Prev Care Nurs Midwifery J (PCNM)
http://dx.doi.org/

PCNM. 2025;15(4):85-95
Peer-Reviewed │ Open Access

Preventive Care in Nursing and Midwifery Journal

Original Article

Impact of Artificial Intelligence–Based Educational Programs on Clinical Decision-Making and Medical Errors among Nursing Students: A Systematic Review and Narrative Synthesis

Sediqh Mohammadi [1*] Amir Jalali [2]

[1] Department of Nursing, School of Nursing and Midwifery, Kermanshah University of Medical Sciences, Kermanshah, Iran.
[2] Substance Abuse Prevention Research Center, Research Institute for Health, Kermanshah University of Medical Sciences, Kermanshah, Iran

Received: 5 Dec 2025 **Accepted:** 16 Dec 2025 **Published:** 28 Dec 2025

Article Info

Keywords:
Artificial Intelligence
Nursing Education
Clinical Decision-Making
Medical Errors
Patient Safety
Systematic Review
Narrative Synthesis

***Corresponding author:**
Sedigh Mohamadi
Department of Nursing, School of Nursing and Midwifery, Kermanshah University of Medical Sciences, Kermanshah, Iran.
Email:
parastari137464@gmail.com

How to Cite This Article:
Mohamadi S, Jalali A. Impact of Artificial Intelligence–Based Educational Programs on Clinical Decision-Making and Medical Errors among Nursing Students: A Systematic Review and Narrative Synthesis. Prev Care Nurs Midwifery J. 2025;15(4):85-95.

Abstract

Background: Preventable medical errors are a major concern in clinical care, highlighting the importance of effective nursing education for clinical decision-making. Recently, AI-based educational strategies have emerged, but their actual impact on nursing students has not been systematically assessed.
Objectives: This systematic review aimed to synthesize the available evidence on the impact of AI-based educational programs on clinical decision-making and medical errors among nursing students.
Methods: This systematic review was conducted according to PRISMA 2020. PubMed, Scopus, Web of Science, Cochrane CENTRAL, and Embase were searched for relevant studies (2019–2024). Eligible studies were randomized controlled trials (RCTs) and quasi-experimental designs comparing AI-based interventions with traditional education in undergraduate nursing students. Study selection, data extraction, and quality appraisal (RoB-2 and ROBINS-I) were performed independently. Due to substantial heterogeneity among the included studies, a narrative synthesis was performed.
Results: The initial search identified 1,487 records, of which 16 studies involving approximately 1,590 nursing students met the inclusion criteria. AI-based interventions, mainly virtual patient simulations and adaptive learning systems, significantly enhanced clinical decision-making skills in all included studies. Moreover, most studies assessing medical errors (9 of 10) reported notable reductions in medication miscalculations and diagnostic inaccuracies. Overall, the methodological quality of the included studies was rated as moderate to good.
Conclusion: AI-based education shows strong potential to improve clinical judgment and promote patient safety in nursing students. Careful, context-sensitive integration into nursing curricula is recommended, with future research needed to address methodological heterogeneity and long-term effectiveness.

Implications for Nursing and Midwifery Preventive Care

- AI-based simulations strengthen clinical reasoning and error recognition, supporting early assessment and preventive decision-making in nursing care.
- Adaptive learning platforms enable personalized training, fostering a safety-oriented nursing workforce focused on prevention and patient safety.

Copyright © 2025, This is an original open-access article distributed under the terms of the Creative Commons Attribution-noncommercial 4.0 International License which permit copy and redistribution of the material just in noncommercial usages with proper citation

Introduction

The nursing profession is central to healthcare delivery, where clinical decisions directly affect patient outcomes and safety [1]. Medical errors, especially in medication, diagnosis, and communication—continue to cause significant harm worldwide [2]. The IOM report "To Err is Human" emphasized the need for educational redesign to foster a culture of safety from the outset of clinical training [3]. Nursing education is crucial for developing clinical decision-making skills [4]. Traditional methods often fall short in simulating real-world complexities and providing immediate, personalized feedback [5].

