JTPED
Journal of Technological Pedagogy and Educational Development
Vol. 2, No. 2, 2025, pp. 12-28
https://ejournal.jtped.org/ojs/index.php/jtped

ChatGPT as a Pedagogical Tool for Clinical Reasoning in Medical Education: A Systematic Narrative Review

Annastasya Nabila Elsa Wulandari [a,1], Purwono Purwono [a,2,*], Alfian Ma'arif [b,3], Marco Antonio Marquez Vera [c,4], Rania Majdoubi [d,5], Wael A. Salah [e,6]

[a] Department of Informatics, Universitas Harapan Bangsa, Purwokerto, 53144, Indonesia
[b] Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[c] Universidad Politécnica de Pachuca, Mexico
[d] SETIME Laboratory, Faculty of Sciences, Ibn Tofail University, Kenitra, Morocco
[e] College of Engineering and Technology, Palestine Technical University – Kadoorie, Tulkarm, Palestine
[1] annastasyanabilae@gmail.com; [2] purwono@uhb.ac.id; [3] alfianmaarif@ee.uad.ac.id; [4] marquez@upp.edu.mx; [5] rania_majdoubi@um5.ac.ma; [6] wael.salah@ptuk.edu.ps
* Corresponding Author

ARTICLE INFO

Article history
Received July 23, 2025
Revised July 20, 2025
Accepted August 13, 2025

Keywords
Chatgpt;
Medical Education;
Pedagogical Tool;
Clinical Reasoning

ABSTRACT

The development of clinical reasoning is a critical objective in medical education, yet traditional pedagogical approaches often face challenges such as limited faculty time, inadequate feedback, and variable mentoring quality. The emergence of conversational generative AI, particularly ChatGPT, offers new opportunities for interactive, reflective, and self-directed learning.This systematic narrative review examines the pedagogical potential of ChatGPT in enhancing clinical reasoning among medical students. Literature was sourced from Scopus, PubMed, Web of Science, and Google Scholar using keywords such as "ChatGPT," "generative AI," "medical education," "clinical reasoning," and "pedagogy," focusing on studies published between 2020 to 2025. The findings indicate that ChatGPT contributes to clinical education by simulating diagnostic scenarios, supporting self-regulated learning, and providing personalized feedback. Key applications include case-based reasoning, virtual dialogue, and integration into LMS, OSCE preparation, and problem-based learning. However, limitations such as AI hallucinations, risk of overreliance, and reduced human mentorship were noted. In conclusion, ChatGPT presents a promising pedagogical tool to support clinical reasoning, but its implementation must be guided by ethical oversight and aligned with evidence-based instructional design. Future research should prioritize empirical validation and multi-institutional implementation frameworks to ensure its responsible and effective use in clinical training.

1. Introduction

The development of clinical reasoning is a central objective in undergraduate and professional medical education. Clinical reasoning refers to the cognitive process through which medical practitioners gather patient information, interpret clinical data, formulate differential diagnoses, and make evidence-based decisions [1]. Despite its importance, cultivating this skill during clinical rotations presents persistent challenges in medical pedagogy.

https://doi.org/10.59247/jtped.v2i2.21

One major difficulty arises from the implicit nature of expert reasoning. Experienced clinicians often rely on intuitive decision-making that is grounded in well-established illness scripts developed over years of practice [2]. This tacit knowledge is seldom articulated in an explicit or structured manner, which makes it difficult for novice learners to comprehend or reproduce the underlying thought processes. Furthermore, clinical environments are typically characterized by time constraints, heavy workloads, and limited opportunities for direct teaching [3]. These factors restrict the availability of guided reflection and feedback, resulting in learning that is often passive and observational rather than interactive and reflective.

In addition, many clinical curricula lack systematic frameworks for teaching reasoning explicitly. While educational strategies such as problem-based learning and case-based learning have been introduced, their implementation during clinical placements is often inconsistent. In the absence of structured mentorship, students may become overwhelmed by complex and ambiguous patient presentations [4]. Without sufficient guidance and support, their clinical reasoning skills may fail to develop or may regress over time.

Advances in educational technology have begun to address some of these gaps. In particular, the emergence of ChatGPT, a conversational generative AI model built on transformer-based architecture, has introduced new possibilities for learner support in clinical education [5]. ChatGPT enables real-time, text-based interactions that simulate human dialogue [6]. This allows students to ask clinical questions, explore alternative diagnoses, and receive instant explanations of medical concepts.

From a pedagogical perspective, ChatGPT has the potential to serve as a cognitive scaffold. It can support students in articulating their reasoning, identifying knowledge gaps, and engaging in self-directed learning [7]. The tool can also function as a reflective learning partner, encouraging deeper exploration of clinical scenarios. Its accessibility outside formal learning environments aligns well with the principles of self-regulated learning.

However, the integration of ChatGPT into clinical education remains a relatively new and evolving area. Questions regarding its accuracy, reliability, and ethical implications must be addressed before widespread adoption can occur [8]. Therefore, this review aims to examine the current literature on ChatGPT's use in medical education, with a focus on its pedagogical role in supporting the development of clinical reasoning. The review synthesizes existing findings, identifies emerging practices, and explores implications for future educational frameworks that incorporate AI-based learning tools.

This review synthesizes current literature on ChatGPT's use in medical education, with a particular focus on its pedagogical role in fostering clinical reasoning. The paper offers a novel contribution by aligning AI capabilities with established educational theories and identifying strategies for effective classroom and clinical integration.

2. Method

This article adopts a narrative review approach with a systematic orientation in the selection and synthesis of literature (Fig. 1). The review was designed to identify, organize, and interpret current evidence on the pedagogical use of ChatGPT and generative artificial intelligence in the context of clinical reasoning within medical education. This approach was chosen due to the interdisciplinary and rapidly evolving nature of the topic, which spans a range of theoretical frameworks and lacks standardized empirical methodologies.

A comprehensive search was conducted across five major academic databases: Scopus, PubMed, Web of Science, Eand Google Scholar. These databases were selected based on their broad indexing of peer-reviewed literature in health professions education, instructional design, and artificial intelligence, ensuring both relevance and comprehensiveness in the search process. The search strategy employed a combination of relevant keywords and Boolean operators, specifically: "ChatGPT" OR "Generative AI" AND "medical education" AND "clinical reasoning" AND "pedagogy." These terms were selected to capture a broad yet focused body of literature encompassing both technological and educational perspectives.

Inclusion criteria were defined to ensure the relevance and quality of the selected studies. Only peer-reviewed journal articles published between 2020 and 2025 were included, reflecting the recent and rapidly evolving nature of generative AI in education. Articles were required to explicitly address the use of ChatGPT or similar generative AI tools in medical education, particularly those that focus on clinical learning or reasoning processes. Studies that addressed AI applications in non-clinical educational domains, opinion pieces without empirical or theoretical grounding, or publications outside the specified timeframe were excluded. To minimize selection and extraction bias, screening and selection were conducted systematically based on predefined inclusion and exclusion criteria. Ambiguities or uncertainties during the process were resolved through iterative evaluation and adherence to the objectives of the review.

Data extraction was performed manually and structured according to key thematic categories. For each article, the following elements were recorded: publication year, country of study, research design or review type, the educational or theoretical framework employed (if any), the type of AI tool utilized, and the reported educational outcomes. Special attention was given to studies that incorporated pedagogical theories such as constructivism, cognitive apprenticeship, or self-regulated learning to interpret the role of AI in clinical education.

To enhance transparency in article selection, a structured narrative flow was followed. The process began with the identification of potentially relevant articles through database searches using predefined keywords. After removing duplicates, titles and abstracts were screened for relevance, followed by a full-text review based on predefined inclusion and exclusion criteria. Ambiguities or disagreements during screening were resolved through consensus. Studies that met all criteria were included in the final thematic synthesis.

