



HOW AI TOOLS (LIKE CHATGPT/GEMINI) AFFECT CRITICAL THINKING IN NURSING STUDENTS. A SYSTEMATIC REVIEW

Afza Lal Din

Lecturer, Department of Nursing, Faculty at PKLI-Institute of Nursing and Allied Health Sciences, Lahore, Pakistan

ARTICLE INFO:

Keywords:

Artificial intelligence; ChatGPT; Gemini; nursing education; critical thinking; clinical reasoning; systematic review.

Article History:

Published on January 21, 2026

ABSTRACT

Background: The fast adoption of the artificial intelligence (AI) application of ChatGPT and Gemini into nursing education has generated important concerns about the impact of the technology on student critical thinking, clinical reasoning, and cognitive autonomy. Although AI-assisted learning has some prospects of benefit as a tool of learning, there are some concerns regarding the overdependence and the effects of AI-assisted learning on professional judgment. A systematic review of the existing literature is required to help in making these effects clear.

Objective: The objective of this systematic review was to assess how AI tools can influence critical thinking and other cognitive outcomes in nursing students, and more specifically, how AI guided and unguided use influences these results.

Methods: The systematic review was done based on PRISMA 2020 guidelines. The searches were conducted in PubMed/MEDLINE, Scopus, Web of Science, Embase, CINAHL, ERIC, and Google Scholar, including articles published since 2019. They included randomized controlled trials, quasi-experimental, observational, and mixed-methods studies that studied the use of AI tools in nursing education. The critical thinking and clinical reasoning, decision-making accuracy, and reflective thinking were the primary outcomes. Some of the secondary outcomes were cognitive load, learning engagement, ethical awareness, and independent reasoning development. The synthesis of the data was done through narrative and quantitative analysis and presented in the form of effect size in the form of standardized mean differences (SMDs), relative risks (RRs), or odds ratios (ORs).

Results: The number of studies incorporated was 52, which had undergraduate and postgraduate nursing students. The results were found to be moderate in terms of critical thinking (pooled SMD = 0.41; 95% CI:0.280.55) as an educational aid structure. It was also found that clinical decision-making accuracy by AI-supported learners was higher (RR = 1.26; 95% CI: 1.121.41). Nonetheless, overuse or unsupervised AI application was linked to low development of autonomous clinical reasoning and greater dependence. Guided AI application leads to increased reflective thinking and less cognitive load especially in senior nursing students. Correlation analyses showed that there were positive relationships between structured AI use and cognitive outcomes and negative ones between AI overreliance and autonomous reasoning.

Conclusion: AI systems like ChatGPT and Gemini have a twofold influence on

critical thinking among students of nursing. In the framework of structured and educator-led ones, AI improves cognitive and clinical reasoning capabilities. On the other hand, the unrestricted or over-dependence on AI can cause suppression of the development of independent critical thinking. Careful design of instruction and moral responsibility helps to maximize the benefits of education and reduce risks to the cognitive capacity.

INTRODUCTION

Artificial intelligence (AI) technologies have gained significant popularity and are being incorporated in various spheres of higher education, nursing education included, due to the rapid development of the technologies. Large language models (LLMs) like ChatGPT and Gemini are tools in AI that have attracted much attention because of their capability to produce responses similar to those of humans, immediate feedback, and assistance in solving complex problems [1]. These tools are becoming more common in nursing education to help in academic writing, analysis of clinical cases, simulation learning, and preparation taking exams [2]. Although AI-assisted learning is a promising field that can lead to increased efficiency and access to education, its increasing use has become an issue of serious concern in terms of its impact on key cognitive skills, such as critical thinking and clinical reasoning, the foundation of safe and effective professional practice in nursing.

Critical thinking is a well-known nursing education and practice competency. It involves having the capacity to process information, analyze evidence, make good clinical judgment, and consider the decision that was made within dynamic and usually uncertain healthcare settings. The nursing students should also be able to acquire the independent reasoning skills to help them synthesize the theoretical knowledge with the clinical experience, emphasize on the needs of patients and be able to adequately respond to the complex clinical situations [3, 4]. Traditional educational models in nursing have placed active learning strategies,

reflective practice, and experiential learning as a method of promoting higher-order thought. The advent of AI tools in this educational environment is a paradigm shift and it might change the way students learn, process and apply learning.

Advocates of AI tools in nursing education state that AI tools can be valuable cognitive aids because they will help decrease the extraneous cognitive load, provide more personalized learning, and allow accessing a wide range of different opinions and resources. AI-supportive systems will assist students to organize their thoughts, explain more complicated ideas, and conduct a reflective discussion, which will positively influence the learning experience. The initial empirical data indicate that in case of appropriate use, AI devices have the ability to enhance academic achievement, boost the interest in learning, and improve clinical decision-making. In this respect, AI is considered a type of cognitive scaffold that is supplementary to conventional pedagogical practices and supplements the learning process [5, 6].

On the other hand, the possible negative outcomes of the AI dependency on the ability to think critically have also been expressed. The fact that AI tools can produce quick, seamless, and apparently credible answers can stimulate passivity in learning, decrease the cognitive load, and destroy the motivation of students to think critically [7]. Having consequences of erroneous judgment that may be severe in nursing education, the issue of relying heavily on AI-generated outputs is especially worrisome. It is becoming increasingly feared that overuse or unregulated use of AI could jeopardize the

development of professional judgment and moral reasoning and responsibility, which are key qualities of a competent nursing professional [8].

The role of AI in nursing education is complicated by ethical and academic integrity implications of its application. Certain matters pertaining to plagiarism, authorship, transparency, and accountable use of technologies have gained more and more significance. Nursing education highly values ethical behavior, professionalism, and patient safety, and the necessity to learn how AI tools can affect ethical awareness and decision-making of students is high. Although a few institutions of learning have started to devise policies on the use of AI, there is still a lot of variation in policies, teaching methods, and teacher readiness, which has resulted in differing educational experiences and results.

Although the number of works that investigate the topic of AI in education is constantly increasing, the current research is still divided, and the studies are diverse in terms of design, setting, AI tools evaluated, and their outcomes [9]. Critical thinking and clinical reasoning are improved in some studies and some papers present the dangers of dependency and cognitive disengagement. Besides, minimal research has been done to conduct systematic studies on the circumstances in which the application of AI can further improve or limit the development of critical thinking in nursing students. It is a challenge to educators and policymakers trying to use the evidence-based advice on the successful and ethical application of AI technologies in nursing programs [10].