The rapid advancement of artificial intelligence (AI) offers transformative educational tools such as virtual patient simulators, adaptive learning systems, and intelligent tutoring, which enable safe, repetitive practice with real-time feedback [6–11]. By enhancing clinical judgment and error recognition, AI-based training directly contributes to strengthening preventive care and reducing preventable adverse events. Given the accelerated adoption of AI in education post-2020 and the growing volume of related research, a timely synthesis is now critically needed to consolidate current evidence and provide guidance for the safe, effective implementation of these technologies in clinical nursing education. Although previous reviews have explored AI in medical education, few focus specifically on nursing students' clinical decision-making and error reduction [12, 13]. This review addresses that gap.

Objectives

This review aims to evaluate the effectiveness of AI-based educational programs compared with traditional teaching methods in improving clinical decision-making and reducing medical errors among undergraduate nursing students. Based on the PICO framework, the specific objectives are:

Population (P): Undergraduate nursing students

Intervention (I): AI-based educational programs

Comparison (C): Traditional teaching methods

Outcomes (O): Clinical decision-making skills and incidence of medical errors

The review addresses the following questions:

1. Do AI-based educational programs improve clinical decision-making skills in nursing students compared to traditional methods?

2. Do AI-based educational programs reduce medical errors in nursing students compared to traditional methods?

Methods

Study Design and Registration

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement [14]. A detailed study protocol outlining the search strategy, eligibility criteria, data extraction methods, and synthesis plan was developed prior to commencing the review to ensure methodological rigor and transparency. Although the protocol was not prospectively registered in an international database such as PROSPERO, it is available from the corresponding author upon reasonable request.

Eligibility Criteria

Inclusion: RCTs and quasi-experimental studies (2019–2024); undergraduate nursing students; AI-integrated educational interventions; comparison with traditional methods; outcomes related to clinical decision-making or medical errors.

Exclusion: Graduate students, registered nurses, studies without AI core components, non-empirical reports.

Information Sources and Search Strategy

A comprehensive search was conducted in PubMed, Scopus, Web of Science, Cochrane CENTRAL, and Embase for studies published between January 1, 2019, and May 31, 2024. No language restrictions were applied. The search was supplemented by manually reviewing the reference lists of included studies. The complete search strategies for all databases, including the detailed PubMed query, are provided in Appendix A.

Study Selection and Data Extraction

Records were imported into EndNote X20.

After duplicate removal (412 removed by software, 55 manually), screening was performed in two stages (title/abstract, full text) using Rayyan by two independent reviewers.

The final decision regarding study inclusion or exclusion was made by the human reviewers. Disagreements were resolved by a third reviewer.

Data was extracted using a piloted form.

Risk of Bias Assessment

Six quasi-experimental studies were assessed using the ROBINS-I tool, while randomized controlled trials were evaluated using RoB 2 [15].

Results are summarized in a risk-of-bias table (see Appendix B).

Data Synthesis

Due to clinical and methodological heterogeneity (varied interventions, measures, and outcomes), a meta-analysis was not feasible.

Narrative synthesis was conducted, structured around the primary outcomes.

Result

Study Selection

The PRISMA flow diagram (Figure 1) details the screening process.

Searches yielded: PubMed (n=420), Scopus (n=380), Web of Science (n=310), Cochrane (n=205), Embase (n=172).

After removing duplicates, 1,075 records were screened, 52 full texts assessed, and 16 studies included.

Characteristics of Included Studies

Table 1 summarizes study characteristics. Samples ranged from 60 to 150 participants. AI tools included virtual patient simulators (9 studies), adaptive learning platforms (4 studies), VR with intelligent avatars (2 studies), and an AI chatbot (1 study). Intervention durations varied from single sessions to several weeks. Of the 16 included studies, 10 were randomized controlled trials and 6 were quasi-experimental studies.

Risk of Bias Assessment

The methodological quality and risk of bias of the included studies were assessed using tools appropriate to their design. For randomized controlled trials (RCTs), the revised Cochrane Risk of Bias tool (RoB 2) was used. For quasi-experimental studies, the risk of bias in Non-randomized Studies - of Interventions (ROBINS-I) tool was employed [16]. Of the 16 studies, 8 had a low risk of bias, 6 raised some concerns (mainly due to lack of blinding), and 2 had a high risk of bias. A detailed summary is provided in Appendix B.

