Progress in Medical Education
ISSN: 3007-0007

Review Article https://doi.org/10.61189/141463mjzwgj

Revolutionizing medical education: The role of generative artificial intelligence in medical education

Wenhui Guo[1], Bing Xu[1], Jiaojiao Feng[1], Zui Zou[1,*], Miao Zhou[2,*]

[1]School of Anesthesiology, Second Military Medical University/Naval Medical University, Shanghai 200433, China.
[2]Department of Anesthesiology, The Affiliated Cancer Hospital of Nanjing Medical University, Jiangsu Cancer Hospital, Jiangsu Institute of Cancer Research, Nanjing Medical University, Nanjing 210009, Jiangsu, China.
[*]The authors contribute equally.

Corresponding authors: Zui Zou and Miao Zhou.

Declaration of conflict of interest: None.

Received August 11, 2025; Accepted September 23, 2025; Published December 31, 2025

Highlights

- This review introduces the core concepts of artificial intelligence (Generative AI) and summarizes the most remarkable Generative AI models applied in medical education.
- This survey systematically presents the main clinical applications of generative AI medical education.
- The challenges and solutions of introducing Generative AI into medical education are discussed to explore future directions for its development and implementation.

Abstract

Generative artificial intelligence (Generative AI) is reshaping both learning and teaching paradigms in medical education. With the advancement of Large Language Models (LLMs)-based tools such as ChatGPT, Gemini, and other medical-domain-specific models, Generative AI shows strong potential to address persistent challenges in medical education, including rigid curricula, unequal access to educational resources, and the diverse learning needs of medical students. This review summarizes the applications of Generative AI across key domains: (1) personalized learning through real-time analysis of student performance; (2) clinical skills training via immersive simulations and virtual patients; (3) automated generation of teaching materials such as clinical cases and assessments; and (4) support for student research and academic writing. Empirical evidence indicates that Generative AI-enhanced instruction can improve knowledge acquisition, clinical reasoning, and overall educational efficiency. However, challenges remain, including the generation of inaccurate or fabricated content, risks to academic integrity, algorithmic bias, data privacy concerns, and unresolved ethical issues regarding AI's role in teaching. Without proper oversight, these risks may compromise educational quality and equity. To ensure responsible adoption, this review advocates for the establishment of institutional policies, enhancement of educators' AI literacy, transparent model validation, and a human-centered design framework that positions Generative AI as a collaborative teaching assistant. When responsibly integrated, Generative AI holds the transformative potential to cultivate future medical professionals equipped with clinical competence, responsibility, and innovative thinking.

Keywords: Generative artificial intelligence, medical education, large language models, clinical simulation, personalized learning, human-centered design

Artificial intelligence (AI) and generative AI

Definition of artificial intelligence and generative AI

The term of artificial intelligence (AI) was first introduced by Jphn McCarthy in 1956 and refers to the ability of machines or software systems to perform tasks that normally require human intelligence [1]. Specifically, AI encompasses various advanced technologies, including machine learning, neural networks, and language modeling. In the context of medical education, AI can assist educators by automating

Address correspondence to: Zui Zou, School of Anesthesiology, Second Military Medical University/Naval Medical University, 800 Xiangyin Road, Shanghai 200433, China. Tel: +86 21 81872031. E-mail: zouzui1980@163.com. **Miao Zhou**, Department of Anesthesiology, The Affiliated Cancer Hospital of Nanjing Medical University, Jiangsu Cancer Hospital, Jiangsu Institute of Cancer Research, Nanjing Medical University, Nanjing 210009, Jiangsu, China. Tel: +86 18217567295. E-mail: zhoumiao@jszlyy.com.cn.

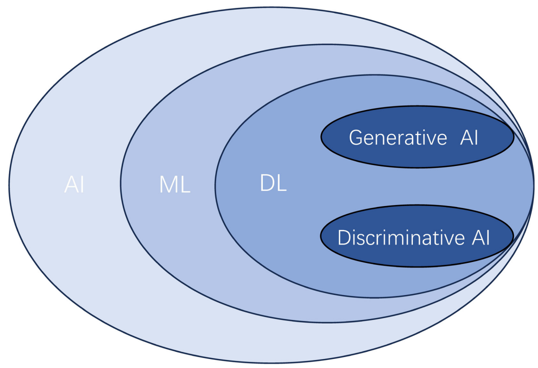

Figure 1. The organization of artificial intelligence technologies.

repetitive tasks, reducing dependence on human resources, and improving the efficiency of teaching and training medical students [2].

Machine learning, as a core branch of AI, enables systems to learn patterns from massive data, allowing them to perform tasks and make decisions based on prior experiences [3]. Deep learning, a specialized branch of machine learning, employs neural networks that mimic the structure and function of the human brain to automatically extract hierarchical features from complex data, providing a foundation for both discriminative and generative models [4]. Discriminant artificial intelligence refers to models designed to classify or distinguish among data types based on probabilistic reasoning. In medicine, such systems are applied, for example, in diagnosing lung tumor by using convolutional neural networks to process chest X-rays or CT images and determine whether a lesion is malignant or benign.

