

Prospects and applications of artificial intelligence and large language models in obstetrics and gynecology education: a narrative review

Kyung Jin Eoh, MD, PhD

Department of Obstetrics and Gynecology, Yongin Severance Hospital, Yonsei University College of Medicine, Yongin, Korea

Purpose: This review examines how artificial intelligence (AI) and large language models (LLMs) can meet the diverse demands of obstetrics and gynecology education. Based on an exploration of their applications, benefits, and challenges, strategies are proposed for effectively integrating these emerging technologies into educational programs.

Current Concepts: Traditional obstetrics and gynecology education relies on lectures, hands-on training, and clinical exposure. However, these approaches often face limitations such as restricted practical opportunities and difficulties in remaining current with rapidly evolving medical knowledge. Recent AI advancements offer enhanced data analysis and problem-solving capabilities, while LLMs, through natural language processing, can supply timely, disease-specific information and facilitate simulation-based training. Despite these benefits, concerns persist regarding data bias, ethical considerations, privacy risks, and potential disparities in healthcare access.

Discussion and Conclusion: Although AI and LLMs hold promise for improving obstetrics and gynecology education by expanding access to current information and reinforcing clinical competencies, they also present drawbacks. Algorithmic transparency, data quality, and ethical use of patient information must be addressed to foster trust and effectiveness. Strengthening ethics education, developing Explainable AI, and establishing clear validation and regulatory frameworks are critical for minimizing risks such as over-diagnosis, bias, and inequitable resource distribution. When used responsibly, AI and LLMs can revolutionize obstetrics and gynecology education by enhancing teaching methods, promoting student engagement, and improving clinical preparedness.

Key Words: Artificial intelligence; Natural language processing; Obstetrics; Gynecology; Education

Introduction

Obstetrics and gynecology constitutes a particularly com-

plex and multifaceted medical specialty that requires a high level of expertise and practical proficiency [1]. Currently, most obstetrics and gynecology education involves didactic lectures, hands-on training, and clinical immersion, in which students primarily gain knowledge and experience through direct patient interactions [2]. However, these traditional teaching methods face notable challenges, including limited opportunities for practical application and the difficulty of keeping pace with rapidly advancing medical knowledge [3].