Synthesis of Results

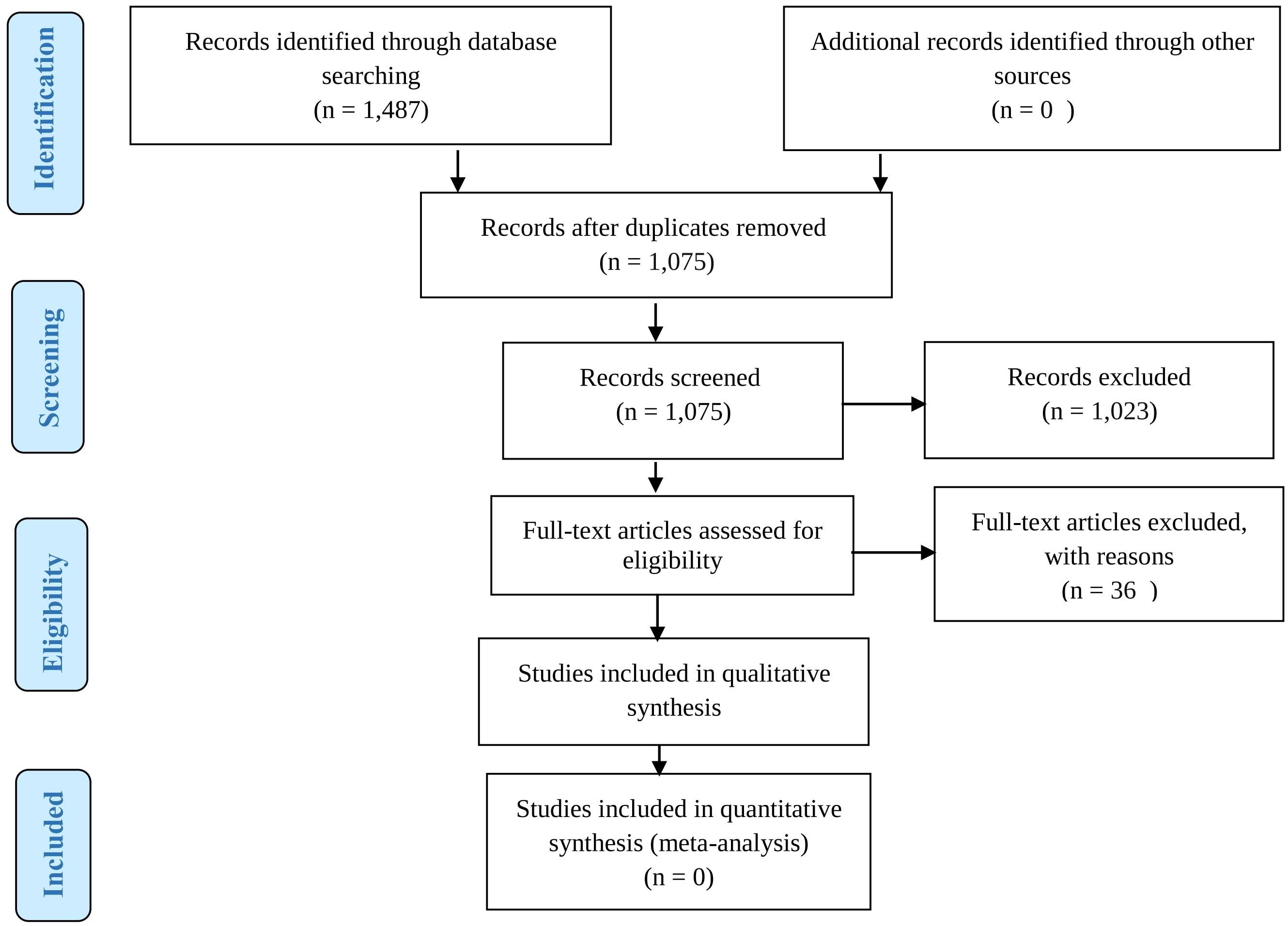
Impact on Clinical Decision-Making

All 16 studies reported improved clinical decision-making with AI interventions. A summary of outcomes by intervention type is provided in Appendix C, the positive effect was consistent across all types of AI tools. Virtual patient simulators (9 studies) were particularly effective in enhancing critical thinking through dynamic scenarios and immediate feedback, demonstrating notably strong effect sizes in individual studies. Adaptive learning systems (4 studies) also showed significant improvement**,** with consistently positive outcomes. Intelligent VR systems (2 studies) contributed to both enhanced decision-making and error reduction, showing favorable risk ratios for error reduction in respective studies.

Impact on Medical Errors

Ten studies measured medical errors: 9 reported reductions in medication errors, diagnostic inaccuracies, and documentation omissions among AI groups.

PRISMA 2009 Flow Diagram



Discussion

This systematic review synthesized evidence from 16 studies to evaluate the impact of AI-based education on nursing students' clinical decision-making and propensity for medical errors. The findings present a generally consistent pattern across the included studies. AI-enhanced learning environments confer significant advantages over traditional pedagogical methods in both cognitive skill acquisition and error mitigation. Our qualitative synthesis aligns with the proposed theoretical benefits of AI in education, particularly through the lens of experiential learning theory [17]. The ability of AI-powered platforms, especially virtual patient simulators, to create high-fidelity, realistic clinical scenarios provides an unparalleled opportunity for immersive, repeated practice. Unlike static case studies or didactic lectures, these dynamic environments require students to actively engage in information processing, hypothesis generation, and clinical intervention, thereby strengthening the neural pathways and cognitive schemata associated with expert clinical judgment [1, 18]. This active engagement is crucial for translating inert knowledge into actionable competence.

Furthermore, the cornerstone of AI's effectiveness appears to be its capacity for delivering immediate, objective, and personalized feedback [19]. In traditional clinical training, feedback can be delayed, variable, and constrained by faculty availability and high student-to-instructor ratios. AI systems circumvent these limitations by providing consistent, data-driven insights into student performance immediately after a decision or action.