The synthesis of findings aimed to identify recurrent themes, variations in implementation, and gaps in the literature that may inform future research and curriculum development. Where possible, studies were cross-compared to assess consistency and emerging best practices related to the pedagogical use of ChatGPT in medical settings.

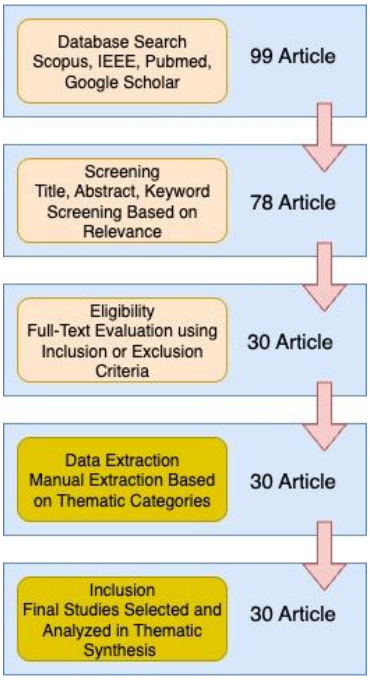

Fig. 1. Structured narrative flow of article selection for the systematic narrative review

3. Theoretical Foundations

3.1. Cognitive Apprenticeship

The cognitive apprenticeship model emphasizes the importance of making expert thinking visible to learners through structured guidance, reflection, and gradual transfer of responsibility [9]. In medical education, it is a well-established approach for cultivating complex reasoning skills such as diagnosis and decision-making [10]. It encourages learners to observe, emulate, and eventually internalize the strategies used by experienced clinicians during patient encounters and diagnostic problem-solving [11]. Traditionally, this model relies on live mentoring, where clinicians explain their thought processes and provide contextual feedback [12]. However, time constraints and inconsistent supervision often hinder its consistent application [13], leaving students without sufficient scaffolding.

The integration of artificial intelligence, particularly conversational agents such as ChatGPT, presents new opportunities to extend the principles of cognitive apprenticeship beyond direct human mentorship [14]. ChatGPT can function as a surrogate expert, engaging students in reflective dialogues that simulate the reasoning patterns of experienced clinicians [15]. When prompted with clinical scenarios, ChatGPT can verbalize step-by-step diagnostic considerations, offer multiple differential diagnoses, and justify treatment decisions based on evidence. This interaction enables learners to observe modeled reasoning in real time and engage in iterative questioning that enhances their understanding [16].

Specifically, ChatGPT supports cognitive modeling by verbalizing links between data and hypotheses, and enables scaffolding by adapting to a student's level of understanding [17]. While students may initially rely heavily on ChatGPT, they can transition to independent reasoning with its support. Nonetheless, ChatGPT lacks genuine clinical judgment and contextual sensitivity. Its probabilistic outputs are not substitutes for lived expertise. Thus, it should be positioned as a supplement to human mentorship, not a replacement [18]. Educators must guide students to critically evaluate AI-generated reasoning and cross-check with clinical evidence.

ChatGPT offers a scalable and consistent alternative to traditional mentorship in modeling expert reasoning. While it does not replace the nuanced judgment of experienced clinicians, it can enhance learners' exposure to expert-like thinking through structured AI-guided dialogue. Future implementations should ensure that AI outputs are critically examined within a supervised learning framework. To operationalize this, ChatGPT could be integrated into case-based tutorials where students compare its reasoning process with clinician feedback.

3.2. Self-Regulated Learning (SRL)

Self-regulated learning (SRL) refers to the process by which learners proactively manage their cognitive, motivational, and behavioral engagement with educational tasks. It encompasses goal setting, strategic planning, self-monitoring, and reflective evaluation [19]. In medical education, SRL has been associated with improved knowledge retention, diagnostic accuracy, and the development of clinical reasoning skills [20]. Given the dynamic and unpredictable nature of clinical training, specifically during clinical clerkships with variability in supervision and feedback, fostering SRL strategies such as goal-setting, mentoring, and structured feedback has been identified as essential to support students navigating complex learning environments [21].

ChatGPT offers significant potential to support SRL by providing learners with an interactive, low-stakes environment in which they can ask questions, clarify doubts, and explore clinical topics at their own pace. As a conversational agent, ChatGPT can function as an accessible learning companion that promotes cognitive engagement through dialogue [22]. Students can use the tool to review pathophysiology, generate differential diagnoses, and practice clinical decision-making without the pressure of formal assessment or time-limited instruction.

One of the key benefits of ChatGPT in the context of SRL is its alignment with metacognitive regulation [23]. Learners can use the tool to monitor their understanding of clinical concepts by testing hypotheses and receiving immediate, text-based explanations. For instance, a student may ask

ChatGPT to explain the rationale behind a particular laboratory test or to compare treatment options for a specific diagnosis [15]. By receiving feedback in real time, learners can identify gaps in their knowledge and adjust their learning strategies accordingly.

ChatGPT facilitates strategic help-seeking, which is a central component of SRL [24]. In traditional settings, students may hesitate to ask questions due to fear of judgment or lack of access to supervisors. ChatGPT reduces these barriers by offering private, judgment-free interactions that encourage inquiry. This can be especially valuable for early-stage learners who are still developing the confidence to engage in clinical discussions.

ChatGPT can support goal-directed behavior by allowing students to customize their learning sessions [25]. A learner preparing for an internal medicine exam may, for example, use ChatGPT to generate practice questions, simulate clinical cases, or summarize key topics. These features empower students to take ownership of their learning and to structure study plans that align with personal goals and timelines.

It is important to acknowledge the limitations of AI-supported SRL. The unregulated nature of ChatGPT may lead to the reinforcement of misconceptions if learners are not trained to critically evaluate its outputs. The absence of human feedback also limits the depth of reflection and emotional support that traditional mentorship provides [26]. Therefore, while ChatGPT can enhance independent learning, it should be integrated into a broader pedagogical framework that includes guided reflection and peer or faculty feedback.

ChatGPT aligns well with the principles of SRL by enabling personalized, low-pressure inquiry that encourages metacognitive monitoring and strategic help-seeking. Compared to conventional self-study methods, its interactivity provides a more responsive and engaging learning experience. However, its integration should be accompanied by guidance to ensure learners are equipped to assess the credibility of information provided.

3.3. Constructivist Learning Theory

Constructivist learning theory posits that learners actively construct knowledge through experience, reflection, and interaction with their environment [27]. Rather than receiving information passively, students build understanding by integrating new concepts into their existing cognitive frameworks. In the context of medical education, constructivism has informed various learner-centered pedagogies, including problem-based learning (PBL), case-based learning (CBL), and clinical simulations [28]. These approaches emphasize active engagement, critical thinking, and the application of knowledge to authentic clinical scenarios.

One of the key components of constructivist learning is dialogic interaction, which involves questioning, clarifying, and negotiating meaning. In clinical education, such interaction traditionally occurs between students and clinical preceptors, or within small group discussions. However, access to consistent dialogic mentorship can be limited in clinical environments due to variability in supervision and institutional constraints [29]. In this regard, ChatGPT offers a promising supplement by enabling continuous, text-based interaction that mimics educational dialogue.

Through conversational exchanges with ChatGPT, students are able to engage in iterative questioning and hypothesis testing, both of which are central to constructivist learning [30]. For example, a student faced with a case of chest pain may ask ChatGPT to compare potential diagnoses, explore pathophysiological mechanisms, and evaluate treatment options. The model's capacity to generate explanatory responses encourages learners to refine their thinking, challenge assumptions, and connect theoretical knowledge with clinical application.