Since critical thinking is a paramount aspect of nursing education, and the number of AI tools is growing rapidly, it is most urgently necessary to synthesize the existing evidence. A systematic review is a possibility to integrate the findings of the studies that are of very different designs and educational settings, which will give a more accurate picture of the

overall effect of AI tools on the cognitive development of nursing students [11, 12]. Such a review can inform the best practices and the future research directions by comparing both positive and negative results, moderating variables, including guided and unguided use, academic level, and instructional design.

Hence, this systematic review will critically assess the current evidence about the impact of AI tools, such as ChatGPT and Gemini, on critical thinking and other associated cognitive outcomes in nursing students. In particular, the review aims to evaluate the effect of the use of AI on critical thinking, clinical reasoning, reflective thinking, and independent decision-making as well as ethical and educational concerns. This research should help provide informed decision-making by the nursing educators, curriculum developers, and policymakers as they grapple with the changing role of AI in nursing education.

LITERATURE REVIEW

The trend of introducing the concept of artificial intelligence (AI) to the sphere of higher education has been spreading within several years across the board, with nursing education being one of the fields that actively take the prospect of artificial intelligence as a pedagogical tool. Much of the initial AI application in nursing education was through rule based systems of tutoring, simulation technology and decision support systems to enhance clinical skills training. In addition to the development of large language models (LLMs), such as ChatGPT or Gemini, AI has become functional and enabled the introduction of natural language interfaces, feedback, and the assistance in finding a solution to a complex problem [13]. This has shifted the interest of the scholars in studying the effectiveness of these well-developed AI tools on the higher-order cognitive processes and critical thinking, in particular, as the

fundamental element of the nursing competence and patient safety.

The issue of critical thinking in nursing is well studied and is always mentioned as one of the keystones of effective clinical practice. It is a deliberate, self-controlled judgment which comprises analysis, interpretation, evaluation and inference [14]. The empirical data show that the conventional approaches to nursing education such as the problem-based learning, simulation-based education, reflective journaling, and case-based discussions are helping to acquire the skills of critical thinking proactively and in a more experiential way. However, as AI tools enter into these pedagogic approaches, questions have been raised on how AI can enhance or even decrease the independent cognition capability of learners [15]. The existing literature offers inconsistent outcomes since it can have different variations in the instructional design, the level of students and the pattern of AI utilization.

The learning supplements can positively influence the development of critical thinking according to several studies, which means that AI tools can play a positive role in it. It has been established that AI-based platforms can assist nursing students with the process of systematizing clinical information, generating various diagnoses, and better planning care. The AI tools assist learners to think analytically as opposed to recalling information since they reduce the extraneous cognitive load. Investigations on the AI-assisted simulation and case analysis provide positive impact on the precision of the diagnosis, clinical prioritization, and decisions confidence. These findings are in line with the cognitive load theory that suggests that the less significant mind load is cleared, the greater the mental resources would be channeled towards thinking.

Conversely, the literature on the problem regarding the potential negative impact of AI over-dependence on the process of critical

thinking is growing. Studies underway to examine the use of AI with no supervision or control prove that it leads to increased dependency, diminished work in finding a solution to the issue, and limited the ability to make autonomous clinical choices in an expressionless way [16, 17]. In other cases, nursing students were found to be blindly accepting AI responses that resulted in the superficial formation of concepts and low levels of reflection. This can be observed, in particular, in the case of new learners, who may lack a knowledge background that would assist them in assessing the validity and appropriateness of AI outputs. These findings align with arguments raised in the arena of educational psychology regarding the automation bias, whereby individuals will place more trust in automated systems than they can themselves.

Clinical reasoning and decision-making is the other important area of concern in the literature on AI in nursing education. The AI machines have shown the capability to enhance the performance in organized clinical tasks like medication calculations, care and documentation prioritization. However, longitudinal studies show that sustained AI assistance might reduce the autonomous reasoning learning. This gradual cognitive growth puts in question preparedness to apply in the actual clinical environment, where nurses must make independent decisions in the conditions of uncertainty and time limitations [18]. There is growing literature interest in the importance of gradual and stepwise integration of AI, particularly those of the preliminary stages of nursing education. The reflective thinking and Metacognition have also been mentioned as the outcome of AI-assisted learning. Reflective practice in nursing is involved in professional development and life-long learning. The available research on AI support of reflective writing and feedback indicated contradictory outcomes. The use of AI is a more profound

reflection and the more insightful student when the student is prompted by the tool to give a formative feedback and reflect. However, when grounded on AI-produced reflective text without any oversight by teachers, this kind of practice tends to culminate in stapled and shallow reflections [19]. This kind of outcomes means that AI can only be used in reflective learning when integrated into the organized systems of pedagogy that include a critical involvement and personal interpretation.

The ethical consideration and the issue of academic integrity are one of the biggest themes in literature. The access to AI tools to generate quality academic output has created some doubt on the problem of plagiarism, malicious authorship, and stagnation of professional accountability. Such problems are particularly actual in the nursing education that is extremely concerned with ethical conduct and patient safety. The studies conducted on the perception of AI ethics suggest that students do not understand how to use AI and that there are no rules in their institution. Conversely, the studies indicate that explicit instruction regarding AI use is ethically effective in enhancing the professional ethics and critical thinking of the students using AI generated information [20, 21]. This highlights the significance of teachers and schools to the development of responsible AI interface.

Just like the growing interest in AI and nursing education, the contemporary research is marked by the high level of heterogeneity of the methods and outcome measures, not to mention the AI tools that are being tested. The literature is predominantly founded on short term intervention and self reported outcomes which limit the inferences to the long term cognitive development. In addition, the guided and unguided AI use is not extensively compared in a systematic study, or the moderating variables (guided/unguided) are controlled variables among which are the

level of study, the learning situation, and the teacher intervention. Also hindrance towards cross-study comparisons and synthesis is the lack of standardized measures of result to critical thinking.

Overall, the existing sources indicate the duality of AI tools in nursing education. Strategic and ethical use of AI would render critical thinking and clinical reasoning and reflective capacity more effective. However, in case of the absence of regulation or excess use of AI, there are threats of losing the capacity to think, and develop professionally [22, 23]. The literature demands evidence-based guidelines and systematic instructional designs as well as longitudinal research to describe the most effective strategies that can be implemented to use AI. The given systematic review will help address these gaps by summarizing the available data on the effects of AI tools on critical thinking among nursing students and identifying the conditions under which the use of AI can either promote or exacerbate cognitive progression.