Received: February 9, 2025 **Accepted:** February 18, 2025

Corresponding author: Kyung Jin Eoh
E-mail: kjeoh2030@yuhs.ac

© Korean Medical Association

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

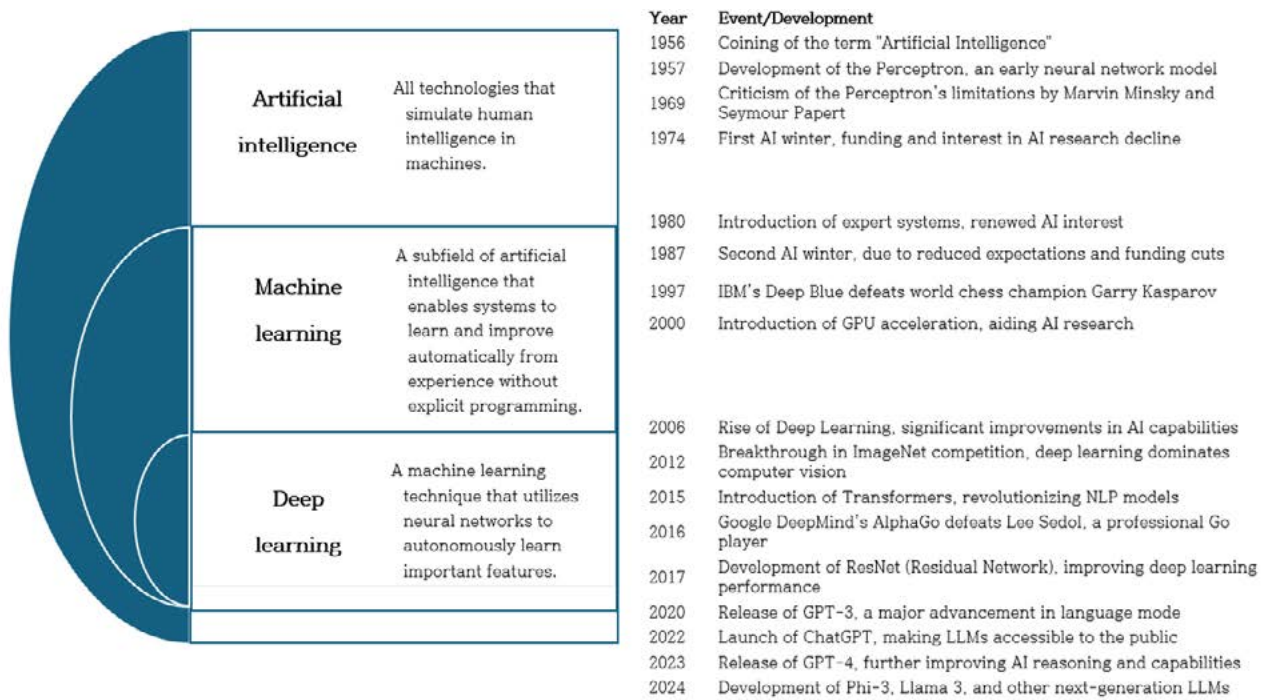


Figure 1. Hierarchy and chronological development of artificial intelligence, machine learning, and deep learning. AI, artificial intelligence; GPU, graphics processing unit; NLP, natural language processing; GPT, Generative Pre-trained Transformer. Illustration by the author.

To address these shortcomings, artificial intelligence (AI) has emerged as an innovative tool with the potential to transform obstetrics and gynecology education [4]. By offering powerful data analysis, problem-solving, and knowledge dissemination capabilities, AI can significantly improve the quality of education and the overall learning experience [5].

The Oxford English Dictionary defines AI as “the theory and development of computer systems capable of performing tasks that typically require human intelligence, such as visual and speech recognition, decision-making, and language translation” [6]. AI encompasses a range of techniques, including probabilistic machine learning, evolutionary computing, expert systems, and fuzzy systems. Its origins date back to 1943, when McCulloch and Pitts proposed an electrical circuit model mimicking the brain's neural networks [7]. The term “artificial intelligence” was formally adopted at the 1956 Dartmouth Conference, paving the way for subsequent breakthroughs such as neural network models, backpropa-

gation, multilayer architectures, and deep learning technologies, all of which have profoundly influenced numerous fields (Figure 1) [8,9].

In the medical domain, AI is mainly applied to improve patient diagnosis, develop treatment plans, and promote patient safety, with significant potential to reduce diagnostic errors and enhance healthcare quality [10]. It is especially valuable for meeting growing healthcare demands driven by aging populations and shortages of medical personnel [11]. Within obstetrics and gynecology education, integrating AI can help mitigate students' anxieties and uncertainties while creating an opportunity to reshape medical education [4,12].

At this critical juncture, it is essential to understand both the current progress and the limitations of AI-based medical technologies, as well as to develop a systematic educational framework grounded in that knowledge. This narrative review aims to assess the transformative effects of AI on medical fields, including obstetrics and gynecology, and to investigate how these innovations can be effectively integrat-