ChatGPT also supports contextualized learning, an essential principle of constructivism, by generating realistic clinical scenarios. It enhances learning retention and deep comprehension helping learners connect symptoms, diagnostics, and interventions within coherent case narratives [31]. By prompting ChatGPT with realistic case descriptions, learners can receive structured feedback that helps them understand the relevance and interrelationships among symptoms, diagnostics, and interventions. This process facilitates deeper comprehension and supports the formation of illness scripts, which are essential for expert-level reasoning [32].

ChatGPT can enhance learner autonomy, allowing students to pursue inquiries based on their own clinical interests and knowledge gaps [33]. This aligns with constructivist principles that prioritize learner agency and intrinsic motivation. When students generate their own questions and direct their own exploration, they become active participants in the learning process, which is especially important in complex, open-ended domains such as clinical medicine.

The increasing use of large language models (LLMs) such as ChatGPT in medical education demonstrates their potential to support constructivist learning by fostering critical thinking, problem-solving, and self-directed engagement with clinical content [34]. LLMs contribute to active knowledge construction through iterative exploration and interactive dialogue with complex medical scenarios.

While ChatGPT enables constructivist dialogue, it is limited in its ability to assess a student's emotional state, provide affective support, or adapt instruction based on subtle nonverbal cues elements that are essential in human teaching. Its explanations, although often coherent, can occasionally oversimplify complex concepts or present inaccurate information without clear reference to sources. To address these limitations, the integration of ChatGPT into constructivist pedagogy should be accompanied by thoughtful instructional design that emphasizes critical evaluation and fosters metacognitive awareness.

4. Themes of the Review

4.1. Support in Diagnosis, Case Analysis, and Personalized Learning

ChatGPT has shown significant potential in supporting diagnostic reasoning and case analysis among medical students (Fig. 2). Through its capacity to simulate clinical dialogues and patient scenarios, the model enables learners to engage in realistic, text-based simulations that reinforce the practical application of medical knowledge. These simulations provide a safe, controlled environment for students to practice diagnostic and therapeutic decision-making, particularly within problem-based learning formats. Studies have demonstrated that ChatGPT can serve as a diagnostic aid by offering real-time suggestions for differential diagnoses and treatment options, effectively enhancing clinical reasoning and decision-making skills during case-based exercises [35]. In several instances, it has been used to simulate virtual patient interactions, offering immediate feedback that helps learners refine their approach to clinical cases and build diagnostic confidence [36].

However, limitations remain, particularly in the interpretation of laboratory results and diagnostic imaging. ChatGPT has shown weaknesses in accurately analyzing lab values, which suggests a need for further refinement of its domain-specific capabilities in these areas [37]. Despite this, the model still provides educational value by simplifying complex medical concepts and guiding students through logical next steps in clinical decision-making. In this way, it can aid in the incremental development of students' interpretive skills.

Beyond its role in diagnosis, ChatGPT contributes to interactive and personalized learning. It supports learner engagement through adaptive, responsive dialogue that mirrors real clinical interactions. By responding to user input with context-specific information, ChatGPT promotes deeper cognitive processing and sustained attention during study sessions [38]. The model provides tailored learning experiences by adjusting content based on the learner's queries, allowing for immediate clarification of misunderstood concepts and on-demand reinforcement of knowledge. This level of personalization can help address individual learning needs and enhance self-regulated learning among medical students, making ChatGPT a valuable complement to traditional instructional strategies.

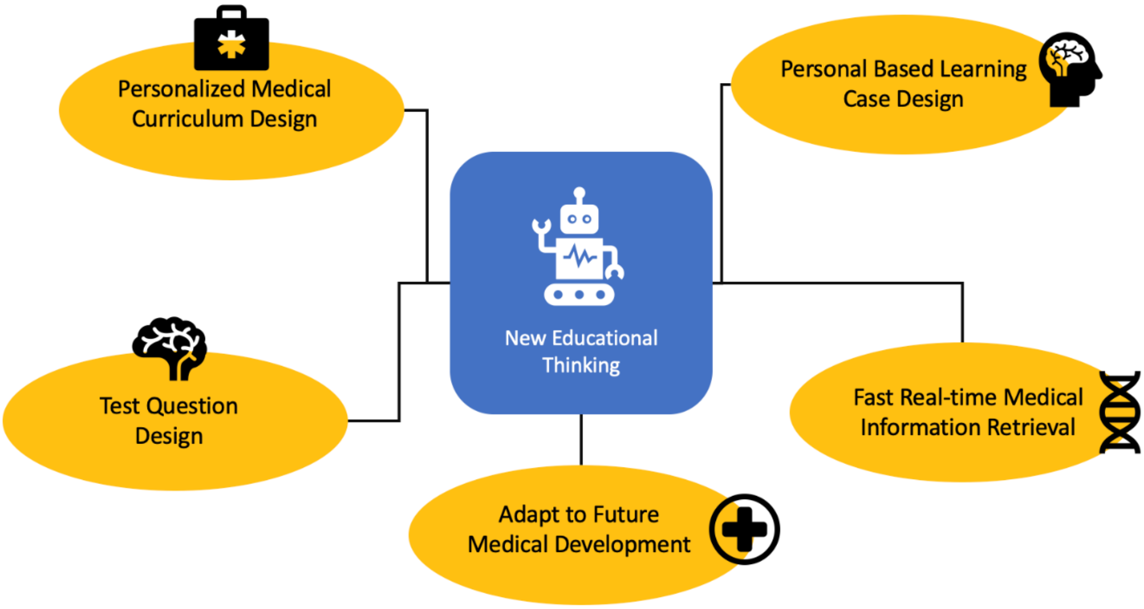

Fig. 2. ChatGPT's Role in Medical Education

4.2. ChatGPT in Reflective and Self-Directed Learning in Clinical Training

ChatGPT, as a large language model, has emerged as a promising tool for enhancing reflective and self-directed learning in clinical training (Fig. 3). Its capabilities support learners in engaging with medical content interactively, promoting deeper understanding and autonomy in clinical reasoning development. One notable application is the use of ChatGPT-4o as a virtual standardized patient, enabling medical interns to simulate patient encounters and manage clinical scenarios without endangering real patients [35]. This simulation-based learning provides a controlled, low-risk environment where students can develop clinical reasoning, problem-solving, and crisis management skills. While such implementations have proven cost-effective and accessible, some technical limitations such as occasional connectivity issues and natural language processing constraints have been reported.

In addition, ChatGPT has been utilized to assist in clinical decision support, particularly in generating differential diagnosis lists and supporting structured clinical judgments [39]. By delivering real-time feedback and access to evidence-informed reasoning, ChatGPT can enhance decision-making accuracy and help learners internalize systematic diagnostic approaches. These functions make it a valuable complement to traditional educational methods.

The interactive nature of ChatGPT also contributes significantly to medical education by allowing learners to pose questions, seek clarification, and receive immediate, relevant responses. Such interactivity supports subjective learning processes, facilitates the development of clinical writing skills, and strengthens clinical judgment [40]. Growing concerns have been raised regarding the potential for academic dishonesty and the risk of overreliance on AI tools, highlighting the need for careful integration into formal curricula.

ChatGPT's strength in processing vast volumes of medical literature makes it highly effective in supporting self-directed learning [7]. Students can use it to explore complex topics such as pathophysiology, treatment algorithms, and emerging guidelines in a personalized and adaptive manner. The platform offers instant access to relevant content, thereby supporting on-demand learning tailored to individual needs [41]. Simulated clinical scenario dialogues further enhance the reflective learning process by allowing students to iteratively test their reasoning, receive feedback, and revise their understanding, ultimately improving their clinical preparedness and patient communication skills.