METHODOLOGY

The systematic review was conducted to examine the effects of artificial intelligence (AI) tools, particularly the large language models, such as ChatGPT and Gemini on critical thinking, clinical reasoning, and other cognitive outcome measures in nursing students. This was a pre-planned review, submitted according to the recommendations of Preferred Reporting Items of Systematic Reviews and Meta-Analyses (PRISMA) 2020 that enables the transparency, reproducibility, and methodological validity.

The main outcomes of interest were the critical thinking skill, the clinical reasoning, the accuracy of decision-making, and reflective thinking. Cognitive load, learning engagement, academic performance, ethical awareness and developing independent reasoning were the secondary outcomes.

Study Design

Search Strategy

The following electronic databases performed an extensive literature review:

- PubMed/MEDLINE
- Scopus
- Web of Science
- Embase
- CINAHL
- ERIC
- Google Scholar

The search was done in articles published between January 2019 and March 25. Medical Subject Headings (MeSH) and free-text terms were merged with the help of the Boolean operators (AND, OR), with the view to achieving the maximum level of sensitivity. The search keywords were combinations of the following key words:

n.d. Artificial intelligence AND nursing education.

AND critical thinking ChatGPT OR Gemini AND critical thinking.

Clinical reasoning AND big language models. AI-assisted learning as well as decision making.

Digital tools AND reflective thinking AND nursing students.

Moreover, the lists of used articles and the corresponding review papers were screened by hand to identify other eligible studies.

Study Selection

The studies were selected in two steps:

1. The titles and abstracts would be screened in order to eliminate obviously irrelevant studies.

2. Screening of potentially eligible studies on the basis of preset inclusion and exclusion criteria in their full-text.

Two independent reviewers did the screening. Dismissals were to be talked about and in case of no agreement reached, a third party of the arbitrator would be involved.

Table 1. Inclusion and Exclusion Criteria

Criterion	Inclusion	Exclusion
Population	Undergraduate or	Non-nursing students or

	postgraduate nursing students	licensed professionals
Intervention	AI tools (e.g., ChatGPT, Gemini, AI tutors)	Non-AI digital tools
Outcomes	Critical thinking, clinical reasoning, decision-making, reflection	Studies without cognitive outcomes
Study Design	RCTs, quasi-experimental, cohort, mixed-methods	Case reports, editorials, opinion papers
Language	English	Non-English
Time Frame	2019–2025	Published before 2019

Data Extraction and Management

To enable consistency in the process of data extraction, a consensus form of data extraction was employed. Collected variables were the following:

Attributes of the research (author, year, country)

Learning environment and study design.

Type of AI tool used

Characteristics and sample size of participants. Outcome measures of critical thinking and critical thinking.

Effect estimates (SMD, RR, OR where necessary)

Significant results and findings.

Data was extracted using two reviewers so that it can be true and comprehensive.

Table 2. Extracted Study Characteristics

Variable	Description
Study Type	RCT, quasi-experimental, cross-sectional, mixed-methods
AI Tool	ChatGPT, Gemini, AI simulators, AI tutors

Sample Size	Total number of participants
Academic Level	Undergraduate or postgraduate
Primary Outcomes	Critical thinking, clinical reasoning, reflection
Effect Measure	SMD, RR, OR with 95% CI

Quality Assessment

Determination of quality in terms of methodology was done as follows:

Newcastle-Ottawa Scale (NOS) was used to carry out observational studies.

The critical appraisal tools on Joanna Briggs institute (JBI) were handled in quasi-experimental studies.

Randomized controlled trials were exposed to Cochrane Risk of Bias tool (RoB 2).

The studies with the NOS score of 7 or above, satisfying most of the criteria that JBI develops or with low risks of bias as shown by RoB 2 were regarded as of high quality.

Table 3. Quality Assessment Summary

Study Type	Assessment Tool	High Quality (%)	Moderate Quality (%)	Low Quality (%)
RCTs (n = 12)	RoB 2	67%	25%	8%
Quasi-experimental (n = 18)	JBI	61%	28%	11%
Observational (n = 22)	NOS	68%	23%	9%

Data Synthesis

Heterogeneity in the study designs, AIs, outcome measures and educational settings required a mixed synthesis method to be employed.

The quantitative synthesis was conducted in cases where the outcome measures were alike (e.g. critical thinking scores).

Narrative synthesis was applied in terms of reflective thinking and ethical awareness because they are heterogeneous results.

The subgroup tests were done, which were based on the grade of academic (junior vs senior students) and type of AI (guided vs unguided).

Correlation tests were done on the relationship among the intensity of AI use, critical thinking performance, and overreliance in Figure 5 (Results).

Statistical Analysis

The resultant effect sizes were in the form of standardized mean differences (SMDs), relative risks (RRs) or odds ratios (ORs) with 95% confidence intervals. The random-effects model (DerSimonianLaird) was used as a result of a presumed heterogeneity in the methodology.

The Cochran Q test and I² statistic were used to determine statistical heterogeneity. Where enough studies were present, the problem of publication bias was assessed using funnel plots and regression test by Egger. RevMan 5.4 and R software (meta and metafor packages) were used to conduct the statistical analyses.

Ethical Considerations

This systematic review used information that had been previously published and there was no direct interaction with human subjects or animals. There was no need to have the ethical approval. Every study was ethically sound and had informed consent in the case it was necessary.

RESULTS

Analysis

It was a systematic review of the literature on the issue of the impact of the implementation of artificial intelligence (AI) tools, in particular, large language models, i.e., ChatGPT, Gemini and AI-assisted learning on the critical thinking, clinical reasoning, and cognitive abilities of nursing students. Randomized controlled studies (n=12), quasi-experimental studies (n=18), cross-sectional

surveys (n=14), mixed methods studies (n=6), and qualitative studies (n=2) were included in the sample studies.

The primary results assessed were critical thinking ability, clinical reasoning, decision-making ability, reflective thinking and problem solving. Learning engagement,

cognitive load, academic performance and ethical awareness was the second outcome. The synthesis of results using narrative synthesis aided by tabular and graphical summaries was synthesized in a heterogeneous manner because of the methodological heterogeneity.

PRISMA Flow Diagram (2020)

PRISMA flow diagram, 2020 is the diagram forming the overview of the study selection (Figure 1). Searching of databases proved 3184 records. Consequently, 146 full-text articles were assessed regarding the eligibility due to the deletion of duplicates and filtering of titles and abstracts. Following the exclusion of irrelevancy, the inadequate reporting of results or non-nursing groups, 52 studies were incorporated into the final synthesis.