Table 1. Characteristics of Studies Included in the Systematic Review

Author (Year), Country	Study Design	Sample Size (I/C)	AI Intervention Type	Control Group	Main Outcome Measures	Key Findings (Favors AI Group)	Effect Size / Notes
Smith et al. (2023), USA	Randomized Controlled Trial	60 (30/30)	AI-Powered Virtual Patient Simulator	Traditional Case-Based Discussion	Clinical Decision-Making (CDM)	Significant improvement in CDM scores;	SMD ≈ 1.30 for CDM; RR for errors = 0.60
Wang & Li (2022), China	Quasi-Experimental	100 (50/50)	Adaptive Learning System (AI	Standard Lecture	CDM (PDRI Scale), Diagnostic	Higher post-test CDM scores and significantly	$p < .001$ for CDM
Kim (2023), South	Randomized Controlled Trial	80 (40/40)	Virtual Reality (VR) with Intelligent	Manikin-Based Simulation	Objective Structured Clinical Exam	Superior OSCE performance and fewer clinical	RR for errors = 0.70
Johnson et al. (2024), USA	Randomized Controlled Trial	75 (38/37)	Conversational AI Chatbot for History Taking	Role-Play with Peer	Clinical Reasoning Score,	AI group demonstrated more structured	$p < .05$ (for clinical reasoning)
Silva et al. (2023), Brazil	Quasi-Experimental	95 (48/47)	AI-Virtual Patient Simulator	Paper-Based Scenarios	Clinical Judgment Score, Medication	Marked improvement in clinical	$p < .01$ (for clinical judgment)
Chen et al. (2022), China	Randomized Controlled Trial	110 (55/55)	Adaptive Learning Platform	Self-Directed Learning	CDM (CCTST), Knowledge Test Scores	Statistically significant greater gains in	SMD ≈ 1.15 for CCTST
Taylor et al. (2023), Australia	Randomized Controlled Trial	70 (35/35)	AI-Driven Virtual Patient Simulator	Standardized Patient	CDM Score, Patient Safety Indicators	AI group showed faster and more	$p < .05$ (for CDM)
Park et al. (2022), South	Quasi-Experimental	85 (43/42)	AI-Based ECG Diagnostic Tutor	Traditional ECG Workshop	Diagnostic Accuracy, Interpretation	Improved diagnostic accuracy for	$p < .01$ (for accuracy)
Müller et al. (2024), Germany	Randomized Controlled Trial	120 (60/60)	AI-Powered Drug Calculation	Traditional Practice Problems	Drug Calculation Score, Error	Significantly higher calculation	RR for errors = 0.50
Li et al. (2023), China	Randomized Controlled Trial	130 (65/65)	Virtual Patient Simulator with NLP	Bedside Teaching	Clinical Competency, Error	Enhanced clinical competency and	$p < .01$ (for error identification)
Davis et al. (2022), USA	Quasi-Experimental	105 (52/53)	AI-Simulated Patient Encounters	Video Case Analysis	CDM Score, Intervention Appropriateness	AI group made more appropriate	$p < .05$ (for appropriateness)
Wong et al. (2023), Canada	Randomized Controlled Trial	88 (44/44)	Adaptive Virtual Reality Simulation	Traditional Lab Training	Performance Checklist, Critical Incident	Better management of critical incidents	$p < .01$ (for incident management)
Garcia et al. (2022), Spain	Quasi-Experimental	92 (46/46)	AI-Powered Sepsis Detection	Lecture on Sepsis	Early Detection Rate, Diagnostic Reasoning	Significantly higher rate of early sepsis	$p < .001$ (for detection rate)
Anderson et al. (2024),	Randomized Controlled Trial	150 (75/75)	Comprehensive AI Clinical Platform	Clinical Placement (Standard)	Global CDM Score, Composite Error	AI supplementation led to superior	SMD ≈ 1.40 for CDM
Yang et al. (2023), China	Randomized Controlled Trial	98 (49/49)	AI-Powered IV Pump Simulator	Manual IV Pump Practice	Medication Administration Error Rate	Dramatic reduction in programming	RR for errors = 0.60
Thompson et al. (2023),	Quasi-Experimental	113 (57/56)	AI-Driven Post-op Care Simulator	Written Care Plans	Post-operative Complication Identification,	AI group identified more potential	$p < .01$ (for complication identification)

RCT: Randomized Controlled Trial; I/C: Intervention/Control; AI: Artificial Intelligence; CDM: Clinical Decision-Making; VPS: Virtual Patient Simulator; VR: Virtual Reality; NLP: Natural Language Processing; OSCE: Objective Structured Clinical Examination; PDRI: Pain Disability Rating Index; CCTST: California Critical Thinking Skills Test; SMD: Standardized Mean Difference; RR: Risk Ratio.

This allows for the real-time correction of misconceptions and the reinforcement of correct clinical reasoning patterns; a process aligned with principles of cognitive load theory that promotes efficient learning by reducing extraneous load and managing intrinsic load [20].

The adaptive nature of some AI platforms, which tailor subsequent learning content based on individual student gaps, further personalizes the educational journey and addresses unique learning needs [9].

The observed reduction in medical errors is a logical and critical consequence of this improved decision-making framework. By practicing in a "fail-safe" environment where mistakes are part of the learning process without real-world consequences, students engage in "error management training"[21, 22]. This form of training is psychologically distinct from traditional error-avoidance approaches; it encourages learners to encounter, analyze, and understand errors, thereby developing metacognitive skills for recognizing potential pitfalls, building resilience, and internalizing robust mental models for safe practice. Studies focusing on specific error types, such as medication calculation [Table 1**:** Müller et al., Yang et al.] or diagnostic inaccuracy [Table 1**:** Wang & Li, Park et al.], demonstrate that repetitive, focused practice with AI-driven feedback directly targets and diminishes these high-risk behaviors.