The integration of ChatGPT into medical education is supported by empirical evidence demonstrating that health profession students generally perceive it as a valuable tool for retrieving medical information and fostering innovative learning approaches [42]. Although concerns remain regarding critical thinking and the need for verification, current literature emphasizes ChatGPT's potential to enhance self-directed learning in medical education. When integrated with appropriate guidance, the tool may effectively contribute to the development of autonomous learning behaviors among medical trainees. This suggests that ChatGPT, when used critically and responsibly, can serve as a valuable partner in cultivating independent learning behaviors among medical trainees.

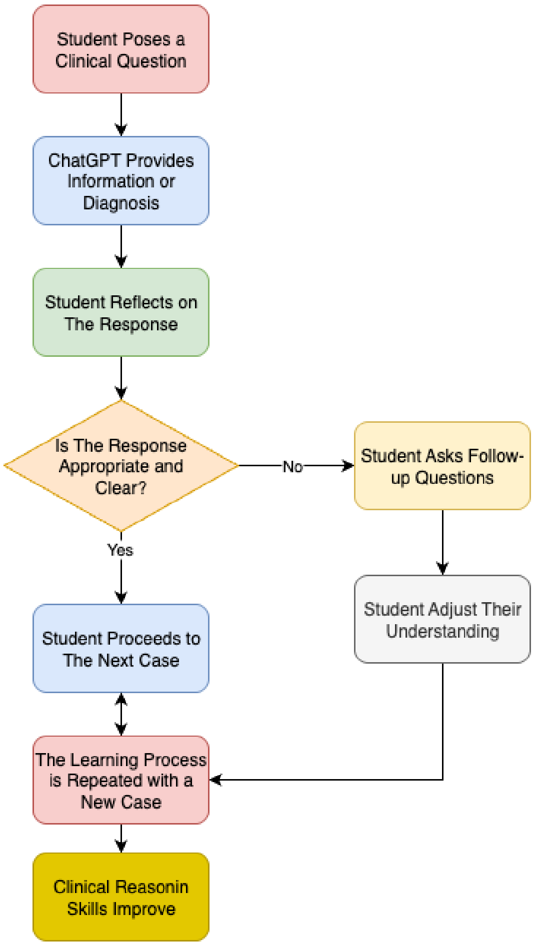

Fig. 3. AI-based reflective and self-directed learning support flow

4.3. Opportunities and Advantages

ChatGPT, a state-of-the-art language model powered by generative artificial intelligence, presents a wide range of opportunities and pedagogical advantages for enhancing clinical reasoning within medical education [5]. Its application as a learning tool has the potential to transform how medical students engage with complex clinical content, reflect on diagnostic processes, and receive real-time cognitive support. Compared to traditional e-learning platforms, ChatGPT offers higher interactivity and adaptability, which can better sustain learner engagement, although its reliance on probabilistic reasoning may limit reliability in complex cases.

One of the primary opportunities offered by ChatGPT lies in its capacity to enable interactive learning. Through natural language conversations, students can ask questions, receive immediate responses, and engage in simulated dialogues that mirror tutor-student interactions [43]. This interactivity allows learners to clarify concepts, practice diagnostic reasoning, and explore clinical cases in a low-risk environment. Prior studies have indicated that such AI-driven engagement supports the development of applied medical knowledge and strengthens the learner's ability to transfer theoretical understanding to clinical scenarios.

ChatGPT also facilitates personalized education, adapting to the unique needs and learning styles of individual students. It can generate customized case scenarios, offer tailored feedback, and deliver practice questions that align with each student's level of proficiency [44]. This adaptability fosters self-directed learning and supports varied educational pathways, particularly beneficial in competency-based curricula and asynchronous learning environments.

ChatGPT also serves as a supplementary tool for clinical decision support, assisting students in generating differential diagnoses and evaluating possible diagnostic pathways [45]. Research has shown that ChatGPT can accurately analyze case inputs and, in certain structured scenarios, perform at levels comparable to or exceeding that of medical students. Although such tools are not substitutes for clinical judgment, they offer a valuable platform for reasoning practice and formative assessment [46].

Another notable contribution of ChatGPT is its role in enhancing clinical reasoning skills. By engaging with structured prompts and clinical case simulations, students are encouraged to think critically, justify their decisions, and refine their diagnostic acumen [47]. The AI supports problem-based learning methods and encourages hypothesis generation, hypothesis testing, and reflection key components of effective clinical reasoning.

In terms of efficiency and accessibility, ChatGPT provides learners with rapid access to synthesized medical information from across disciplines, supporting quick reference and concept reinforcement [48]. This feature can accelerate the learning process, particularly for students in time-constrained clinical settings. Additionally, as a freely accessible or low-cost tool, ChatGPT contributes to reducing barriers in medical education and offers a scalable solution for institutions with limited resources [25].

There are also significant advantages for educators. ChatGPT can support faculty members in designing assessments, developing case-based teaching materials, and generating question banks [49]. Its ability to automate routine tasks, such as content generation or basic feedback provision, allows educators to focus on higher-order instructional activities and individualized mentorship [50]. To leverage these advantages, educators are encouraged to embed ChatGPT within clinical case discussions, assign AI-assisted reflection prompts, and integrate the tool into formative assessment practices to enhance reasoning development.

Finally, while concerns about ethical use, data privacy, and potential algorithmic bias remain valid, these issues can be addressed through robust institutional policies and responsible implementation practices [51]. Ensuring transparency in the use of AI tools, adhering to data protection regulations, and continuously validating the accuracy of ChatGPT outputs are essential steps to mitigate potential risks and to promote trust in AI-assisted education [52].

4.4. Limitations and Pedagogical Risks

While ChatGPT offers notable advantages in supporting clinical reasoning and learner autonomy, its use in medical education is not without significant limitations and pedagogical risks. One of the primary concerns is the hallucination of clinical information, wherein the model generates inaccurate or fabricated responses that may appear coherent and convincing [53]. Such hallucinations pose a serious threat to the integrity of clinical learning, as students may unknowingly absorb incorrect information [54]. These errors are particularly dangerous in medical education, where precision and evidence-based reasoning are paramount. In some studies, ChatGPT has produced clinically inappropriate suggestions or inconsistently applied medical logic, raising ethical concerns about its role in critical decision-making contexts [55].

The potential for overdependence on AI is another prominent risk [56]. While ChatGPT facilitates rapid access to medical knowledge, its convenience may discourage students from engaging in deeper learning processes such as hypothesis generation, critical evaluation of evidence, and reflective thinking. There is a growing concern that students may prioritize efficiency over understanding, resulting in a superficial grasp of complex concepts [57]. This trend may impair the development of core competencies, particularly in areas that require nuanced clinical judgment. As a result, the importance of human supervision and validation becomes critical. Educators must play an active role in guiding students to question, verify, and contextualize AI-generated content within the broader framework of clinical reasoning and ethical practice.

Another significant risk is the erosion of human mentorship [58]. The increasing reliance on AI tools in educational contexts could reduce meaningful interactions between students and experienced clinicians [59]. These interactions are essential not only for the acquisition of technical skills but also for the modeling of professional behaviors, communication, and ethical decision-making. Clinical mentors provide irreplaceable insights derived from real-world experience, often conveying tacit knowledge that cannot be replicated by an AI system [60]. Mentorship fosters the humanistic and ethical dimensions of medical practice, aspects that remain largely outside the scope of generative AI.

In light of these limitations, the integration of ChatGPT in medical education must be approached with caution and pedagogical intent. While it can augment learning, it should not be seen as a replacement for critical thinking, faculty engagement, or professional mentorship [61]. Clear guidelines, robust validation mechanisms, and reflective learning strategies are essential to mitigate the risks associated with its use and to ensure that AI supports, rather than supplants, the human elements of clinical education [62].