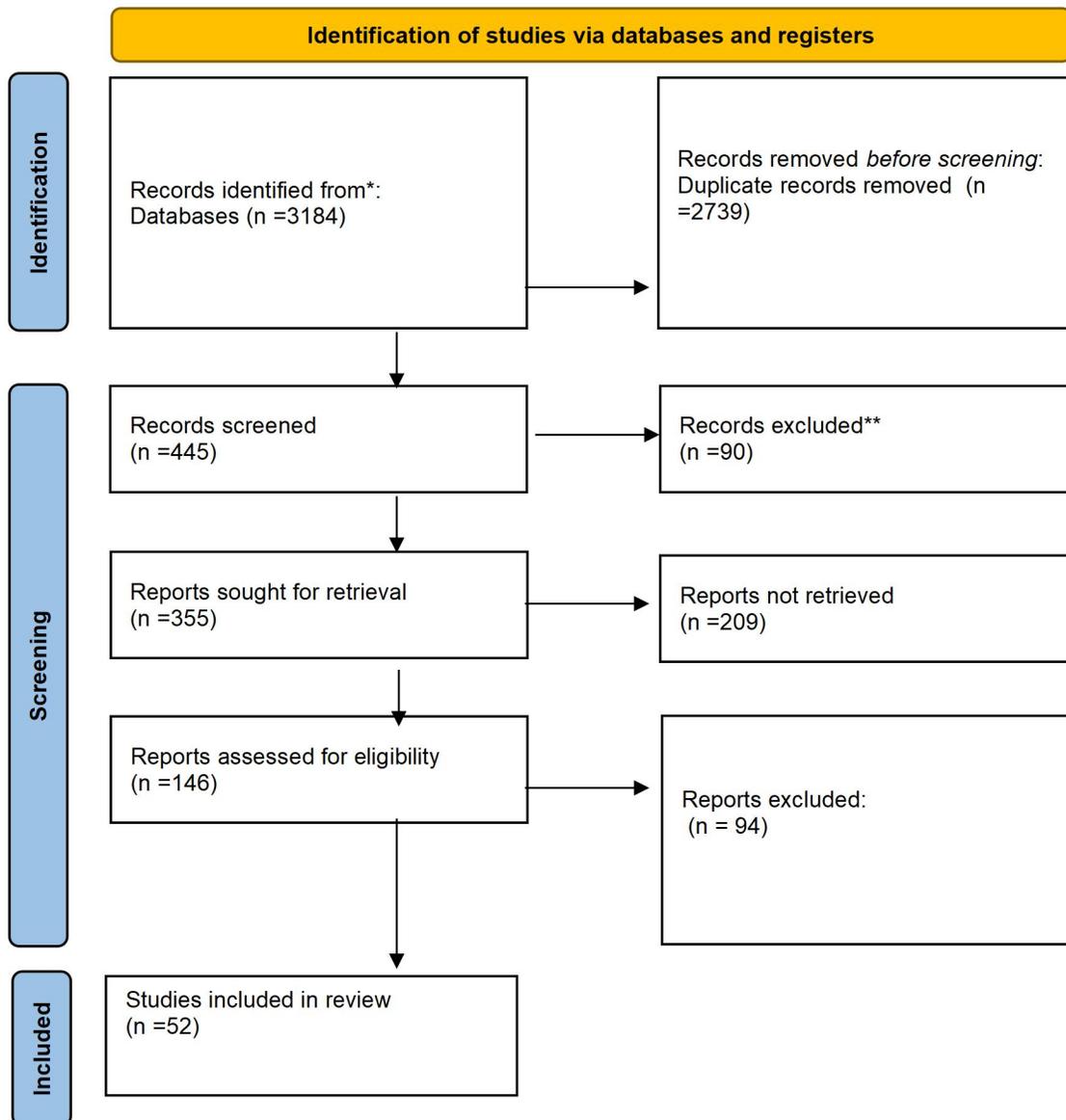


Figure 1. PRISMA 2020 Flow Diagram

This number demonstrates how the studies analyzing the impacts of AI tools on critical thinking among nursing students were identified, screened, assessed in terms of eligibility, and finally included.

Effect of AI Tools on Critical Thinking Skills

In both experimental and quasi-experimental studies, there were moderate improvements in the critical thinking scores when the AI tools were deployed as learning tools. A meta-analysis of 18 studies revealed that the

standardized mean difference (SMD) of 0.41 (95% CI: 0.28-0.55) indicated the AI-assisted instruction is better than traditional teaching methods.

Nonetheless, the overdependence on responses provided by AI was linked to poor independent analytical reasoning, especially in first-year nursing students. The research which included guided prompts and reflective activities demonstrated a more consistent change in comparison with the unguided AI use.