Generation AI is a kind of AI technology that not only judges or classifies data but also creates new data in various modalities, including text, image, audio, video, and even 3D models [5]. The output of such systems are collectively known as *artificial intelligence-generated content* (AIGC), which fundamentally differs from discriminative AI by producing rather than categorizing data [6]. By aligning language models with human intentions through techniques like Reinforcement Learning from Human Feedback (RLHF), generative AI expands beyond narrow, task-specific applications toward more generalized cognitive capabilities [7, 8].

As machine learning, deep learning, discriminative AI, and generative AI are all subfields of AI, and given that "AI" is a broader and more general term, some studies may not clearly distinguish these domains, leading to the widespread use of "AI" as a general descriptor [9]. Understanding the relationships among these components is essential for grasping the staged evolution of intelligent systems [10]. A conceptual framework (**Figure 1**) illustrates the relationship among these technologies.

Landmark models and applications of generative AI

Large models refer to AI systems characterized by a vast number of parameters, enabling enhanced learning capacity [11]. Such large-scale architectures demonstrate superior performance in handling complex tasks. Many advanced Generative AI systems are built upon these large models to produce high-quality outputs. A Large Language Model (LLM) is a subset of large models specifically designed for natural language processing (NLP), enabling machines to understand, generate, and interact using human language [11].

Generative AI has demonstrated transformative potential across multiple domains, including conversational chatbots, image synthesis, software development assistance, and scientific research (e.g., accelerating drug discovery) [12]. Prominent general-purpose models include OpenAI's GPT-3.5 and GPT-4, Google's Gemini, and Stability AI's Midjourney for image generation [13-15]. A landmark in this evolution was the launch of ChatGPT by OpenAI in late 2022, a conversational agent powered by GPT-3.5 and GPT-4 architectures [16]. In China, DeepSeek represents a significant achievement, featuring a hybrid open-source and proprietary architecture optimized for Chinese linguistic and cultural contexts [17]. Like ChatGPT, DeepSeek is built upon large language model technology but fine-tuned to better capture Chinese syntactic and semantics, making it particularly suitable for applications in education, healthcare, and technical fields. The mapping between Generative AI and its application-level models, including ChatGPT and DeepSeek, is illustrated in **Figure 2**.

While general-purpose LLMs such as GPT and Gemini show impressive capabilities, they often lack domain-specific expertise and may pro-

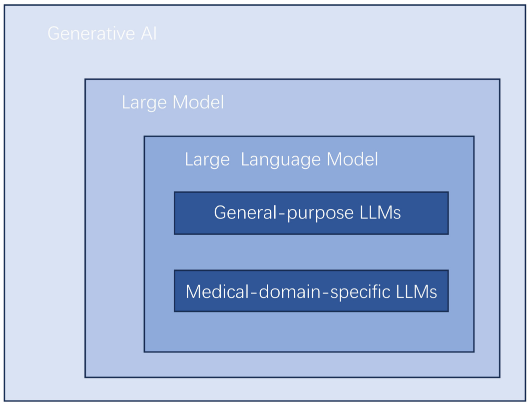

Figure 2. Mapping generative AI to application-level models. AI, artificial intelligence.

duce inaccurate or non-evidence-based content when applied to specialized fields like medicine [18, 19]. To address this limitation, several medical-domain-specific Generative AI models have been developed, including Google's Med-PaLM and Med-PaLM 2 and Med-Gemini, which are designed to improve factual accuracy and clinical reasoning in medical contexts [20-22]. Earlier domain-adaptive efforts include pre-trained language models such as BioBERT and ClinicalBERT, which have demonstrated strong performance in biomedical text understanding tasks [23]. However, these models are not technically considered LLMs. They are based on the Bidirectional Encoder Representations from Transformers (BERT) architecture possess far fewer parameters, and operates as encoder-only systems optimized for natural language understanding (e.g., entity recognition, text classification) rather than open-ended text generation.

As summarized in **Table 1**, medical-domain-specific models often exhibit a narrower scope. They are optimized for biomedical or clinical tasks and struggle with interdisciplinary reasoning or creative educational applications beyond their training domain. Although they achieve higher factual accuracy, their conversational flexibility and interactive naturalness are markedly lower than general-purpose models such as ChatGPT. Furthermore, challenges persist regarding data accessibility and update frequency: many medical models are trained on proprietary or restricted datasets, limiting transparency, reproducibility, and timely integration of emerging medical knowledge.