Table 1. Summary of large language models in various domains

LLMs	Developer	Summary/key applications	Primary domain	Year of development
ChatGPT	OpenAI	NLP and text generation, widely used for chatbots, content creation, and code generation	General AI, NLP	2022
GPT 3, 3.5, 4	OpenAI	Advanced versions of ChatGPT, providing improved contextual understanding and text generation	General AI, NLP	2020
Gemini	Google	Google's LLM, focused on search enhancement and conversational AI	General AI, NLP	2023
DeepSeek	DeepSeek	Chinese-developed AI assistant with DeepThink capabilities and enhanced contextual understanding	General AI, NLP	2023
BioGPT	Luo et al.	Biomedical text generation and literature mining, specialized for clinical data processing	Biomedical AI	2022
BioBERT	Lee et al.	Fine-tuned BERT for biomedical text generation and literature extraction	Biomedical AI	2019
BioMedLM	Stanford (Bolton et al.)	Biomedical LLM from Stanford, optimized for medical text analysis	Biomedical AI	2023
ChatCAD	Wang et al.	AI-driven diagnosis system analyzing medical images for clinical support	Medical imaging	2022
ClinicalBERT	Huang et al.	Clinical text processing model trained on electronic health records	Clinical data	2019
BlueBERT	Peng et al.	Biomedical BERT variant trained on PubMed abstracts and clinical notes	Biomedical AI	2019
Clinical Camel	Toma et al.	Llama-based model fine-tuned for biomedical research applications	Biomedical AI	2023
BioMegatron	Shin et al.	BERT variant trained specifically on PubMed abstracts and biomedical datasets	Biomedical AI	2020
BioMedRoBERTa	Gururangan et al.	Fine-tuned RoBERTa model optimized for biomedical research applications	Biomedical AI	2020
MedAlpaca	Han et al.	Alpaca-based model tailored for medical research and data processing	Biomedical AI	2023
Med-PaLM	Google (Singhal et al.)	Google's medical AI model trained on diverse biomedical literature and research data	Biomedical AI	2023

LLM, large language model; GPT, Generative Pre-trained Transformer; AI, artificial intelligence; NLP, natural language processing; BERT, Bidirectional Encoder Representations from Transformers; BioMedLM, Biomedical Language Model; ChatCAD, Chat-based Clinical AI Diagnosis; BioMedRoBERTa, Biomedical Robustly Optimized BERT Pretraining Approach; Med-PaLM, Medical Pathways Language Model.

ed into medical education.

This study did not involve human subjects; thus, neither institutional review board approval nor informed consent was required.

Overview of Large Language Models and Their Applications in Obstetrics and Gynecology Education

Large language models (LLMs) are a specialized branch of AI capable of understanding and generating natural language by training on extensive datasets. Notable examples include OpenAI's Generative Pre-trained Transformer-4 (GPT-4) and Google's Gemini, which excel at analyzing large volumes of text and producing new information based on that analysis. In medicine, LLMs are used in a variety of ways, such as analyzing patient records and supporting medical consultations (Table 1) [13,14].

The potential applications of LLMs in obstetrics and gynecology education are extensive. For example, LLMs can

quickly provide relevant information when students inquire about specific conditions or can serve as simulation tools to replicate complex diagnostic processes. These uses can create a more effective learning environment, helping bridge the gap between theoretical knowledge and practical skills [15].

As machine learning becomes increasingly prevalent in future healthcare systems, the healthcare delivery landscape is expected to change significantly. Wartman and Combs [16] identified key characteristics of future healthcare, including the creation of environments that facilitate care regardless of patient location, team-based care involving diverse providers, the use of multiple information sources (including AI), and the efficient management of AI models in clinical settings. Adapting to these changes will require healthcare professionals to develop new competencies.

In addition, the World Economic Forum's 4C framework—critical thinking, communication, collaboration, and creativity—represents essential competencies for future healthcare practitioners, underscoring the need for pedagog-

Table 2. Potential roles of large language models in clinical training and knowledge acquisition in obstetrics and gynecology