To provide a balanced overview of ChatGPT's role in medical education, it is important to juxtapose its pedagogical strengths with the associated risks. While ChatGPT offers unique advantages in promoting clinical reasoning and learner autonomy, it also introduces potential challenges that must be addressed through thoughtful implementation. Table 1 summarizes the key opportunities and limitations identified across the reviewed literature.

Table 1. Pedagogical Opportunities and Risks of ChatGPT in Clinical Education

Aspect	Opportunities / Strengths	Risks / Limitations
Clinical Reasoning	Supports case analysis and differential diagnosis	May generate hallucinated or inaccurate diagnoses
Self-Directed Learning	Enables autonomous, interactive learning	Risk of superficial understanding without critical evaluation
Feedback and Reflection	Offers instant clarification and simulated dialogue	Lacks emotional nuance and human judgment
Accessibility & Scalability	Useful for low-resource settings, easy to integrate into LMS	Potential for reduced reliance on human mentorship

4.5. Best Practices and Implementation Framework

To ensure the pedagogically sound and ethically responsible use of ChatGPT in medical education, institutions must develop structured guidelines and implementation frameworks that align with curricular objectives and professional standards. As the integration of generative AI becomes increasingly prevalent, clear institutional policies are necessary to guide both faculty and students in the effective use of these tools.

One essential component is the development of ethical guidelines and usage protocols. Institutions should provide training that educates learners about the limitations of ChatGPT, including the risks of misinformation, hallucinations, and ethical dilemmas related to academic integrity [63]. This includes setting boundaries for appropriate use in coursework, patient-related discussions, and assessments. Faculty should also be equipped with resources to evaluate AI-supported learning and to model responsible use of AI tools in clinical contexts.

Another critical step is the integration of ChatGPT into formal learning environments, such as Learning Management Systems (LMS). Embedding ChatGPT functionalities within LMS platforms can create structured and trackable opportunities for interactive learning [64]. Students can engage in guided AI-assisted case discussions, receive formative feedback, or complete reflection exercises based on simulated clinical conversations. These features can be monitored by instructors to ensure alignment with learning goals and to assess student progress.

In clinical skills training, the use of ChatGPT can be incorporated into Objective Structured Clinical Examinations (OSCEs) as a supplementary tool [65]. For example, it can simulate patient responses or function as a pre-assessment tutor, allowing students to prepare for OSCE scenarios through AI-based practice sessions. Although ChatGPT cannot replace live standardized patients, it can serve as a valuable addition in preparatory phases, particularly for reinforcing clinical reasoning steps and interview strategies.

ChatGPT can be meaningfully embedded into Problem-Based Learning (PBL) modules. Within PBL settings, students often encounter ambiguous clinical cases that require collaborative inquiry and hypothesis testing [66]. By using ChatGPT during group discussions or self-study periods, learners can access immediate explanations, explore alternative diagnoses, and clarify unfamiliar concepts [32]. This enhances their ability to contribute meaningfully to team-based learning and strengthens the self-directed aspects of PBL.

To maximize educational outcomes, institutions should adopt a blended AI-human approach in which ChatGPT complements, rather than replaces, traditional instruction and mentorship. Regular faculty facilitation, reflective debriefing sessions, and peer discussions are essential to contextualize AI interactions and maintain human-centered learning [67]. Continuous evaluation of AI integration through student feedback, performance metrics, and longitudinal studies is necessary to refine best practices and ensure alignment with educational values.

5. Discussion

The integration of ChatGPT into medical education presents a significant shift in how clinical reasoning is taught, practiced, and assessed. Findings from the reviewed literature suggest that ChatGPT holds strong pedagogical potential to support medical students across a variety of learning domains, including diagnosis, case analysis, self-directed learning, and critical reflection. While the tool presents numerous advantages, it also introduces notable limitations and risks that require thoughtful mitigation through structured implementation and faculty guidance.

Across several studies, ChatGPT was found to be an effective aid in diagnostic reasoning and case-based learning. Its capacity to simulate patient scenarios and clinical dialogue supports experiential learning in a safe, low-stakes environment. Through its real-time responsiveness and adaptability, ChatGPT enables learners to explore differential diagnoses, justify decision-making, and refine their approach to clinical cases. These capabilities align closely with principles of problem-based learning and cognitive apprenticeship, where reasoning must be made explicit and iteratively practiced [32]. ChatGPT's ability to deliver personalized feedback and generate learner-specific content enhances both engagement and cognitive retention, promoting the development of self-regulated learning skills.

Beyond diagnostic support, the tool also facilitates reflective and autonomous learning. By providing interactive responses and on-demand access to complex clinical content, ChatGPT supports students in clarifying difficult concepts and testing their clinical reasoning in a manner that encourages independent exploration. It has been successfully used in simulation-based learning as a virtual standardized patient, allowing students to rehearse communication strategies, decision-making, and case management [35]. These functions extend the classroom experience and are particularly valuable for preparing learners before high-stakes clinical encounters or assessments.

Several limitations temper these pedagogical benefits. Among the most concerning is the phenomenon of hallucinated responses, wherein ChatGPT generates clinically inaccurate or fabricated content [68]. Such misinformation, if unrecognized, may undermine learning and pose a threat to the development of sound clinical judgment. The growing convenience of AI may foster overreliance, discouraging deeper engagement and independent reasoning. When used as a substitute rather than a complement to critical thinking, such tools risk promoting superficial learning particularly in the absence of adequate scaffolding and mentorship.

The erosion of faculty-student interaction is another potential consequence of excessive AI integration. Clinical education relies not only on knowledge transmission but also on the modeling of professionalism, ethics, and context-aware reasoning domains that generative AI cannot fully replicate. The value of human mentorship in shaping clinical identity and fostering the moral dimensions of practice remains irreplaceable [32]. The inclusion of ChatGPT in clinical education must not displace faculty involvement but should rather enhance opportunities for reflection, formative feedback, and metacognitive discussion.

To maximize the pedagogical value of ChatGPT, best practices for its integration are essential. These include embedding AI use within formal curricular structures, setting clear ethical guidelines, and establishing institutional policies on appropriate usage. When incorporated into Learning Management Systems, Objective Structured Clinical Examinations (OSCEs), and problem-based learning modules, ChatGPT can serve as a supportive layer in the broader educational ecosystem. Blended approaches that pair AI-driven interaction with faculty-led debriefing and collaborative discussion appear to be the most effective in maintaining academic rigor and professional standards.

Given the current findings, future research should prioritize empirical investigations to substantiate ChatGPT's impact on clinical reasoning and educational outcomes. Rigorous methodologies such as randomized controlled trials, quasi experimental studies, and mixed-method designs are needed to assess its efficacy in diverse learning environments. Multi-institutional and cross-context studies are particularly critical to ensure the generalizability and contextual relevance of implementation frameworks. These efforts will provide a robust evidence base for the scalable, responsible integration of generative AI into medical education.

6. Limitation of The Review

Several limitations should be acknowledged in this review. The available body of literature on the use of ChatGPT and generative AI in medical education remains limited due to the novelty of the topic. The rapid emergence of large language models has only recently begun to attract scholarly attention within the context of clinical training, resulting in a relatively small number of peer-reviewed studies that directly examine their pedagogical implications.

The majority of the reviewed articles are descriptive in nature, consisting primarily of commentaries, conceptual discussions, and early-stage exploratory studies. While these contributions offer valuable insights and theoretical perspectives, they often lack empirical validation or rigorous methodologacal design. Few studies employ experimental or quasi-experimental approaches that systematically measure learning outcomes or compare AI-assisted instruction with traditional educational methods.