Applications of generative AI in medical education

The traditional medical education system faces significant challenges amid the exponential expansion of medical knowledge and the rapid advancements in clinical practice, including outdated curricula, unequal access to educational resources, and difficulty in addressing the diverse learning needs of medical students who vary in learning styles, prior knowledge, and career goals [24]. Generative AI offers a transformative solution by enabling more adaptive and personalized learning pathways in medical education [25].

A variety of generative AI tools have been introduced into medical education, differing in design, function, and clinical applicability. As summarized in **Table 1**, comparative evaluation of leading AI platforms is important for medical educators and students to make informed decisions. These platforms represent the forefront of generative AI in global medical education, each offering unique strengths and trade-offs in accessibility, accuracy, language support, and domain adaptation. It is important to note that limitations—such as hallucinations, data privacy risks, and algorithmic biases—are common across all generative AI systems. The attribution reflects the degree to which these issues have been highlighted. For example, ChatGPT has been extensively studied for its accuracy and reliability in generating medical information [26, 27]. In contrast, *DeepSeek* has shown limited adoption outside of China, with its current implementation primarily concentrated in Chinese-speaking regions. As cloud-based AI systems, *Gemini* and *Claude* introduce heightened privacy risks when handling confidential educational or clinical information [28, 29]. Despite its strong general reasoning and linguistic performance, *Claude*—as a general-purpose LLMs—has not been explicitly fine-tuned for medical applications. Consequently, studies have shown that its performance in clinical knowledge tasks behind domain-specific models such as *Med-PaLM*, which have been purpose-built for medical education and healthcare contexts [20].

Having established the transformative potential of Generative AI in addressing challenges in traditional medical education, it is now essential to explore its specific educational applications.

Table 1. Comparison of generative AI tools (platforms) in medical education

Tools	Developer	LLM Base	Language Support	Key Features	Limitations
ChatGPT-4	OpenAI	GPT-4	Multilingual	Dialogue, content generation, and multimodal capabilities [13]	Risk of hallucinations, data privacy concerns, and subscription-based access [27]
DeepSeek	DeepSeek AI	DeepSeek series	Bilingual (Chinese in dominance and English)	Code generation and logical reasoning [17]	Limited global adoption
Gemini	Google	PaLM 2/ Gemini	Multilingual	Search-integrated	Privacy concerns [21]
Claude 3	Anthropic	Claude 3	English-focused	Long-context	Less medical fine-tuning [20]
Med-PaLM 2	Google Health	PaLM 2 (medical-tuned)	English	High accuracy in medical QA, physician-level performance [20]	Cannot handle non-clinical or interdisciplinary queries; reduced conversational flexibility; not publicly available

As illustrated in **Table 1**, although generative AI platforms differ in architecture and regional adoption, their shared capacity to generate human-like content underpins their diverse pedagogical uses. The following subsections examine these applications through four primary pathways: delivering personalized learning experiences, simulating clinical environments, automating content creation, and supporting academic research.

Delivering personalized learning experiences

Generative AI can dynamically assess students' learning progress in real time by analyzing multiple data sources such as test scores, forum interactions, and self-assessment records. Based on these inputs, the system recommends individualized learning materials [30]. This adaptive process enables learners to progress at their own pace: when students perform well, the recommended content increases in difficulty; conversely, when they encounter difficulties, the system provides simplified materials and additional guidance. Some educational platforms employ intelligent question-and-answer modules to further personalize instruction. AI tutors on these platforms can even adapt explanations to each learner's academic level [31]. The AI tutors use simpler terminology for early-year students and more advanced concepts for senior trainees. A recent study investigating the integration of AI into medical education analyzed how undergraduate medical students in the UK used AI for self-directed learning [32]. The study considered factors such as age, gender, and academic year, and found that AI-assisted learning tools enhanced medical students' engagement, motivation, and learning efficiency. These findings underscore the potential of Generative AI to foster personalized, student-centered medical training [32].

Creating medical simulation scenarios

Generative AI has shown remarkable potential in clinical skills training by simulating realistic clinical scenarios. These systems allow students to repeatedly practice disease diagnosis, treatment planning, and patient communication using AI-generated cases. Notably, urology interns who received Generative AI-assisted clinical guidance via conversational chatbots demonstrated improved performance compared with traditional training methods [33]. When combined with virtual reality (VR) technology, Generative AI can simulate surgical operations and complex clinical scenarios beyond traditional 2D interfaces. This integration provides medical students with an immersive, safe, and repeatable 3D environment in which they can refine procedural and decision-making skills. The system can track students' operational actions in real time and provide immediate feedback, which is similar to an instructor during bedside teaching [34]. Learners receive corrective prompts through voice or visual cues and can review playback recordings for iterative self-correction. Recent advances in multimodal Generative AI have further enhanced the fidelity of these simulations [35]. By combining large language models with image and speech technologies, multimodal systems can create highly realistic virtual patients that exhibit diverse cultural backgrounds and emotional states such as anger, grief, or distrust. The trainees can use this system to practice empathy, communication, and conflict resolution in challenging encounters.