Role	Specific applications	Expected benefits
Complex medical data analysis	Providing possible diagnoses and treatment options based on patient history and test results	Aiding in accurate diagnosis and treatment planning
Clinical process simulation	Interacting with virtual patients to practice diagnosis and treatment	Enhancing communication skills and adaptability in real clinical settings
Enhancing problem-solving skills	Presenting complex medical cases and guiding step-by-step problem resolution	Improving clinical reasoning and comprehensive problem-solving abilities
Personalized learning and feedback	Offering supplementary explanations tailored to the learner's level and providing detailed feedback on incorrect answers	Increasing learning efficiency and boosting confidence
Improved access to academic resources	Translating and summarizing the latest research and global medical knowledge	Creating a learning environment based on the most up-to-date information

ical reform in medical education [17]. A survey of medical students in Busan found that 97% considered AI education indispensable, while a separate survey at Stanford Medical School reported that 73% of students recognized the need for additional training to prepare for medical innovation. Nevertheless, only 18% of students and residents felt that current education adequately equips them for evolving medical technologies, emphasizing the urgent need to improve educational programs [17].

Prominent organizations such as the National Academy of Medicine, the American Medical Association, and the United Kingdom's National Health Service have identified essential competencies for medical education. These include the ability to extract and synthesize relevant information from a rapidly expanding body of medical knowledge, engage in self-directed lifelong learning, understand machine learning principles while recognizing their limitations, deliver precision medicine through data utilization, choose and implement suitable machine learning models, collaborate with diverse healthcare stakeholders, and empathize and communicate effectively with patients [18].

These competencies are especially important in obstetrics and gynecology education, where LLMs can streamline students' and residents' access to complex medical information, thus improving the overall learning experience. It is therefore essential to investigate methods for incorporating these technologies into educational programs.

Potential Roles of LLMs in Clinical Training and Knowledge Acquisition

LLMs have considerable potential to play key roles in medical education, particularly in clinical training and knowledge acquisition. Beyond basic data analysis, these models provide innovative, practical tools for learners and practitioners alike. The following sections detail potential applications of LLMs (Table 2).

1. Support for analysis and interpretation of complex medical data

LLMs can be instrumental in analyzing and interpreting large amounts of medical data [19]. For example, when students or residents enter a patient's medical history, lab results, and imaging data, LLMs can suggest possible diagnoses or recommend various treatment options. In cases involving rare diseases or complicated presentations, LLMs can deliver accurate, evidence-based information from the most recent literature, helping learners make well-informed clinical decisions.

2. Simulation tools for modeling clinical processes

LLMs can function as simulation tools, allowing students to virtually practice different scenarios they may encounter in clinical settings [20,21]. For instance, learners can engage in conversations with an LLM that simulates patient interactions, enabling them to identify symptoms, request appro-

Table 3. Ethical and practical limitations of AI and LLM utilization

Issue area	Details
Data bias and reliability of outcomes	- Utilization of large datasets, but potential for biased learning - Requirement for diverse and high-quality data
Patient safety and medical ethics	- Potential for incorrect AI recommendations - Need for a critical evaluation system for AI-generated results - Importance of medical ethics education for students
Technical limitations of deep learning	- Possibility of errors, lack of representativeness, and bias issues - "Black box" nature of decision-making processes - Need for the development of XAI
Privacy and data security	- Necessity for data anonymization - Implementation of enhanced data protection protocols
Over-diagnosis and healthcare disparities	- Risk of increased unnecessary treatments - Potential rise in medical costs - Efforts required to ensure fair access to AI technologies

AI, artificial intelligence; LLM, large language model; XAI, Explainable AI.

priate tests, or develop treatment plans. This form of training goes beyond rote memorization by cultivating the comprehensive reasoning and communication skills essential in real clinical practice.

3. Enhancement of problem-solving skills and clinical reasoning

LLMs can also support the development of learners' problem-solving skills [22]. By presenting complex medical cases and guiding students through the resolution process, these models offer hands-on opportunities to refine diagnostic reasoning. For instance, an LLM may pose challenging diagnostic puzzles, leading learners step by step toward the best possible conclusion. This approach encourages systematic thinking and better prepares students for real-world clinical demands. In a recent study comparing GPT model performance with that of obstetrics and gynecology residents, no significant difference in exam scores was found (GPT-4 mean, 79.3 vs. resident mean, 82.1; $P=0.246$), with particular strengths observed in problem-solving and data interpretation tasks [4]. As LLMs evolve, their range of applications and effectiveness is likely to expand further.