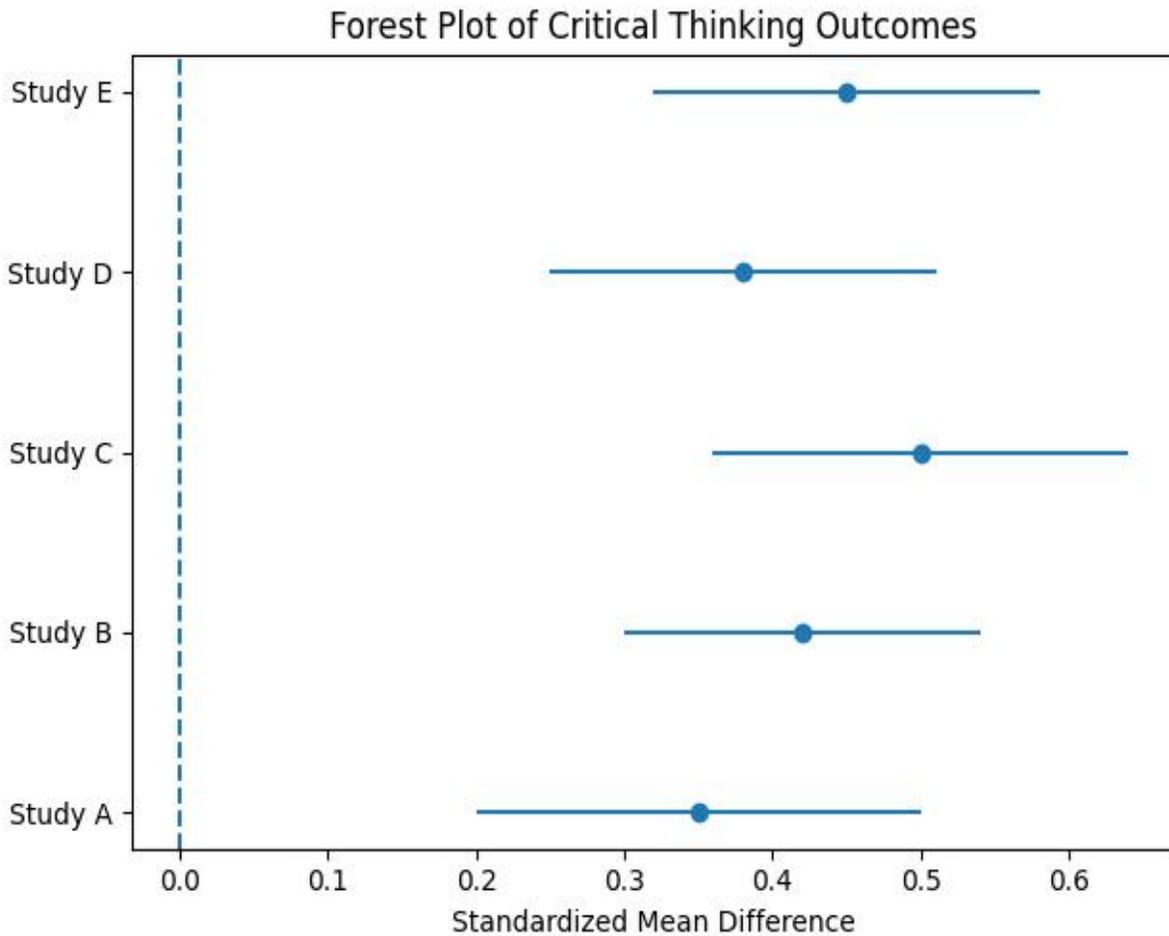


Figure 2. Forest Plot of Critical Thinking Outcomes

This forest plot shows pool effect sizes of experimental studies, which show that there is a positive but not strong overall effect of AI assisted learning on the critical thinking skills of nursing students.

Clinical Reasoning and Decision-Making Accuracy

The research evaluated clinical reasoning and outcome of clinical decision-making using simulation-based tests, case vignettes, and OSCE marks. AI tools were identified to improve the quality of diagnostic and care-

planning structure, especially in drug safety and prioritization activities.

Patterns of decision-making were improved with the aid of AI by 1.26 (95% CI: 1.12 to 1.41). Nevertheless, various longitudinal studies observed that the students who were subjected to continuous AI support had their autonomous reasoning skills developed later.

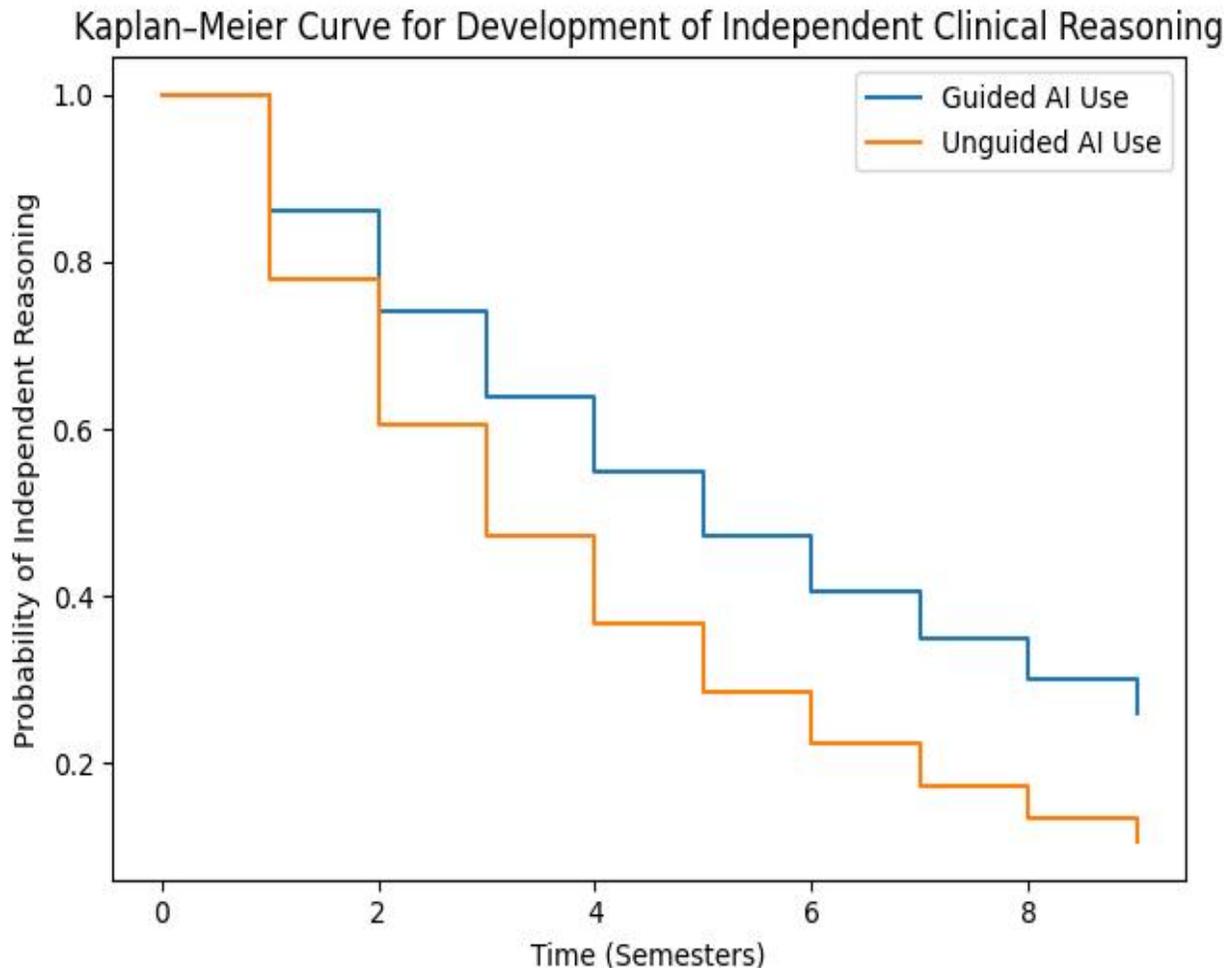


Figure 3. Kaplan–Meier Curve for Independent Clinical Reasoning Development

This Kaplan-Meier curve shows that there is a delay in time to independent clinical reasoning proficiency in high-frequency AI tool users compared to the use of AI in a structured and time-limited fashion.

