadership and a clear vision, GenAI can be harnessed to advance a future of nursing that is both technologically enabled and grounded in compassionate, person-centred care.

References

Alruwaili, A. N., Alshammari, A. M., Alhaiti, A., Elsharkawy, N. B., Ali, S. I., & Ramadan, O. M. E. (2025). Neonatal nurses' experiences with generative AI in clinical decision-making: a qualitative exploration in high-risk NICUs. *BMC Nursing, 24*(1), 386.

Aveyard, H. (2023). Ebook: Doing a Literature Review in Health and Social Care: A Practical Guide 5e.

Aveyard, H., & Bradbury-Jones, C. (2021). Writing for Publication: Reporting your Literature Review. *Writing for Publication in Nursing and Healthcare: Getting it Right*, 111-121.

Aveyard, H., & Bradbury-Jones, C. (2019). An analysis of current practices in undertaking literature reviews in nursing: findings from a focused mapping review and synthesis. *BMC medical research methodology*, *19*(1), 105.

Bhuyan, S. S., Sateesh, V., Mukul, N., Galvankar, A., Mahmood, A., Nauman, M., Rai, A., Bordoloi, K., Basu, U., & Samuel, J. (2025). Generative Artificial Intelligence Use in Healthcare: Opportunities for Clinical Excellence and Administrative Efficiency. *Journal of Medical Systems*, *49*(1), 10. https://doi.org/10.1007/s10916-024-02136-1

Bienefeld, N., Keller, E., & Grote, G. (2024). Human-AI Teaming in Critical Care: A Comparative Analysis of Data Scientists' and Clinicians' Perspectives on AI Augmentation and Automation. *Journal of Medical Internet Research, 26*(1), e50130.

Chen, C., Lam, K. T., Yip, K. M., So, H. K., Lum, T. Y. S., Wong, I. C. K., Yam, J. C., Chui, C. S. L., & Ip, P. (2025). Comparison of an AI Chatbot With a Nurse Hotline in Reducing Anxiety and Depression Levels in the General Population: Pilot Randomized Controlled Trial. *JMIR Human Factors, 12,* e65785.

Chen, C.-J., Liao, C.-T., Tung, Y.-C., & Liu, C.-F. (2024). Enhancing Healthcare Efficiency: Integrating ChatGPT in Nursing Documentation. In J. Mantas et al. (Eds.), *Digital Health and Informatics Innovations for Sustainable Health Care Systems* (p. 851). IOS Press.

Chenard, S. W., Mika, A. P., Polkowski, G. G., Engstrom, S. M., Wilson, J. M., & Martin, J. R. (2024). ChatGPT provides safe responses to post-operative concerns following total joint arthroplasty. *Current Orthopaedic Practice, 35*(6), 244–249.

English, E., Laughlin, J., Sippel, J., DeCamp, M., & Lin, C.-T. (2024). Utility of Artificial Intelligence-Generative Draft Replies to Patient Messages. *JAMA Network Open, 7*(10), e2438573.

Elhilali, A., Brügger, V., Tschannen, I., Hautz, W., & Krummrey, G. (2025). AI-Enhanced Speech Recognition in Triage. In T. Bürkle et al. (Eds.), *Healthcare of the Future* 2025 (pp. 31–34). IOS Press.

Garcia, P., Ma, S. P., Shah, S., Smith, M., Jeong, Y., Devon-Sand, A., Tai-Seale, M., Takazawa, K., Clutter, D., Vogt, K., Lugtu, C., Rojo, M., Lin, S., Shanafelt, T., Pfeffer, M. A., & Sharp, C. (2024). Artificial Intelligence-Generated Draft Replies to Patient Inbox Messages. *JAMA Network Open, 7*(3), e243201.

Hoelscher, S. H. (2024). How Texas nurses can navigate the AI frontier. *Disruptive Healers, 4.* https://issuu.com/texasnurses/docs/tn-issue4-2024-digital/s/61345402

Huang, T., Safranek, C., Socrates, V., Chartash, D., Wright, D., Dilip, M., Sangal, R. B., & Taylor, R. A. (2024). Patient-Representing Population's Perceptions of GPT-Generated Versus Standard Emergency Department Discharge Instructions: Randomized Blind Survey Assessment. *Journal of Medical Internet Research, 26,* e60336.

Jackson, D. (2025), Retaining Human Connectedness in Increasingly AI Driven Environments. *Journal of Advanced Nursing*. https://doi.org/10.1111/jan.70041

Joanna Briggs Institute (JBI). (2017). *Checklist for qualitative research*. https://jbi.global/critical-appraisal-tools

Lee, G., & Shin, Y. H. (2025). Development and evaluation of a question-answering chatbot to provide information for patients with coronary artery disease after percutaneous coronary intervention. *Journal of Korean Academy of Nursing, 55*(2), 153–164.