4. Provision of personalized learning and feedback

LLMs can adapt learning experiences to individual learn-

ers' needs and abilities [23]. If a student demonstrates gaps in understanding a certain topic, LLMs can supply additional explanations and resources or deliver detailed, constructive feedback on incorrect answers. This personalized approach can improve knowledge retention and increase learners' confidence.

5. Improved accessibility to diverse academic resources

LLMs can also improve learners' access to cutting-edge research and authoritative academic sources [24]. By automating time-consuming tasks such as literature searches and interpretations, these models allow students to focus on more in-depth study and training. For learners who do not speak English, LLMs can help overcome language barriers and provide entry to a wealth of international medical knowledge [25].

Ethical and Practical Limitations of AI and LLM Utilization and Proposed Solutions

Although AI and LLMs offer promising opportunities in both medical and educational settings, their implementation also raises various ethical and practical challenges. Understanding and proactively managing these issues is essential for the safe and effective use of AI and LLMs in obstetrics and gynecology education (Table 3).

1. Data bias and reliability of outcomes

AI and LLMs are trained on extensive datasets, but any biases inherent in these datasets—resulting from particular populations or contexts—can lead to inaccurate outcomes [26]. Such biases increase the risk of incorrect information

being applied in medical diagnosis and treatment. For example, a model trained only on data from one region or device might fail to reflect the diverse healthcare needs of various patient groups. Addressing this issue requires improving the diversity and quality of training data and regularly validating model performance.

2. Patient safety and medical ethics

Incorrect AI recommendations can have serious repercussions for patient care [27]. In obstetrics and gynecology, misjudging fetal conditions or providing inappropriate treatment suggestions can jeopardize both maternal and fetal well-being. Thus, systems that allow clinicians to critically evaluate and confirm AI outputs are essential. Moreover, medical ethics education must be integrated to ensure that students understand the ethical implications of using AI in clinical practice.

3. Technical limitations of deep learning

Deep learning models rely heavily on data quality, and those based on medical data may contain errors, lack representativeness, or exhibit bias [28]. For instance, a model trained under specific obstetric and gynecologic conditions may not maintain the same accuracy in different settings or with different equipment. Addressing this challenge demands higher-quality data and research to validate model performance across varied clinical environments. In addition, the opaque “black box” nature of deep learning can make it difficult to explain AI-generated decisions [29]. Healthcare professionals need insight into the reasoning behind AI outputs to establish trust. Ongoing efforts to develop Explainable AI (XAI) are crucial, and obstetrics and gynecology students must learn to interpret these technical constraints critically.

4. Privacy and data security

Because AI and LLMs rely on large datasets, privacy

and data security have become paramount concerns [27]. Obstetrics and gynecology data often involve sensitive information such as birth records, genetic conditions, and reproductive health, making robust safeguards essential. Failure to properly anonymize or securely transmit data can lead to breaches. Collaboration between hospitals and developers is necessary to strengthen data protection policies and security measures in order to forestall such vulnerabilities.

5. Over-diagnosis and healthcare disparities

Although AI can excel at early disease detection, it may also increase over-diagnosis, leading to unnecessary interventions [30]. Such overuse can inflate healthcare costs and cause undue patient anxiety. In obstetrics and gynecology, it is crucial to apply AI judiciously to avoid excessive interventions. Additionally, disparities in AI access and usage across different healthcare settings can exacerbate inequalities [31]. If AI remains limited to certain institutions or populations, health equity may suffer. Collaboration among governments and institutions is needed to promote fair access to AI technologies.