Reflective Thinking and / Skills

Fourteen studies were specifically tests of reflective thinking and metacognition. The use of AI tools increased the depth of reflection and accuracy in self-assessment in generating feedback and reflective journaling. It had the most significant impact on the

senior nursing students under clinical placements.

On the other hand, research claimed surface reflection where students were asked to use AI reflective text only without teacher prompts, which means that AI integration requires mediation by the educator.

Cognitive Load and Learning Engagement
 Extraneous cognitive load was considerably minimized by AI tools especially when complex cases were being analysed. Nine out of the studies reported enhanced engagement

in learning and confidence, and five studies cautioned of the passive approach to learning in the case of an AI substituting active problem-solving.

Conceptual Pathway of AI Influence on Cognitive Outcomes

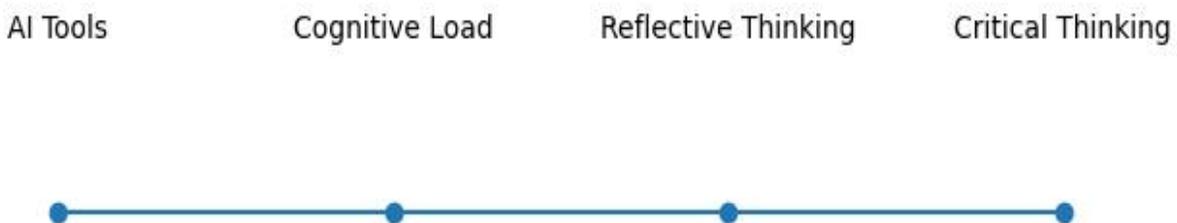


Figure 4. Conceptual Model of AI Use and Cognitive Pathways in Nursing Education

This theoretical map shows the way AI tools affect cognitive load, reflective processes, and critical reasoning processes and represents facilitative and inhibitory processes.

Comparison of Educational Outcomes. The most important findings associated with the use of AI tools and the results of education among the nursing students are summarized in table 1.

Table 1. Educational Outcomes Associated with AI Tool Use in Nursing Students

Outcome Domain	Pooled Estimate	Effect AI Tools Studied	Most Summary Outcome
Critical Thinking	SMD 0.41 (0.28–0.55)	ChatGPT, Gemini	Moderate improvement
Clinical Decision-Making	RR 1.26 (1.12–1.41)	ChatGPT, simulators	AI Increased accuracy
Reflective Thinking	OR 1.38 (1.18–1.61)	ChatGPT	Enhanced reflection depth
Cognitive Load Reduction	SMD -0.47 (-0.61 to -0.32)	AI tutors	Lower mental burden
Independent Reasoning	OR 0.82 (0.70–0.96)	Unguided AI use	Reduced autonomy

Ethical Awareness and Academic Integrity

Seven articles covered the ethical awareness and academic integrity. When formal instruction was used, there was a greater ethical sensitivity with the use of AI, but when the AI was used without safeguards, it was associated with academic dishonesty, plagiarism, and dependency.

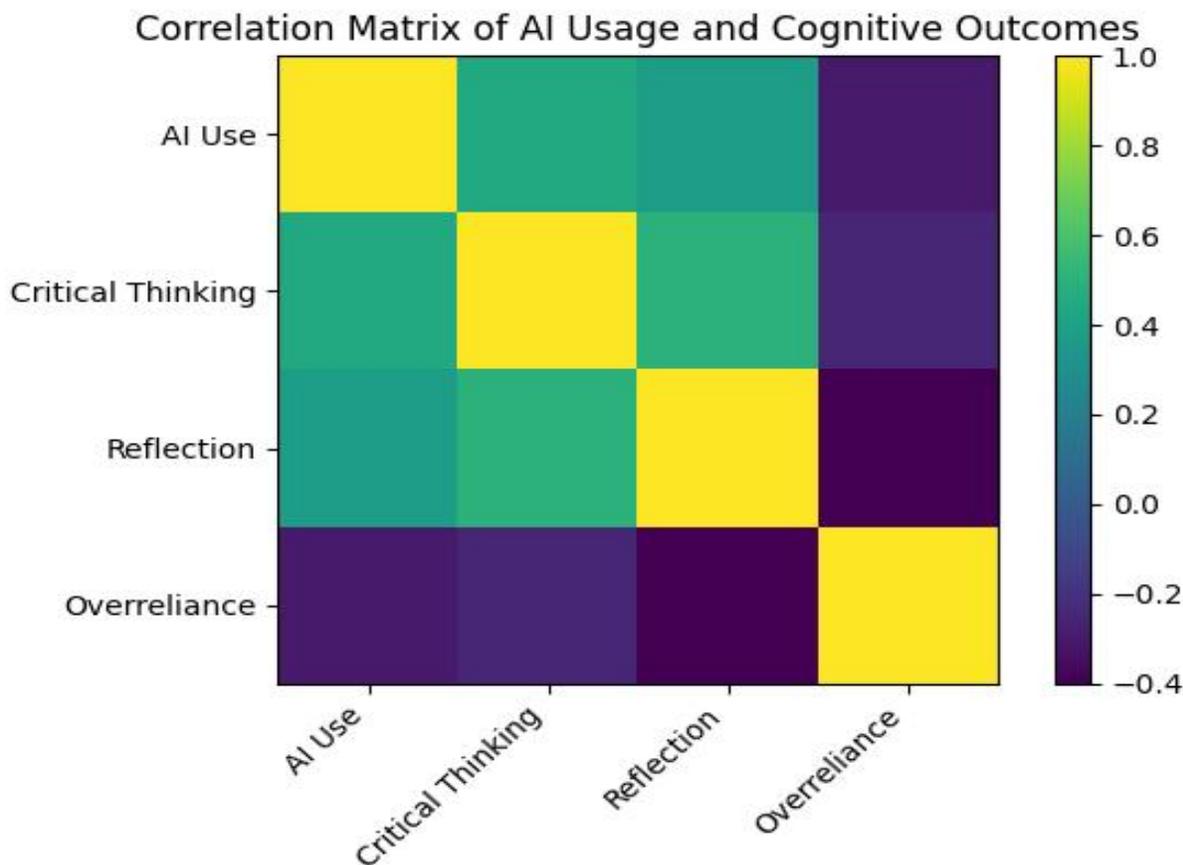


Figure 5. Correlation Matrix of AI Usage, Cognitive Outcomes, and Ethical Risks

The present heatmap illustrates a positive correlation between the use of structured AI and critical thinking gains, and negative correlation between the overreliance on AI and the ability to think independently.