Lee, D., Seong, M., Ju, H., & Park, M. (2024). Development of a Nursing Diagnosis/Record Generative AI System Based on Virtual Patient Data. In G. Strudwick et al. (Eds.), *Innovation in Applied Nursing Informatics* (pp. 678–680). IOS Press.

McCormack, B. (2025) Digital Healthcare and the Illusion of Progress. *Journal of Advanced Nursing*. https://doi.org/10.1111/jan.70024

McDonald, T. (2024). Do generative artificial intelligence company strategies of moving fact and breaking things in civil society cancel their social licence to operate? A nurse's evaluation of chatbot impacts. *Pacific Rim International Journal of Nursing Research, 28*(4). https://doi.org/10.60099/prijnr.2024.268964

Michalowski, M., Topaz, M. and Peltonen, L.M. (2025), An AI-Enabled Nursing Future With no Documentation Burden: A Vision for a New Reality. *Journal of Advanced Nursing.* https://doi.org/10.1111/jan.16911

Morone, G., De Angelis, L., Cinnera, A. M., Carbonetti, R., Bisirri, A., Ciancarelli, I., ... & Negrini, F. (2025). Artificial intelligence in clinical medicine: a state-of-the-art overview of systematic reviews with methodological recommendations for improved reporting. *Frontiers in Digital Health*, *7*, 1550731.

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

Rogers, J., & Baker, M. (2025). AI in practice: Opportunities, obstacles, and outlook. *The Nurse Practitioner, 50*(6), 34-38. doi: 10.1097/01.NPR.0000000000000326

Ronquillo, C. E., Peltonen, L-M., Pruinelli, L., Chu, C. H., Bakken, S., Beduschi, A., Cato, K., Hardiker, N., Junger, A., Michalowski, M., Nyrup, R., Eng, S. R., Reed, D. N., Salakoski, T., Salantera, S., Walton, N., Weber, P., Wiegand, T., & Topaz, M. (2021). Artificial Intelligence in Nursing: Priorities and Opportunities From an International Invitational Think-Tank of the Nursing and Artificial Intelligence Leadership Collaborative. *Journal of Advanced Nursing*, *77*, 3707–3717. https://doi.org/10.1111/jan.14855

Rosen, S., & Saban, M. (2023). Can ChatGPT assist with the initial triage? A case study of stroke in young females. *International Emergency Nursing*, 70, 101340. https://doi.org/10.1016/J.IENJ.2023.101340

Rosen, S., & Saban, M. (2024). Evaluating the reliability of ChatGPT as a tool for imaging test referral: A comparative study with a clinical decision support system. *European Radiology*, 1, 1–12. https://doi.org/10.1007/s00330-023-10230-0

Ruksakulpiwat, S., Thorngthip, S., Niyomyart, A., Benjasirisan, C., Phianhasin, L., Aldossary, H., Ahmed, B. H., & Samai, T. (2024). A Systematic Review of the Application of Artificial Intelligence in Nursing Care: Where are We, and What's Next? *Journal of Multidisciplinary Healthcare*, *17*, 1603. https://doi.org/10.2147/JMDH.S459946

Saban, M., & Dubovi, I. (2024). A comparative vignette study: Evaluating the potential role of a generative AI model in enhancing clinical decision-making in nursing. *Journal of Advanced Nursing*. https://doi.org/10.1111/jan.16101

Shah-Mohammadi, F., & Finkelstein, J. (2025). GPT-4 in Clinical Practice: Assessing Its Capability for Symptom Extraction from Cancer Patient Notes. In J. Mantas et al. (Eds.), *Envisioning the Future of Health Informatics and Digital Health* (p. 86). IOS Press.

Shen, M., Shen, Y., Liu, F., & Jin, J. (2025). Prompts, privacy, and personalized learning: Integrating AI into nursing education—A qualitative study. *BMC Nursing*, *24*, 470. https://doi.org/10.1186/s12912-025-03115-8

Soddu, M., De Vito, A., Madeddu, G., Nicolosi, B., Provenzano, M., Ivziku, D., & Curcio, F. (2025). Assessing the Accuracy, Completeness and Safety of ChatGPT-40 Responses on Pressure Injuries in Infants: Clinical Applications and Future Implications. *Nursing Reports, 15*(4), 130.

Watson, A. L. (2024). Ethical considerations for artificial intelligence use in nursing informatics. *Nursing Ethics, 31*(6), 1031-1040. https://doi.org/10.1177/09697330241230515

Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing, 52*(5), 546-553. https://doi.org/10.1111/j.1365-2648.2005.03621.x

Yim, D., Khuntia, J., Parameswaran, V., & Meyers, A. (2024). Preliminary evidence of the use of generative AI in health care clinical services: systematic narrative review. *JMIR Medical Informatics*, *12*(1), e52073.