Automating the creation of medical teaching content

Another major application of Generative AI in medical education lies in its support for automated generation of teaching materials. Traditional approaches to lesson planning and examination design are heavily rely on educators, often time-consuming, and may lack innovation. In contrast, Generative AI-based tools can autonomously produce high-quality educational resources—such as clinical case scenarios, and quizzes, tailored to students' learning progress and curricular content. A randomized controlled trial evaluated the quality of clinical cases and examination questions generated by *ChatGPT* [36]. In this study, medical students were randomly assigned to two groups: one studied cases generated by ChatGPT, while the other group learned from cases written by faculty members. Researchers then compared students' comprehension and test performance. The results indicated that the quality of AI-generated clinical cases was comparable to those written by teachers, and that the AI-created test questions were generally acceptable, although minor inaccuracies were observed [36].

Supporting thesis writing and scientific research

Generative AI also demonstrates considerable potential in enhancing research training and academic writing. Traditionally, inexperienced students rely heavily on experienced mentors, engaging in repeated revisions to improve their work. Many early-stage medical students struggle with literature retrieval, research design, and manuscript development. Generative AI provides valuable assistance in these areas by enabling users to efficiently access relevant academic literature, generate structured research outlines, refine manuscript organization, and create preliminary drafts of academic writing. For instance, Generative AI systems can automatically retrieve scholarly articles based on user-defined research topics, extract key insights, and generate logically structured introductions and literature reviews [26, 37]. This capability not only enhances students' efficiency in literature synthesis but also supports the development of critical thinking and academic reasoning, and scientific writing skills. Furthermore, in the realm of research methodology, Generative AI can recommend appropriate statistical tests, experimental design frameworks, and questionnaire development strategies based on the research objectives, ultimately advancing the rigor and higher-caliber of student-led research projects [38].

Challenges of generative AI in medical education

Although Generative AI has demonstrated significant potential in medical education—enabling personalized learning, enhancing teaching efficiency, and supporting academic writing—its widespread adoption also raises critical technical, ethical, and pedagogical challenges that must be carefully addressed.

Hallucinations and factual inaccuracy

A major concern regarding Generative AI in medical education is the accuracy and scientific validity of AI-generated content. Given the complexity and high-stakes nature of clinical knowledge, Generative AI models may produce factually incorrect or misleading information—phenomena collectively known as "hallucinations". Such outputs can distort students' understanding and, if uncorrected, may negatively affect their clinical reasoning and future patient care [39, 40]. A multi-center study evaluated the ability of 137 general practice trainees to identify hallucinations in clinical responses generated by ChatGPT-4o, revealing a mean detection accuracy of only 55% [41]. This underscores the urgent need to strengthen AI literacy among both educators and medical students, including the ability to critically evaluate, verify, and contextualize AI-generated information. Cultivating critical thinking is essential to improve learners' discernment and ensure the reliability of AI-assisted education [41].

To mitigate these risks, a combination of technical safeguards and pedagogical interventions is recommended.

Prompt engineering can be employed to constrain model behavior by instructing the AI to generate responses strictly grounded in evidence and to cite authoritative sources such as clinical guidelines [42].

Retrieval-Augmented Generation (RAG) architectures can enhance reliability by integrating responses with verified information form medical databases. Notably, the recently proposed SelfRewardRAG framework combines RAG with large language models to improve medical reasoning through self-verification mechanisms, producing more accurate and trustworthy outputs [43].

Human expert supervision remains indispensable. AI-generated content, especially those involving diagnostic reasoning or therapeutic recommendations, should undergo expert review by qualified clinicians before incorporation into teaching materials or assessments [44].

Data privacy and confidentiality risks

Medical education often involves the use of sensitive health information and real patient data. The incorporation of such data in model training and interaction processes raises substantial ethical and legal challenges [45]. Documented security failures in Generative AI applications have revealed systemic vulnerabilities, including data leakage, unauthorized access, and infrastructure breaches [46]. These vulnerabilities raise serious concerns about compliance with major privacy regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union. Furthermore, they heighten the possibility of malicious alteration of patient records or manipulation of AI-assisted clinical decision-making systems. A cross-sectional study conducted at ABWA Medical College in Pakistan examined medical students' concern about data privacy in relation to the use of ChatGPT in medical education [47]. Approximately 22% of students expressed apprehension regarding potential privacy breaches and the inaccurate information recording. Ensuring ethical data handling is essential to safeguard patient rights and maintain public trust. To address these concerns, several technical and organizational safeguards are recommended. At the user level, reframing or paraphrasing queries to obscure identifiable patient details, as well as employing robust data anonymization and encryption protocols, can substantially reduce privacy risks [48]. Moreover, establishing human oversight mechanisms to review privacy-related content generated or processed by AI systems provides an added layer of protection and accountability [49].