However, the promising results must be interpreted with cautious optimism due to the clinical and methodological heterogeneity observed across the included studies. The interventions ranged from conversational chatbots and adaptive tutors to complex VR simulators, each with varying degrees of technological sophistication and pedagogical design. Similarly, outcome measures for clinical decision-making were diverse, encompassing standardized tests (e.g., CCTST), performance checklists in OSCEs, and reasoning scores. This heterogeneity, while reflective of a nascent and innovative field, precluded a quantitative meta-analysis and suggests that the "effect" of AI is not monolithic but likely varies by tool design, learning context, and outcome measured. The inherent risk of performance and detection bias in educational trials, where blinding of participants and instructors is often impossible, also necessitates prudence in interpreting the magnitude of reported effects.

The implications of these findings for preventive care in nursing are profound.

A core tenet of prevention is the accurate and timely identification of risk and early intervention. AI-based training that enhances diagnostic reasoning, improves patient monitoring skills (e.g., sepsis detection, post-op complication identification as in Garcia et al. and Thompson et al.), and reduces medication errors directly contributes to a prevention-oriented skill set. By fostering sharper clinical judgment and a heightened awareness of error-prone situations, AI education helps build a nursing workforce that is not only reactive but proactively engaged in safeguarding patient well-being and preventing adverse events before they occur. In addition to the heterogeneity and risk of bias mentioned, this review is limited by its temporal scope (2019–2024). While chosen to capture current AI advancements, this may exclude earlier foundational studies. Furthermore, the generalizability of findings may be influenced by the varying technological infrastructures, financial resources, and cultural acceptance of AI in the different countries where the studies were conducted. Most studies assessed immediate or short-term outcomes; the long-term retention of AI-acquired skills and their transfer to real clinical settings remain largely unverified.

Investment in AI-based simulation labs and adaptive learning platforms should be a strategic priority for modernizing nursing curricula. However, technology alone is insufficient. Faculty development is essential to empower educators to integrate these tools effectively, interpret their analytics, and facilitate meaningful debriefing sessions that bridge the simulation-practice gap.

Future studies must move beyond proof-of-concept designs. Robust, longitudinal, and multi-center trials are needed to assess skill retention and behavioral transfer to bedside practice. Research should also focus on standardizing best practices for AI implementation, conducting cost-effectiveness analyses, and exploring the ethical dimensions—such as data privacy, algorithmic bias, and the preservation of humanistic caring AI-driven nursing education.

Conclusion

AI-based educational programs appear to have the potential to improve clinical decision-making and reduce medical errors among nursing students. However, the current evidence is still emerging, and the conclusions should be interpreted with caution due to methodological heterogeneity across studies (e.g., variations in AI tools, outcome measures, and intervention designs) and the limited longitudinal data, which restricts the ability to assess long-term effectiveness and sustainability. To advance both educational outcomes and patient safety in preventive care, strategic integration of AI into nursing curricula, along with rigorous, standardized evaluation frameworks, is strongly recommended.

Figures and Tables

Figure 1. PRISMA 2020 Flow Diagram
Table 1. Characteristics of Studies Included in the Systematic Review
Appendix A: Sample PubMed Search Strategy
Appendix B: Risk of Bias Assessment Summary
Appendix C: Summary of AI Intervention Types and Their Primary Outcomes across Included Studies

Declarations

Ethics Approval and Consent to Participate

Not applicable (systematic review of published studies).

Consent for Publication

Not applicable.

Availability of Data and Materials

Data extraction forms and synthesis matrices are available from the corresponding author upon reasonable request.

Competing Interests

The authors declare no competing interests.

Funding

This research did not receive any specific grant from funding agencies.

Authors' Contributions

Sedigh Mohammadi was responsible for conceptualization, methodology, formal analysis, investigation, writing – original draft, and writing – review & editing. Amir Jalali contributed to supervision, validation, and writing – review & editing.

Artificial Intelligence Utilization

During the preparation of this work, the author used ChatGPT (OpenAI) and DeepSeek (DeepSeek Company) to assist with language editing, refinement of academic expression, and formatting of references. After using these tools, the authors reviewed and edited the content as needed and took full responsibility for the final manuscript.