There is a scarcity of longitudinal data assessing the sustained impact of ChatGPT on clinical reasoning, critical thinking, or professional development over time. Most findings are based on short-term observations or user perceptions, which may not fully capture the long-term educational effects or potential unintended consequences.

Due to the narrative nature of this review, the selection of studies may be influenced by the availability of English-language sources and publication bias favoring positive results. Although efforts were made to ensure a comprehensive and balanced synthesis, the absence of a formal meta-analysis limits the generalizability of the conclusions.

Future research should aim to address these gaps through well-designed empirical studies, including randomized controlled trials, mixed-methods evaluations, and longitudinal assessments of AI integration in diverse medical education contexts.

7. Conclusion

The integration of ChatGPT and generative AI technologies into medical education presents a transformative opportunity to enhance clinical reasoning, support self-directed learning, and supplement pedagogical practices in clinical training. As a conversational agent, ChatGPT can function as a reflective learning partner, offering students real-time feedback, simulating clinical dialogues, and facilitating access to complex medical knowledge in a personalized and interactive manner. These capabilities align closely with established educational theories such as cognitive apprenticeship, self-regulated learning, and constructivist pedagogy.

While the potential benefits are promising, significant limitations and pedagogical risks must be addressed. Concerns regarding misinformation, overreliance on AI, diminished human mentorship, and ethical challenges highlight the importance of cautious, guided implementation. ChatGPT should be positioned not as a replacement for traditional instruction, but rather as a complementary tool that enhances learner engagement and autonomy under appropriate supervision.

To maximize its educational value, institutions must establish clear ethical guidelines, integrate AI use into structured curricula, and train both faculty and students in responsible usage. Embedding ChatGPT into learning management systems, OSCE preparation, and problem-based learning modules can enrich clinical education when supported by reflective activities and human facilitation. Educators may consider using ChatGPT as a clinical case discussion partner, a diagnostic reasoning coach, or a scaffolding tool for students in simulation-based training.

As the field continues to evolve, future research should prioritize empirical validation, longitudinal studies, and the development of pedagogically informed frameworks for AI integration. Particularly, randomized controlled trials and multi-institutional comparative studies are needed to evaluate ChatGPT's impact on learning outcomes, diagnostic accuracy, and professional readiness. Through deliberate design and critical oversight, ChatGPT has the potential to become a meaningful and ethical asset in the advancement of clinical reasoning and professional development in medical education.

Declarations

Author Contribution: All authors contributed equally to the conception, drafting, and revision of this manuscript. All authors read and approved the final version.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] A. M. Leitão and R. Z. Esteves, "Perception of medical students on the development of the clinical reasoning competence," *Rev Bras Educ Med*, vol. 47, no. 1, 2023, https://doi.org/10.1590/1981-5271v47.1-20220127.ing.

[2] N. Irvine, R. Van Der Meer, and I. Megiddo, "Expert decision-making in clinicians: An auto-analytic ethnographic study of operational decision-making in urgent care," *PLoS One*, vol. 20, no. 1, p. e0311748, Jan. 2025, https://doi.org/10.1371/journal.pone.0311748.

[3] Z. Maqsood, M. Sajjad, and R. Yasmin, "Effect of feedback-integrated reflection, on deep learning of undergraduate medical students in a clinical setting," *BMC Med Educ*, vol. 25, no. 1, p. 66, Jan. 2025, https://doi.org/10.1186/s12909-025-06648-3.

[4] F. L. Wagner et al., "Current status and ongoing needs for the teaching and assessment of clinical reasoning – an international mixed-methods study from the students` and teachers` perspective," *BMC Med Educ*, vol. 24, no. 1, p. 622, Jun. 2024, https://doi.org/10.1186/s12909-024-05518-8.

[5] M. Sallam, “ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns,” *Healthcare*, vol. 11, no. 6, p. 887, Mar. 2023, https://doi.org/10.3390/healthcare11060887.

[6] F. Puleio, G. Lo Giudice, A. M. Bellocchio, C. E. Boschetti, and R. Lo Giudice, “Clinical, Research, and Educational Applications of ChatGPT in Dentistry: A Narrative Review,” *Applied Sciences*, vol. 14, no. 23, p. 10802, Nov. 2024, https://doi.org/10.3390/app142310802.

[7] W. Shalong et al., “Enhancing self-directed learning with custom GPT AI facilitation among medical students: A randomized controlled trial,” *Med Teach*, vol. 47, no. 7, pp. 1126–1133, Jul. 2025, https://doi.org/10.1080/0142159X.2024.2413023.

[8] M. A. Adarkwah, S. A. Badu, E. A. Osei, E. Adu-Gyamfi, J. Odame, and K. Schneider, “ChatGPT in healthcare education: a double-edged sword of trends, challenges, and opportunities,” *Discover Education*, vol. 4, no. 1, p. 14, Jan. 2025, https://doi.org/10.1007/s44217-024-00393-3.

[9] A. Shah and A. G. Soosai Raj, “A Review of Cognitive Apprenticeship Methods in Computing Education Research,” in *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, New York, NY, USA: ACM, Mar. 2024, pp. 1202–1208, https://doi.org/10.1145/3626252.3630769.

[10] B. T. Robbins et al., “Assessing Cognitive Apprenticeship Impact on Clinical Reasoning in Third-Year Student Pharmacists,” *Am J Pharm Educ*, vol. 88, no. 1, p. 100625, Jan. 2024, https://doi.org/10.1016/j.ajpe.2023.100625.

[11] M. Tariq, S. Iqbal, S. I. Haider, and A. Abbas, “Using the cognitive apprenticeship model to identify learning strategies that learners view as effective in ward rounds,” *Postgrad Med J*, vol. 97, no. 1143, pp. 5–9, Jan. 2021, https://doi.org/10.1136/postgradmedj-2020-137519.

[12] M. Salajegheh, A. Rooholamini, and A. Norouzi, “Investigating the role of clinical exposure on motivational self-regulation skills in medical students based on cognitive apprenticeship model,” *BMC Med Educ*, vol. 24, no. 1, p. 257, Mar. 2024, https://doi.org/10.1186/s12909-024-05253-0.

[13] T. Minocha, T. Bhagatwala, G. Mirzoyan, G. McDowell, and S. C. Fankhauser, “Empowering future scientists: mentors employ various strategies to engage students in professional science disciplinary literacy practices,” *Disciplinary and Interdisciplinary Science Education Research*, vol. 7, no. 1, p. 5, Feb. 2025, https://doi.org/10.1186/s43031-025-00124-0.

[14] H. Yusuf, A. Money, and D. Daylamani-Zad, “Pedagogical AI conversational agents in higher education: a conceptual framework and survey of the state of the art,” *Educational technology research and development*, vol. 73, no. 2, pp. 815–874, Apr. 2025, https://doi.org/10.1007/s11423-025-10447-4.

[15] K. Y. Chan, T. H. Yuen, and M. Co, “Using ChatGPT for medical education: the technical perspective,” *BMC Med Educ*, vol. 25, no. 1, p. 201, Feb. 2025, https://doi.org/10.1186/s12909-025-06785-9.

[16] M. B. Garcia, R. S. de Almeida, D. P. Acut, R. P. P. de Almeida, P. S. Garcia, and E. Stefani, “Teaching Medicine With Generative Artificial Intelligence (GenAI),” *in Teaching in the Age of Medical Technology*, *IGI Global*, 2025, pp. 123–156, https://doi.org/10.4018/979-8-3373-1519-5.ch005.