Institutional strategies are equally vital. Deploying localized AI models such as Llama 3 within secure institutional infrastructures—operating behind firewalls without external data transmission—can effectively minimize exposure to third-party servers and enhance compliance with data protection standards [50]. Unlike publicly accessible platforms like *ChatGPT*, these on-premise models offers stronger control over data governance and security. Alternatively, organizations may utilize enterprise-grade AI solutions provided by trusted vendors. For example, several academic medical centers have successfully implemented private Azure OpenAI Studio deployments with secure, API-enabled endpoints, ensuring HIPAA-compliant access to LLMs for researchers handling sensitive patient data [51, 52].

Academic integrity and misuse

Furthermore, the integration of Generative AI into educational settings introduces new challenges to academic integrity. Students may become overly reliant on AI tools to complete assignments and generate research papers. This behavior potentially undermine the development of critical thinking problem-solving, and independent analytical skills [53]. A systematic review examining the dual role of Generative AI in higher education analyzed studies published between 2021 and 2024 and assessed its influence on both innovation and academic ethics [54]. The review found that while Generative AI significantly improves study efficiency and creativity, it concurrently introduces substantial risks of academic misconduct.

A critical challenge in addressing these risks lies in the detection of AI-generated content. Some plagiarism detection tools, such as GPTZero, are currently used in academic institutions to identify AI-assisted writing; However, their accuracy remains limited, particularly when AI-generated content has been heavily modified [55]. Furthermore, research indicates that digital watermarking technologies are also imperfect [56]. These marks can be easily removed through minor textual or visual modifications. Therefore, sole reliance on detection

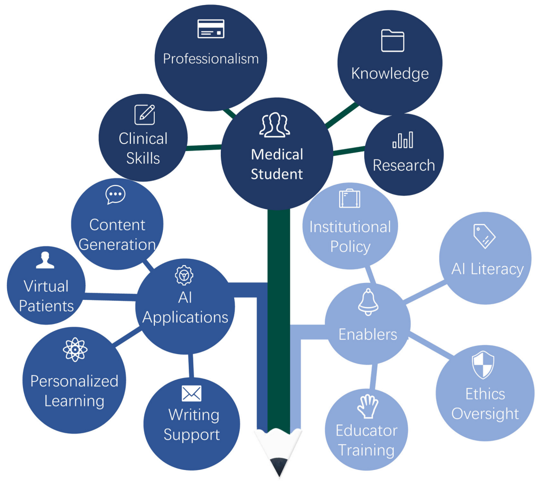

Figure 3. A human-centered framework for integrating generative AI in medical education. AI, artificial intelligence.

tools is insufficient. Rather than imposing outright bans on AI use, educators should adopt new assessment methods, such as oral examinations, practical performance assessments, and longitudinal evaluation of the learning process. Institutions should also require students to explicitly disclose the scope and manner of AI use in their academic work, ensuring transparency and accountability. Equally important is the cultivation of AI literacy and ethical awareness. Students must be educated on the responsible use of AI technologies. To preserve academic authenticity and equity, clear policies and guidance are also required to regulate the responsible use of AI-assisted technologies [57].

Algorithmic bias, health equity, and educational fairness

The application of Generative AI in medical education requires heightened vigilance toward hidden data biases, many of which remain hardly unrecognized by users [58]. A study based on a large-scale language corpora have demonstrated that human biases are often encoded within linguistic patterns, which are then learned and amplified by NLP models [59]. Consequently, Generative AI systems may inadvertently reproduce and reinforce gender, racial, and cultural stereotypes through biased language representations [60]. Evidence further indicates that such biases are not confined to structured datasets, they also exist within unstructured datasets, including visual data. When these data sources are used to train AI systems, they may compromise the fairness, transparency, and reliability of algorithmic decision-making [61]. In the context of medical education, underrepresentation of certain populations—whether by race, language, or socioeconomic status—can lead to unequal model performance and biased outputs. Students from diverse cultural, linguistic, social or economic backgrounds may experience differential learning outcomes when interacting with AI-driven educational tools [62]. Research suggests that such algorithmic bias may substantially influence students' academic pathways and reinforce pre-existing systemic inequities within educational systems. Therefore, ensuring fairness and inclusivity in AI-generated educational content is essential for equitable learning outcomes.

Future of generative AI in medical education

In the era of large language models, Generative AI has demonstrated substantial potential to address persistent challenges in medical education, including rigid curricula, resource shortages, and diverse learner needs. This review highlights four primary areas of applications, including personalized learning, virtual clinical training, automated content generation, and research support, performing well in improving teaching quality and accessibility. While promising, the adoption of Generative AI in clinic teaching entails notable technical, ethical, and pedagogical risks. A preventive and collaborative approach is therefore essential. Medical educators and institutional leaders must actively adapt to these technological transformations by enhancing their digital literacy and embracing human-AI collaborative teaching models [63].