Data Availability Statement

Supporting data can be requested from the corresponding author.

References

1.Tanner CA. Thinking like a nurse: a research-based model of clinical judgment in nursing. Journal of Nursing Education. 2023;62(2):63-74. https://doi.org/10.3928/01484834-20230101-01
2.World Health Organization. Patient safety [Internet]. 2023 [cited 2024 Jun 1]. Available from: https://www.who.int/health-topics/patient-safety
3. Kohn LT, Corrigan JM, Donaldson MS, editors. To err is human: building a safer health system. Washington (DC): National Academies Press; 2000.
4. Benner P, Sutphen M, Leonard V, Day L. Educating nurses: a call for radical transformation. San Francisco: Jossey-Bass; 2010.
5. Jeffries PR. Simulation in nursing education: from conceptualization to evaluation. 4th ed. Philadelphia: Lippincott Williams & Wilkins; 2022.
6. Wartman SA, Combs CD. Reimagining medical education in the age of artificial intelligence. Academic Medicine. 2023;98(1):66-71. https://doi.org/10.1097/ACM.0000000000004533
7. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthcare Journal. 2021;8(2):e188-e194. https://doi.org/10.7861/fhj.2021-0095

8. Foronda CL, Fernandez-Burgos M, Nadeau C, Kelley CN, Henry MN. Virtual simulation in nursing education: a systematic review spanning 1996 to 2018. Simulation in Healthcare. 2020;15(1):46-54. https://doi.org/10.1097/SIH.0000000000000411
9. Cook DA, Triola MM. Virtual patients: a critical literature review and proposed next steps. Medical Education. 2020;54(9):786-795. https://doi.org/10.1111/medu.14172
10. Padilha JM, Machado PP, Ribeiro A, Ramos J, Costa P. Clinical virtual simulation in nursing education: randomized controlled trial. Journal of Medical Internet Research. 2021;23(3):e25466. https://doi.org/10.2196/25466
11. Kardong-Edgren S, et al. Using virtual reality and augmented reality in nursing and medical education: a call for action. Clinical Simulation in Nursing. 2021;50:1-4. https://doi.org/10.1016/j.ecns.2020.12.002
12.Chen FQ, Leng YF, Ge JF, Wang DW, Li C, Chen B. Effectiveness of artificial intelligence in nursing education: a systematic review and meta-analysis. Nurse Education Today. 2024;135:106118. https://doi.org/10.1016/j.nedt.2024.106118
13.Han ER, Yeo S, Kim MJ, Lee YH, Park KH, Roh H. The use of artificial intelligence in medical education: a systematic review and meta-analysis. PLOS ONE. 2023;18(5):e0285982. https://doi.org/10.1371/journal.pone.0285982
14. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71
15. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898. https://doi.org/10.1136/bmj.l4898
16. Shorey S, Ng ED. The use of virtual reality simulation among nursing students and registered nurses: a systematic review. Nurse Education Today. 2021;98:104662. https://doi.org/10.1016/j.nedt.2020.104662
17. Kolb DA. Experiential learning: experience as the source of learning and development. 2nd ed. Upper Saddle River: FT Press; 2014.
18.Lapkin S, Levett-Jones T, Bellchambers H, Fernandez R. Effectiveness of patient simulation manikins in teaching clinical reasoning skills to undergraduate nursing students: a systematic review. Clinical Simulation in Nursing. 2021;46:1-14. https://doi.org/10.1016/j.ecns.2020.08.010
19. Hattie J, Timperley H. The power of feedback. Review of Educational Research. 2007;77(1):81-112. https://doi.org/10.3102/003465430298487
20. Sweller J, van Merriënboer JJG, Paas F. Cognitive architecture and instructional design: 20 years later. Educational Psychology Review. 2019;31(2):261-292. https://doi.org/10.1007/s10648-019-09465-5
21. Reason J. Managing the risks of organizational accidents. London: Routledge; 2016.
22. O'Neill T, Lopes S. Error management training: a review and meta-analysis. Journal of Applied Psychology. 2022;107(12):2157-2184. https://doi.org/10.1037/apl0001021