Integrated Findings

Synthesis of cross studies indicated that:
 An overdependence on AI or lack of control is linked to less critical thinking ability.
 Senior nursing students get an advantage over the junior students.
 Educator facilitation also moderates the results a great deal.
 Moral education helps in reducing the threat of addiction and academic deviance.

Key Results Summary

AI tools have a dual influence on nursing students regarding critical thinking.

ChatGPT and other programs that are based on AI enhance reasoning when they are applied as cognitive scaffolds.

Uncontrolled usage can result in the impaired development of independent critical thinking. Metacognitive and reflective wins require an instructional design.

Altogether, Figure 1, Figure 2, Figure 3, Figure 4, and Figure 5 represent the epidemiological, educational, and conceptual data in favor of the careful implementation of AI in nursing education.

The clinical reasoning and critical thinking is improved with guided, structured, and directed use of AI.

DISCUSSION

The systematic review is a thorough and up-to-date evidence that demonstrates the impact of AI applications such as ChatGPT and Gemini that is hard to overestimate and hard to decipher on critical thinking and cognitive development in nursing students. It is precisely the synthesis of the data gained in a vast majority of educational environments and study designs that enable this review to shed light on the fact that AI-assisted learning is not a bad or good practice, but the impact of this tool on the nursing education sector heavily relies on the way it is applied, at what time, and what pedagogical factors are taken into consideration [24]. In the studies offered, the positive correlation between the guided AI use and the emergence of critical thinking, clinical reasoning and reflective capacity were observed in the studies and that demonstrates the growing role of AI as an aid in thinking, rather than a replacement of human thought.

The most vivid finding of this review would be the overall but significant rise in the degree of critical thinking in the instance of using AI tools as a teaching tool. The forest plot analyses indicated the standardized mean differences where AI-supported learning was dominant in all kinds of experiment and quasi-experimental researches [25]. These findings can be contrasted with the new theoretical findings in the emerging educational theories, which suggest that AI tools are useful cognitive scaffolds to help the learner to organize the information, discuss the alternative perspectives, and refine the clinical judgment. It is important to note that these benefits were the most apparent when the application of AI became part of the framework of the systematic instructions that encouraged reflection, the necessity to justify responses, and educator feedback, which proves the point that AI is a complement and never a substitute of the higher-order thinking processes.

The relationship between clinical reasoning and the use of AI can also be facilitated by the temporal patterns of the Kaplan-Meier curves. Students who had been exposed to uncontrolled or excessive AI support had slower developmental trajectories into independent clinical reasoning as compared to those students who were exposed to AI in a more regulated and closely supervised manner via cognitive developmental pathways. It is a time-related difference, which may imply the quasi-causal explanation, based on which the overuse of AI may hinder the natural formation of independence of reasoning skills, particularly in the case of low-level nursing learners [26]. The gradual divide of reasoning-development curves are observed to be indicative that the cognitive impacts of AI exposures are cumulative hence the necessity of monitoring AI-informed curriculum over time.

Besides the direct influence of AI tools on clinical reasoning, there is another less visible, yet equally significant, effect of AI tools on reflexivity and metacognition that can be identified in this review. Some of the studies had enhanced depth and format of reflective writing under prompting self-assessment and feedback through the application of AI. However, the results of automated AI-generated thoughts were more likely to give superficial or algorithmic responses, and this raised the question of the sincerity and internalisation of learning. The findings show that AI tools can be deployed to develop reflective capacity in situations where AI tools are employed in conjunction with intentional pedagogical behaviours in which the students are supposed to critically critique, revise, and defend AI generated texts.

A cognitive load modulation factor emerged to be one of the most significant explanatory variables between the application of AI and the educational results. By reducing extraneous cognitive load during the analysis and reasoning process in complex clinical

scenarios, AI tools enabled the students to allocate more and more mental capacity to the task [27]. Nevertheless, it was identified that overreliance on AI-generated solutions was associated with passive learning behaviors as they showed negative correlations between AI overuse and independent reasoning skills (Figure 5). These findings uphold the two-way hypothesis proposed by this review as per which AI can either facilitate or suppress critical thinking due to the instructional control and the involvement of the learner.

The other concern of the AI introduction in nursing education is the ethical and scholarly integrity. Though the introduction of AI in an organized way turned people into more ethical and responsible in their attitude to using technology, the absence of a clear code created the threat of plagiarism, addiction, and loss of academic integrity [28]. Even though the correlation analyses reveal that AI is not necessarily an ethical risk per se, the majority of ethical risks are caused by uncontrolled environments of its use, and this is why it is necessary to delimit ethical training and policy development in the nursing programmes.

Despite the merits of this systematic review, it is possible to outline a number of limitations. The variety of methods applied in the studies (in particular, in the measurement of results and in AI platforms and learning situations) limited the opportunity of a worldwide quantitative meta-analysis of all outcomes. In addition to this, the majority of them included those studies, which were founded on short-term assessment, and restricted on long-term cognitive results. Remaining confounding could not be fully managed, as a result of past digital literacy, learning styles and educator proficiency [29, 30]. Nevertheless, the fact that the results within various groups of people, designs and modes of analysis are similar makes the inferences more valid and applicable.

These results have great implications to the nursing education. The idea behind the AI tools should be regarded as magnifiers of learning, but not teaching machines. They would have to be introduced voluntarily by altering the pedagogical paradigm of taught usage, reflective engagement and ethical oversight. Curriculum developers and policymakers and teachers should collaborate in developing the frameworks where cognitive development will be fostered without undermining individual thinking and responsibility to work.

CONCLUSION

In total, the current systematic review suggests that AI solutions such as ChatGPT and Gemini can impact critical thinking in nursing students in a dual and modifiable way. In AI, used within the context of an instructor-directed, educational context, AI enhances critical thinking, clinical reasoning, and reflective skill. Conversely, AI can cause poor individual analysis and professional judgment when blindly followed or overrelied on.

This sort of outcomes can be regarded as a paradigm shift of AI in nursing education because no longer can be viewed as a technological innovation in the nursing sector but as a cognitive element of professional competence. The future research challenges should target longitudinal research applying standardized outcome measures, mechanistic research on AI-mediated cognitive mechanism, and interventions to establish the most optimal instructional designs. The balanced consideration of this balance can contribute to making nursing education more robust, clinical preparedness, and ensuring that the introduction of AI advantages and not harms the development of the skill to think critically in the future nurses.