To ensure responsible and sustainable implementation, it is imperative to establish robust ethical and governance frameworks. Without clear oversight and accountability, AI risks being misused or perpetuating inequities in educational access and evaluation. Establishing transparent policies and continuous monitoring mechanisms can safeguard fairness and uphold the integrity of the medical education ecosystem.

To guide this transformation, we introduce a human-centered conceptual framework (**Figure 3**) that positions the medical student at the core of the educational process. The tree-structured framework conceptualizes AI as a supportive tool that enhances, rather than replaces, human learning and professional development. At the core of the model lies the medical student, symbolizing the central role of the individual learner. From this central node, four principal branches extend, representing the foundational domains of medical education: knowledge, clinical skills, research, and professionalism. On the left branch, Generative AI-enabled applications, such as personalized learning systems, virtual patient simulations, automated content generation, and scientific writing support tools, are integrated into each domain. The right branch represents the systemic enablers required for responsible adoption, including institutional governance, educator training, AI literacy initiatives, and ethical oversight committees. Together, these structural supports form a stable foundation for sustainable and accountable Generative AI integration in medical education.

Generative AI should be regarded not as a substitute for educators, but as a collaborative ally that enhances teaching effectiveness. When deployed within a well-defined ethical and regulatory framework, Generative AI can significantly improve medical education. The correct and appropriate use of generative AI will ultimately help cultivate a new generation of medical professionals distinguished by clinical competence, ethical responsibility, and creative thinking.

Conclusion

Generative AI, including domain-specific models tailored for medical applications, is reshaping the landscape of medical education. Its technical attributes have brought new solutions to long-standing challenges in traditional medical education, offering new possibilities for personalized learning, clinical simulation, automated content generation, and research support.

Despite its transformative potential, the integration of Generative AI into medical education is not without risks. Issues such as hallucinations, data privacy risks, academic misconduct, and algorithmic biases highlight the need for careful considerations and responsible governance in its deployment.

To fully utilize the advantages of Generative AI while minimizing its risks, this review proposes a human-centered conceptual framework positioning AI as an auxiliary tool rather than replacement for medical educators. Achieving this vision requires robust institutional governance, comprehensive educator training, widespread AI literacy programs, and consistent ethical oversight to foster equitable and transparent adoption.

Looking ahead, collaboration among educators, technologists, and policymakers will be critical to realize the full potential of Generative AI while safeguarding the integrity and humanistic values that define medical education. Under ethical and governance frameworks, Generative AI holds the potential to cultivate a new generation of medical professionals distinguished by clinical competence, ethical awareness, and innovative thinking.

Author contributions: Wenhui Guo: Conceptualization, Methodology, Writing – Original Draft, Writing – Review & Editing. Bing Xu and Jiaojiao Feng: Investigation, Visualization. Zui Zou and Miao Zhou: Supervision, Writing – Review & Editing.

References

[1] McCarthy J, Minsky ML, Rochester N, et al. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. AI Magazine 1955;27:12.
[2] Tolsgaard MG, Pusic MV, Sebok-Syer SS, et al. The fundamentals of Artificial Intelligence in medical education research: AMEE Guide No. 156. Med Teach 2023;45(6):565-573.
[3] Kumar KJ, Jairam K, Ambedkar C, et al. An In-Depth Study Of Machine Learning In Artificial Intelligence. Educational Administration: Theory and Practice 2023;29(4):2401-2408.
[4] Yao JF, Huang YHJ, Yang CY, et al. Deep Learning Applications. 2024 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) 2024;1-4.
[5] Poland S, Sharad S, Wigen-Toccalino G, et al. Powering Innovation: The Comprehensive Impact of Generative AI. Climate and Energy 2024;40(7):1-11.
[6] Banh L, Strobel G. Generative artificial intelligence. Electronic Markets 2023;33:63.
[7] Kim Y, Son K, Kim S, et al. Beyond Prompts: Learning from Human Communication for Enhanced AI Intent Alignment. arXiv 2024;arXiv:2405.05678.

[8] Ouyang L, Wu J, Jiang X, et al. Training language models to follow instructions with human feedback. arXiv 2022; arXiv:2203.02155.

[9] García-Peñalvo FJ, Vázquez-Ingelmo A. What Do We Mean by GenAI? A Systematic Mapping of The Evolution, Trends, and Techniques Involved in Generative AI. International Journal of Interactive Multimedia and Artificial Intelligence 2023;8(4):7-16.

[10] Guettala M, Bourekkache S, Kazar O, et al. Generative Artificial Intelligence in Education: Advancing Adaptive and Personalized Learning. Acta Informatica Pragensia 2024;13(3):460-489.

[11] Zhao WX, Zhou K, Li J, et al. A Survey of Large Language Models. arXiv 2023;arXiv:2303.18223.