[17] L.-H. Hsieh, W.-C. Liao, and E.-Y. Liu, “Feasibility assessment of using ChatGPT for training case conceptualization skills in psychological counseling,” *Computers in Human Behavior: Artificial Humans*, vol. 2, no. 2, p. 100083, Aug. 2024, https://doi.org/10.1016/j.chbah.2024.100083.

[18] T. Nguyen, “ChatGPT in Medical Education: A Precursor for Automation Bias?,” *JMIR Med Educ*, vol. 10, p. e50174, Jan. 2024, https://doi.org/10.2196/50174.

[19] G. Alfakhry et al., “Scaffolding Self-Regulated Learning in Operative Dentistry Through Self-Assessment Training,” *J Med Educ Curric Dev*, vol. 11, Jan. 2024, https://doi.org/10.1177/23821205241226820.

[20] S. Li, J. Zheng, S. P. Lajoie, H. Li, D. Pu, and H. Wu, “The Relationship Between Self-Regulated Learning Competency and Clinical Reasoning Tendency in Medical Students,” *Med Sci Educ*, vol. 33, no. 6, pp. 1335–1345, Oct. 2023, https://doi.org/10.1007/s40670-023-01909-6.

[21] S. Fatima, W.-H. Hong, M. N. Mohd Noor, C. C. Foong, and V. Pallath, “Evaluating the Instructional Strategies Influencing Self-Regulated Learning in Clinical Clerkship Years: A Mixed Studies Review,” *Teach Learn Med*, pp. 1–19, Feb. 2025, https://doi.org/10.1080/10401334.2025.2468953.

[22] J. Cross et al., "Assessing ChatGPT's Capability as a New Age Standardized Patient: Qualitative Study," *JMIR Med Educ*, vol. 11, pp. e63353–e63353, May 2025, https://doi.org/10.2196/63353.

[23] N. A. Dahri et al., "Extended TAM based acceptance of AI-Powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study," *Heliyon*, vol. 10, no. 8, p. e29317, Apr. 2024, https://doi.org/10.1016/j.heliyon.2024.e29317.

[24] M. Zhu, "Leveraging ChatGPT to Support Self-Regulated Learning in Online Courses," *TechTrends*, Apr. 2025, https://doi.org/10.1007/s11528-025-01075-z.

[25] R. Scherr, F. F. Halaseh, A. Spina, S. Andalib, and R. Rivera, "ChatGPT Interactive Medical Simulations for Early Clinical Education: Case Study," *JMIR Med Educ*, vol. 9, p. e49877, Nov. 2023, https://doi.org/10.2196/49877.

[26] M. Lan and X. Zhou, "A qualitative systematic review on AI empowered self-regulated learning in higher education," *NPJ Sci Learn*, vol. 10, no. 1, p. 21, May 2025, https://doi.org/10.1038/s41539-025-00319-0.

[27] J. C. Trullàs, C. Blay, E. Sarri, and R. Pujol, "Effectiveness of problem-based learning methodology in undergraduate medical education: a scoping review," *BMC Med Educ*, vol. 22, no. 1, p. 104, Dec. 2022, https://doi.org/10.1186/s12909-022-03154-8.

[28] D. Jiang et al., "Effect of integrated case-based and problem-based learning on clinical thinking skills of assistant general practitioner trainees: a randomized controlled trial," *BMC Med Educ*, vol. 25, no. 1, p. 62, Jan. 2025, https://doi.org/10.1186/s12909-025-06634-9.

[29] M. Zhang and W. Hu, "Application of PBL combined with CBL teaching method in clinical teaching of vascular surgery," *PLoS One*, vol. 19, no. 8, p. e0306653, Aug. 2024, https://doi.org/10.1371/journal.pone.0306653.

[30] C. W. Safranek, A. E. Sidamon-Eristoff, A. Gilson, and D. Chartash, "The Role of Large Language Models in Medical Education: Applications and Implications," *JMIR Med Educ*, vol. 9, p. e50945, Aug. 2023, https://doi.org/10.2196/50945.

[31] A. Higashitsuji, T. Otsuka, and K. Watanabe, "Impact of ChatGPT on case creation efficiency and learning quality in case-based learning for undergraduate nursing students," *Teaching and Learning in Nursing*, vol. 20, no. 1, pp. e159–e166, Jan. 2025, https://doi.org/10.1016/j.teln.2024.10.002.

[32] Z. Hui, Z. Zewu, H. Jiao, and C. Yu, "Application of ChatGPT-assisted problem-based learning teaching method in clinical medical education," *BMC Med Educ*, vol. 25, no. 1, p. 50, Jan. 2025, https://doi.org/10.1186/s12909-024-06321-1.

[33] K. Naznin, A. Al Mahmud, M. T. Nguyen, and C. Chua, "ChatGPT Integration in Higher Education for Personalized Learning, Academic Writing, and Coding Tasks: A Systematic Review," *Computers*, vol. 14, no. 2, p. 53, Feb. 2025, https://doi.org/10.3390/computers14020053.

[34] A. Lawson McLean, "Constructing knowledge: the role of AI in medical learning," *Journal of the American Medical Informatics Association*, vol. 31, no. 8, pp. 1797–1798, Aug. 2024, https://doi.org/10.1093/jamia/ocae124.

[35] S. Öncü, F. Torun, and H. H. Ülkü, "AI-powered standardised patients: evaluating ChatGPT-4o's impact on clinical case management in intern physicians," *BMC Med Educ*, vol. 25, no. 1, p. 278, Feb. 2025, https://doi.org/10.1186/s12909-025-06877-6.

[36] A. Aster, S. V. Ragaller, T. Raupach, and A. Marx, "ChatGPT as a Virtual Patient: Written Empathic Expressions During Medical History Taking," *Med Sci Educ*, vol. 35, no. 3, pp. 1513–1522, Feb. 2025, https://doi.org/10.1007/s40670-025-02342-7.

[37] A. C. Alexander, S. Somineni Raghupathy, and K. M. Surapaneni, "An assessment of the capability of ChatGPT in solving clinical cases of ophthalmology using multiple choice and short answer questions," *Advances in Ophthalmology Practice and Research*, vol. 4, no. 2, pp. 95–97, May 2024, https://doi.org/10.1016/j.aopr.2024.01.005.

[38] Y. Wu, Y. Zheng, B. Feng, Y. Yang, K. Kang, and A. Zhao, "Embracing ChatGPT for Medical Education: Exploring Its Impact on Doctors and Medical Students," *JMIR Med Educ*, vol. 10, p. e52483, Apr. 2024, https://doi.org/10.2196/52483.

[39] T. Hirosawa, Y. Harada, K. Mizuta, T. Sakamoto, K. Tokumasu, and T. Shimizu, "Evaluating ChatGPT-4's Accuracy in Identifying Final Diagnoses Within Differential Diagnoses Compared With Those of Physicians: Experimental Study for Diagnostic Cases," *JMIR Form Res*, vol. 8, p. e59267, Jun. 2024, https://doi.org/10.2196/59267.

[40] J. Cross et al., "Assessing ChatGPT's Capability as a New Age Standardized Patient: Qualitative Study," *JMIR Med Educ*, vol. 11, pp. e63353–e63353, May 2025, https://doi.org/10.2196/63353.

[41] F. E. Çiçek, M. Ülker, M. Özer, and Y. S. Kıyak, "ChatGPT versus expert feedback on clinical reasoning questions and their effect on learning: a randomized controlled trial," *Postgrad Med J*, vol. 101, no. 1195, pp. 458–463, Apr. 2025, https://doi.org/10.1093/postmj/qgae170.

[42] L. Moskovich and V. Rozani, "Health profession students' perceptions of ChatGPT in healthcare and education: insights from a mixed-methods study," *BMC Med Educ*, vol. 25, no. 1, p. 98, Jan. 2025, https://doi.org/10.1186/s12909-025-06702-0.