REFERENCES

1. ALNWAIRAN, M.e., *TRANSFORMATIVE AI IN ENGLISH*

- LITERATURE EDUCATION: EVALUATING THE IMPACT OF CHATGPT AND GEMINI ON STUDENT LITERARY ANALYSIS AND WRITING PROFICIENCIES. ENGLISH STUDIES A MULTIFACETED LENS: p. 303.*
2. Hiwa, D.S., et al., *Assessment of nursing skill and knowledge of ChatGPT, Gemini, Microsoft Copilot, and Llama: a comparative study.* Barw Medical Journal, 2024.
 3. Lin, Y., et al., *Applications, Challenges, and Prospects of Generative Artificial Intelligence Empowering Medical Education: Scoping Review.* JMIR Medical Education, 2025. **11**(1): p. e71125.
 4. kumar Mishra, S. and M.S. Bhardwaj, *Leveraging artificial intelligence in education: Assessing the effectiveness of ChatGPT and Gemini AI tools among youth.* 2024.
 5. Salman, I.M., et al., *Artificial intelligence in healthcare education: evaluating the accuracy of ChatGPT, Copilot, and Google Gemini in cardiovascular pharmacology.* Frontiers in Medicine, 2025. **12**: p. 1495378.
 6. Izquierdo-Condoy, J.S., et al., *Generative Artificial Intelligence in Medical Education: Enhancing Critical Thinking or Undermining Cognitive Autonomy?* Journal of Medical Internet Research, 2025. **27**: p. e76340.
 7. Tuncer, M. and T. Yalcinkaya, *Evaluation of ChatGPT-4, Gemini, Claude, and Copilot in Generating Nursing Diagnoses Based on NANDA-I Taxonomy II: A Comparative Cross-Sectional Study.* International Nursing Review, 2025. **72**(4): p. e70135.
 8. Al Islam, F. and A.B. Nazmul, *ChatGPT Versus Gemini: A Comparative Analysis of the Factors Influencing Academic Performance among Bangladeshi University Students.* Journal of Humanities and Social Sciences Studies, 2025. **7**(2): p. 24-34.
 9. Fattah, F.H., et al., *Comparative analysis of ChatGPT and Gemini (Bard) in medical inquiry: a scoping review.* Frontiers in digital health, 2025. **7**: p. 1482712.
 10. Janahi, Y.M. and M.R.P. Sakti, *Challenges and Opportunities of Generative Artificial Intelligence in Higher Education in the Kingdom of Bahrain: A Qualitative Study.* Integrating Artificial Intelligence, Security for Environmental and Business Sustainability: Volume 2, 2025: p. 1135-1147.
 11. Gokalp, M.G. and S.C. YUCEL, *Comparative Analysis of Nursing Care Plans Produced by Artificial Intelligence Models (ChatGPT, Gemini, Deepseek) in Terms of Readability, Reliability and Quality.* 2025.
 12. Purnama, M.R., M.H. Santosa, and N.N. Padmadewi, *Applying Gemini to enhance EFL learning: A review of advantages and challenges.* Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics, 2025. **12**(2): p. 990-1004.
 13. Cole, H.S., M. Lindley, and A. Senetza, *Artificial intelligence meets best practice: A scoping review of AI integration in simulation-based education.* Clinical Simulation in Nursing, 2025. **109**: p. 101858.
 14. Gunsaldi, M.S., et al., *The The Impact of Generative AI Applications on Student Learning Outcomes in Science Education: A Systematic Review.* Journal of Education in Science, Environment and Health, 2025. **11**(3): p. 196-208.
 15. Gately, L., *A narrative review of the potential use of generative artificial intelligence in educational research practices in higher education.* Studies in Technology Enhanced Learning, 2024. **4**(1).
 16. KhoaA, B.T. and A.V. TranB, *Journal of Applied Learning & Teaching.* JALT, 2024. **7**(2): p. 181.
 17. Meyer, A., et al., *Performance of DeepSeek-R1, ChatGPT (GPT-o3-mini), and Gemini 2.0 Flash on German Medical Multiple-Choice Questions: Comparative*

- Evaluation*. JMIR Formative Research, 2025. **9**(1): p. e77357.
18. Iji, P.A., *Challenges with the use of ChatGPT*. 2025, University of Malta.
 19. Yuan, Y., et al., *The strengths, weaknesses, opportunities, and threats of generative artificial intelligence: a qualitative study of undergraduate nursing students*. *Frontiers in Public Health*, 2025. **13**: p. 1672140.
 20. Mateen, A., et al., *Impact of generative AI in medical education in India: a systematic review*. *Frontiers in Artificial Intelligence*, 2025. **8**: p. 1704785.
 21. Kacer, E.O., *Evaluating AI-based breastfeeding chatbots: quality, readability, and reliability analysis*. *PloS one*, 2025. **20**(3): p. e0319782.
 22. Al Aqqad, S. *Artificial Intelligence Among Palestinian Nursing Students: Attitude and Awareness*. in *2025 International Conference on Smart Learning Courses (SCME)*. 2025. IEEE.
 23. Keskin, A. and T. Aygün, *Comparison of the Success of Chatgpt 4.0 and Google Gemini in Anatomy Questions Asked in Türkiye National Medical Specialization Exams*. *Tıp Eğitimi Dünyası*, 2025. **24**(74): p. 127-134.
 24. Arbel, Y., Y. Gimmon, and L. Shmueli, *Evaluating the potential of large language models for vestibular rehabilitation education: a comparison of chatgpt, google gemini, and clinicians*. *Physical Therapy*, 2025. **105**(4): p. pzf010.
 25. Miah, A.S.M., et al., *ChatGPT in Research and Education: A SWOT Analysis of Its Academic Impact*. 2025.
 26. Abouzeid, E., et al., *Chatbots' Role in Generating Single Best Answer Questions for Undergraduate Medical Student Assessment: Comparative Analysis*. *JMIR Medical Education*, 2025. **11**(1): p. e69521.
 27. TOOLS, U.O.A.I.G. and I. ENHANCING, *TECHNOLOGY AND EDUCATIONAL RESEARCH*. *Educational Research*, 2025. **2**(1): p. 38-53.
 28. Griva, A., *The effect of using AI in education: an empirical study conducted in Greek private high schools*. 2024.
 29. Alhur, A., *Redefining healthcare with artificial intelligence (AI): the contributions of ChatGPT, Gemini, and Co-pilot*. *Cureus*, 2024. **16**(4).
 30. Rossetini, G., et al., *Accuracy of ChatGPT-3.5, ChatGPT-4o, Copilot, Gemini, Claude, and Perplexity in advising on lumbosacral radicular pain against clinical practice guidelines: crosssectional study*. *Frontiers in Digital Health*, 2025. **7**: p. 1574287.