[12] Kaswan KS, Dhatterwal JS, Malik K, et al. Generative AI: A Review on Models and Applications. 2023 International Conference on Communication, Security and Artificial Intelligence (ICCSAI) 2023;699-704.

[13] OpenAI, Achiam J, Adler S, et al. GPT-4 Technical Report. arXiv 2023;arXiv:2303.08774.

[14] Team G, Anil R, Borgeaud S, et al. Gemini: A Family of Highly Capable Multimodal Models. arXiv 2023; arXiv:2312.11805.

[15] Yakymiv V, Piskozub Y. Research on the use of AI for Selecting Abstractions for Natural Language Image Generation Tools. International Journal of Computing 2024;23(4):637-654.

[16] Roumeliotis KI, Tselikas ND. ChatGPT and Open-AI Models: A Preliminary Review. Future Internet 2023;15(6):192.

[17] Xiong L, Wang H, Chen X, et al. DeepSeek: Paradigm Shifts and Technical Evolution in Large AI Models. IEEE/CAA Journal of Automatica Sinica 2025;12(5):841-858.

[18] Pal A, Sankarasubbu M. Gemini Goes to Med School: Exploring the Capabilities of Multimodal Large Language Models on Medical Challenge Problems & Hallucinations. Clinical Natural Language Processing Workshop 2024;21-46.

[19] Sun X, Tang W, Huang Z, et al. From general to specific: Tailoring large language models for real-world medical communications. Clin Transl Med 2025;15(1):e70157.

[20] Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. Nature 2023;620(7972):172-180.

[21] Saab K, Tu T, Weng WH, et al. Capabilities of Gemini Models in Medicine. arXiv 2024;arXiv:2404.18416.

[22] Zhang P, Shi J, Kamel Boulos MN. Generative AI in Medicine and Healthcare: Moving Beyond the 'Peak of Inflated Expectations'. Future Internet 2024;16(12):462.

[23] Lee J, Yoon W, Kim S, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics 2020;36(4):1234-1240.

[24] Buja LM. Contemporary Medical Education: Revolution versus Evolution. In: Firstenberg MS, Stawicki SP, editors. Medical Education for the 21st Century. Rijeka. IntechOpen, 2021.

[25] Coates WC. Precision education - a call to action to transform medical education. Int J Emerg Med 2025;18(1):21.

[26] Mitra NK, Chitra E. Glimpses of the Use of Generative AI and ChatGPT in Medical Education. Education in Medicine Journal 2024;16(2):155-164.

[27] Boscardin CK, Gin B, Golde PB, et al. ChatGPT and Generative Artificial Intelligence for Medical Education: Potential Impact and Opportunity. Acad Med 2024;99(1):22-27.

[28] Beaty KA, Chow JM, Cunha RLF, et al. Managing sensitive applications in the public cloud. IBM Journal of Research and Development 2016;60(2-3):4:1-4:13.

[29] Esposito C, Santis AD, Tortora G, et al. Blockchain: A Panacea for Healthcare Cloud-Based Data Security and Privacy? IEEE Cloud Computing 2018;5(1):31-37.

[30] Wei Y, Jiang YH, Liu J, et al. The Advancement of Personalized Learning Potentially Accelerated by Generative AI. arXiv 2024;arXiv:2412.00691.

[31] Maity S, Deroy AJA. Generative AI and Its Impact on Personalized Intelligent Tutoring Systems. arXiv 2024;arXiv:2410.10650.

[32] Sunmboye K, Strafford H, Noorestani S, et al. Exploring the influence of artificial intelligence integration on personalized learning: a cross-sectional study of undergraduate medical students in the United Kingdom. BMC Med Educ 2025;25(1):570.

[33] Hui Z, Zewu Z, Jiao H, et al. Application of ChatGPT-assisted problem-based learning teaching method in clinical medical education. BMC Med Educ 2025;25(1):50.

[34] Chance EA. The combined impact of AI and VR on interdisciplinary learning and patient safety in healthcare education: a narrative review. BMC Med Educ 2025;25(1):1039.

[35] Chu SN, Goodell AJ. Synthetic Patients: Simulating Difficult Conversations with Multimodal Generative AI for Medical Education. arXiv 2024;arXiv:2405.19941.