[43] D. García-Torres, M. A. Vicente Ripoll, C. Fernández Peris, and J. J. Mira Solves, "Enhancing Clinical Reasoning with Virtual Patients: A Hybrid Systematic Review Combining Human Reviewers and ChatGPT," *Healthcare*, vol. 12, no. 22, p. 2241, Nov. 2024, https://doi.org/10.3390/healthcare12222241.

[44] D. Ali, Y. Fatemi, E. Boskabadi, M. Nikfar, J. Ugwuoke, and H. Ali, "ChatGPT in Teaching and Learning: A Systematic Review," *Educ Sci (Basel)*, vol. 14, no. 6, p. 643, Jun. 2024, https://doi.org/10.3390/educsci14060643.

[45] K. Shikino et al., "Evaluation of ChatGPT-Generated Differential Diagnosis for Common Diseases With Atypical Presentation: Descriptive Research," *JMIR Med Educ*, vol. 10, pp. e58758–e58758, Jun. 2024, https://doi.org/10.2196/58758.

[46] A. Rao et al., "Assessing the Utility of ChatGPT Throughout the Entire Clinical Workflow: Development and Usability Study," *J Med Internet Res*, vol. 25, p. e48659, Aug. 2023, https://doi.org/10.2196/48659.

[47] A. A. Peralta Ramirez, S. Trujillo López, G. A. Navarro Armendariz, S. A. De la Torre Othón, M. R. Sierra Cervantes, and J. A. Medina Aguirre, "Clinical Simulation with ChatGpt: A Revolution in Medical Education?," *Journal of CME*, vol. 14, no. 1, Dec. 2025, https://doi.org/10.1080/28338073.2025.2525615.

[48] A. Skryd and K. Lawrence, "ChatGPT as a Tool for Medical Education and Clinical Decision-Making on the Wards: Case Study," *JMIR Form Res*, vol. 8, p. e51346, May 2024, https://doi.org/10.2196/51346.

[49] D. Rouabhia, "Artificial Intelligence Driven Course Generation: A Case Study Using ChatGPT," *Atras Journal*, vol. 5, no. 3, pp. 287–302, 2024, https://doi.org/10.70091/atras/AI.18.

[50] H. Crompton and D. Burke, "The Educational Affordances and Challenges of ChatGPT: State of the Field," *TechTrends*, vol. 68, no. 2, pp. 380–392, Mar. 2024, https://doi.org/10.1007/s11528-024-00939-0.

[51] I. M. García-López, C. S. González González, M.-S. Ramírez-Montoya, and J.-M. Molina-Espinosa, "Challenges of implementing ChatGPT on education: Systematic literature review," *International Journal of Educational Research Open*, vol. 8, p. 100401, Jun. 2025, https://doi.org/10.1016/j.ijedro.2024.100401.

[52] I. M. García-López and L. Trujillo-Liñán, "Ethical and regulatory challenges of Generative AI in education: a systematic review," *Front Educ (Lausanne)*, vol. 10, Jun. 2025, https://doi.org/10.3389/feduc.2025.1565938.

[53] D. Roustan and F. Bastardot, "The Clinicians' Guide to Large Language Models: A General Perspective With a Focus on Hallucinations," *Interact J Med Res*, vol. 14, p. e59823, Jan. 2025, https://doi.org/10.2196/59823.

[54] F. Aljamaan et al., "Reference Hallucination Score for Medical Artificial Intelligence Chatbots: Development and Usability Study," *JMIR Med Inform*, vol. 12, p. e54345, Jul. 2024, https://doi.org/10.2196/54345.

[55] J. Zhou et al., "Integrating AI into clinical education: evaluating general practice trainees' proficiency in distinguishing AI-generated hallucinations and impacting factors," *BMC Med Educ*, vol. 25, no. 1, p. 406, Mar. 2025, https://doi.org/10.1186/s12909-025-06916-2.

[56] J. Goddard, "Hallucinations in ChatGPT: A Cautionary Tale for Biomedical Researchers," *Am J Med*, vol. 136, no. 11, pp. 1059–1060, Nov. 2023, https://doi.org/10.1016/j.amjmed.2023.06.012.

[57] M. Gerlich, "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking," *Societies*, vol. 15, no. 1, p. 6, Jan. 2025, https://doi.org/10.3390/soc15010006.

[58] L. Çetinkaya, "Redefining Mentorship in Medical Education with Artificial Intelligence: A Delphi Study on the Feasibility and Implications," *Teach Learn Med*, pp. 1–11, Jun. 2025, https://doi.org/10.1080/10401334.2025.2521001.

[59] C. Preiksaitis and C. Rose, "Opportunities, Challenges, and Future Directions of Generative Artificial Intelligence in Medical Education: Scoping Review," *JMIR Med Educ*, vol. 9, p. e48785, Oct. 2023, https://doi.org/10.2196/48785.

[60] F. Shahzad, S. Chilba, and A. Arslan, "Tacit knowledge exchange among senior management educators: A qualitative study," *The International Journal of Management Education*, vol. 22, no. 2, p. 100973, Jul. 2024, https://doi.org/10.1016/j.ijme.2024.100973.

[61] S. R. Alli, S. Q. Hossain, S. Das, and R. Upshur, "The Potential of Artificial Intelligence Tools for Reducing Uncertainty in Medicine and Directions for Medical Education," *JMIR Med Educ*, vol. 10, pp. e51446–e51446, Nov. 2024, https://doi.org/10.2196/51446.

[62] B. Pohn, L. Mehnen, S. Fitzek, K.-E. (Anna) Choi, R. J. Braun, and S. Hatamikia, "Integrating artificial intelligence into pre-clinical medical education: challenges, opportunities, and recommendations," *Front Educ (Lausanne)*, vol. 10, Mar. 2025, https://doi.org/10.3389/feduc.2025.1570389.

[63] J. Haltaufderheide and R. Ranisch, "The ethics of ChatGPT in medicine and healthcare: a systematic review on Large Language Models (LLMs)," *NPJ Digit Med*, vol. 7, no. 1, p. 183, Jul. 2024, https://doi.org/10.1038/s41746-024-01157-x.

[64] A. V Thomae, C. M. Witt, and J. Barth, "Integration of ChatGPT Into a Course for Medical Students: Explorative Study on Teaching Scenarios, Students' Perception, and Applications," *JMIR Med Educ*, vol. 10, pp. e50545–e50545, Aug. 2024, https://doi.org/10.2196/50545.

[65] S. M. Misra and S. Suresh, "Artificial Intelligence and Objective Structured Clinical Examinations: Using ChatGPT to Revolutionize Clinical Skills Assessment in Medical Education," *J Med Educ Curric Dev*, vol. 11, Jan. 2024, https://doi.org/10.1177/23821205241263475.

[66] X. Tong et al., "The application of problem-based learning (PBL) guided by ChatGPT in clinical education in the Department of Nephrology," *BMC Med Educ*, vol. 25, no. 1, p. 1048, Jul. 2025, https://doi.org/10.1186/s12909-025-07427-w.

[67] C. B. Divito, B. M. Katchikian, J. E. Gruenwald, and J. M. Burgoon, "The tools of the future are the challenges of today: The use of ChatGPT in problem-based learning medical education," *Med Teach*, vol. 46, no. 3, pp. 320–322, Mar. 2024, https://doi.org/10.1080/0142159X.2023.2290997.

[68] Q. Y. Lee, M. Chen, C. W. Ong, and C. S. H. Ho, "The role of generative artificial intelligence in psychiatric education– a scoping review," *BMC Med Educ*, vol. 25, no. 1, p. 438, Mar. 2025, https://doi.org/10.1186/s12909-025-07026-9.