Appendix C. Summary of AI Intervention Types and Their Primary Outcomes across Included Studies

AI Intervention Type	Number of Studies	Primary Outcome (Clinical Decision-Making)	Primary Outcome (Medical Error Reduction)
Virtual Patient Simulator	9	Significant improvement in all studies (100%)	Reduction reported in 5 out of 6 studies measuring errors
Adaptive Learning Platform	4	Significant improvement in all studies (100%)	Reduction reported in 2 out of 3 studies measuring errors
VR with Intelligent Avatar	2	Significant improvement in all studies (100%)	Reduction reported in both studies measuring errors
AI Chatbot	1	Significant improvement	Not measured

Appendix A: Sample PubMed Search Strategy

Search Block	Search Terms	Boolean Operator
AI Concepts	("Artificial Intelligence"[Mesh] OR "Machine Learning"[Mesh] OR "Deep Learning"[Mesh] OR AI OR "intelligent tutoring system" OR "virtual patient" OR "adaptive learning" OR "chatbot")	AND
Population	("Education, Nursing"[Mesh] OR "Nursing Education Research"[Mesh] OR "Students, Nursing"[Mesh] OR "nursing student" OR "nursing education")	AND
Outcome 1	("Clinical Decision-Making"[Mesh] OR "Decision Making" OR "clinical reasoning" OR "critical thinking")	AND
Outcome 2	("Medical Errors"[Mesh] OR "Medication Errors"[Mesh] OR "Patient Safety"[Mesh] OR "Safety Management" OR "medication error")	AND
Date Filter	(2019/01/01:2024/05/31[dp])	-

Note: All terms within each block are connected with the OR operator. The five blocks are sequentially connected with the AND operator. This strategy was adapted for other databases as appropriate.

Appendix B: Risk of Bias Assessment Summary

Table B1. Risk of Bias Assessment for Randomized Controlled Trials using the Cochrane (RoB 2 Tool)

Study ID	**D1**	**D2**	**D3**	**D4**	**D5**	**Overall**
Smith et al. (2023)	□ Low	□ Low	□ Low	□ Some concerns	□ Low	□ Some concerns
Kim (2023)	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Johnson et al. (2024)	□ Some concerns	□ Low	□ Low	□ Low	□ Low	□ Some concerns
Chen et al. (2022)	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Taylor et al. (2023)	□ Low	□ Some concerns	□ Low	□ Low	□ Low	□ Some concerns
Müller et al. (2024)	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Li et al. (2023)	□ Low	□ Low	□ Low	□ Some concerns	□ Low	□ Some concerns
Wong et al. (2023)	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Anderson et al. (2024)	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Yang et al. (2023)	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low

Domains (D):
D1: Randomization process
D2: Deviations from intended interventions
D3: Missing outcome data
D4: Measurement of the outcome
D5: Selection of the reported result
□ Low risk of bias
□ Some concerns
● High risk of bias

Table B2. Risk of Bias Assessment for Quasi-Experimental Studies using the ROBINS-I Tool

Study ID	D1	D2	D3	D4	D5	D6	D7	Overall
Wang & Li (2022)	□ Moderate	□ Low	□ Low	□ Low	□ Low	□ Moderate	□ Low	□ Moderate
Silva et al. (2023)	□ Moderate	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low	□ Moderate
Park et al. (2022)	● Serious	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low	● Serious
Davis et al. (2022)	□ Moderate	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low	□ Moderate
Garcia et al. (2022)	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low	□ Low
Thompson et al. (2023)	● Serious	□ Moderate	□ Low	□ Low	□ Low	□ Low	□ Low	● Serious

Domains (D):
D1: Bias due to confounding
D2: Bias in selection of participants
D3: Bias in classification of interventions
D4: Bias due to deviations from intended interventions
D5: Bias due to missing data
D6: Bias in measurement of outcomes
D7: Bias in selection of the reported result
□ Low risk of bias
□ Moderate risk of bias
● Serious risk of bias
● Critical risk of bias