[36] Coşkun Ö, Kıyak YS, Budakoğlu I. ChatGPT to generate clinical vignettes for teaching and multiple-choice questions for assessment: A

randomized controlled experiment. Med Teach 2025;47(2):268-274.
[37] Almansour M, Soliman M, Aldekhyyel R, et al. An Academic Viewpoint (2025) on the Integration of Generative Artificial Intelligence in Medical Education: Transforming Learning and Practices. Cureus 2025;17(3):e81145.
[38] Bahroun Z, Anane C, Ahmed V, et al. Transforming Education: A Comprehensive Review of Generative Artificial Intelligence in Educational Settings through Bibliometric and Content Analysis. Sustainability 2023;15(17): 12983.
[39] Karabacak M, Ozkara BB, Margetis K, et al. The Advent of Generative Language Models in Medical Education. JMIR Med Educ 2023;9:e48163.
[40] Perkins M, Pregowska A. Artificial Intelligence in Medical Education: Technology and Ethical Risk. SSRN Electronic Journal 2023;17.
[41] Zhou J, Zhang J, Wan R, et al. Integrating AI into clinical education: evaluating general practice trainees' proficiency in distinguishing AI-generated hallucinations and impacting factors. BMC Med Educ 2025;25(1):406.
[42] Sallam M. ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare (Basel) 2023; 11(6):887.
[43] Hammane Z, Ben-Bouazza FE, Fennan A, et al. SelfRewardRAG: Enhancing Medical Reasoning with Retrieval-Augmented Generation and Self-Evaluation in Large Language Models. 2024 International Conference on Intelligent Systems and Computer Vision (ISCV) 2024;1-8.
[44] Thurzo A. How is AI Transforming Medical Research, Education and Practice? Bratislava Medical Journal 2025;126:243-248.
[45] Gawankar S, Nair S, Pawar V, et al. Patient Privacy and Data Security in the Era of AI-Driven Healthcare. 2024 8th International Conference on Computing, Communication, Control and Automation (ICCUBEA) 2024;1-6.
[46] Sallam M, Al-Mahzoum K, Sallam M. Generative Artificial Intelligence and Cybersecurity Risks: Implications for Healthcare Security Based on Real-life Incidents. Mesopotamian Journal of Artificial Intelligence in Healthcare 2024;2024:184-203.
[47] Abdullah HMA, Naeem NI, Malkana GA, et al. Exploring the Ethical Implications of ChatGPT in Medical Education: Privacy, Accuracy, and Professional Integrity in a Cross-Sectional Survey. Cureus 2024;16(12):e75895.
[48] Xu T, Weng H, Liu F, et al. Current Status of ChatGPT Use in Medical Education: Potentials, Challenges, and Strategies. J Med Internet Res 2024;26:e57896.
[49] Li J. A Privacy Preservation Model for Health-Related Social Networking Sites. J Med Internet Res 2015;17(7):e168.
[50] Jonnagaddala J, Wong ZS. Privacy preserving strategies for electronic health records in the era of large language models. NPJ Digit Med 2025;8(1):34.
[51] Ng MY, Helzer J, Pfeffer MA, et al. Development of Secure Infrastructure for Advancing Generative AI Research in Healthcare at an Academic Medical Center. Res Sq 2024; rs.3.rs-5095287.
[52] Ng MY, Helzer J, Pfeffer MA, et al. Development of secure infrastructure for advancing generative artificial intelligence research in healthcare at an academic medical center. J Am Med Inform Assoc 2025;32(3):586-588.
[53] Khatri B, Karki P. Artificial Intelligence (AI) in Higher Education: Growing Academic Integrity and Ethical Concerns. Nepalese Journal of Development and Rural Studies 2023;20(01): 1-7.
[54] Bittle K, El-Gayar O. Generative AI and Academic Integrity in Higher Education: A Systematic Review and Research Agenda. Information 2025;16(4):296.
[55] Liu J, Hui K, Al Zoubi F, et al. The great detectives: humans versus AI detectors in catching large language model-generated medical writing. Int J Educ Integr 2024;20:8.
[56] Jiang Z, Zhang J, Gong NZ. Evading Watermark based Detection of AI-Generated Content. Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security 2023;1168-1181.
[57] Francis NJ, Jones S, Smith DP. Generative AI in Higher Education: Balancing Innovation and Integrity. Br J Biomed Sci 2025;81:14048.
[58] Araujo T, Helberger N, Kruikemeier S, et al. In AI we trust? Perceptions about automated decision-making by artificial intelligence. AI & Soc 2020;35:611-623.
[59] Caliskan A, Bryson JJ, Narayanan A. Semantics derived automatically from language corpora contain human-like biases. Science 2017; 356(6334):183-186.
[60] Agarwal R, Bjarnadottir M, Rhue L, et al. Addressing algorithmic bias and the perpetuation of health inequities: An AI bias aware framework. Health Policy and Technology 2023;12(1):100702.

[61] Ibrahim H, Liu X, Zariffa N, et al. Health data poverty: an assailable barrier to equitable digital health care. Lancet Digit Health 2021; 3(4):e260-e265.
[62] Boateng O, Boateng B. Algorithmic bias in educational systems: Examining the impact of AI-driven decision making in modern education. World Journal of Advanced Research and Reviews 2025;25(01):2012-2017.
[63] Ali M. The Role of AI in Reshaping Medical Education: Opportunities and Challenges. Clin Teach 2025;22(2):e70040.