6. Proposed solutions

Several strategies can help overcome the limitations of AI and LLMs. First, improving data collection methods to ensure a diverse representation of patient populations can enhance model accuracy. Second, developing XAI can make decision-making processes transparent to healthcare professionals. Third, systematically incorporating medical ethics education will better prepare students to address AI-related ethical dilemmas. Finally, continuous validation and regulatory oversight are necessary to assess AI model performance and govern their use in medical settings. A comprehensive understanding of these challenges is crucial for effectively integrating AI and LLMs into obstetrics and gynecology education, ensuring they serve as safe and reliable tools for both

learning and clinical practice.

The Future of Obstetrics and Gynecology Education: Prospects and Directions for LLM Utilization

LLMs represent a transformative resource that can significantly reshape obstetrics and gynecology education, offering a promising future [4,15]. Beyond simply providing study materials, these technologies can increase the quality of education by engaging learners and simulating real clinical scenarios.

As LLMs become increasingly sophisticated, students will be able to simulate more complex and varied clinical situations [20,21]. For example, virtual practice in managing high-risk pregnancies, performing emergency cesarean sections, or addressing hemorrhage can significantly improve students' clinical reasoning and decision-making skills. Such personalized and precise learning experiences mark a substantial leap beyond traditional methods. Additionally, LLMs offer a high degree of personalization in obstetrics and gynecology education [23]. By adjusting content according to each learner's proficiency, interests, and learning speed, these models can maintain engagement and enhance educational outcomes. For instance, beginners might focus on foundational concepts and procedures, while advanced learners tackle complex case-based scenarios to sharpen their problem-solving abilities.

LLMs also have the capacity to expand educational accessibility in diverse multinational and multicultural contexts [25]. By providing translated materials and culturally appropriate resources, they enable high-quality education for medical students and practitioners around the globe.

This technological evolution signals a paradigm shift that will affect not only obstetrics and gynecology education but

also the broader realm of medical instruction. More than just tools, LLMs can promote collaboration among students, residents, and faculty, ultimately improving patient care [14]. In the future, leveraging LLMs will be crucial for advancing the efficiency and effectiveness of obstetrics and gynecology education and setting new standards for medical training.

AI and LLMs hold transformative potential for obstetrics and gynecology education by improving access to the latest knowledge and refining clinical competencies. Nevertheless, issues such as data bias, ethical concerns, and privacy risks call for cautious and thoughtful implementation. Incorporating ethics education, advancing XAI, and instituting strong validation frameworks are crucial to minimizing over-diagnosis and preventing inequalities in healthcare. Used responsibly, these technologies can reshape the teaching environment, enhance student engagement, and elevate clinical readiness, ultimately defining new benchmarks for medical education in this dynamic field.

ORCID

Kyung Jin Eoh, <https://orcid.org/0000-0002-1684-2267>

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

References

1. Fialkow MF, Snead CM, Schulkin J. Benefits and barriers to teaching medical students in an Ob-Gyn clinic. *Health Serv Res Manag Epidemiol* 2018;5:2333392817753518.
2. Duhm L, Wittek A, Walter A, et al. Factors influencing medical students' interest in obstetrics and gynecology: a longitudinal study on career preferences. *Arch Gynecol Obstet* 2025; 311:903-914.
3. Metheny WP, Ernest JM, Bachicha J, et al. Have we met the educational challenges of obstetrics and gynecology?: a response to the Association of Professors of Gynecology and Obstetrics Initiative of 1986. *Am J Obstet Gynecol* 2002;187: 1405-1412.
4. Eoh KJ, Kwon GY, Lee EJ, et al. Efficacy of large language models and their potential in obstetrics and gynecology edu-

- cation. *Obstet Gynecol Sci* 2024;67:550-556.
5. McAdams RM, Green TL. Equitable artificial intelligence in obstetrics, maternal-fetal medicine, and neonatology. *Obstet Gynecol* 2024 Mar 28 [Epub]. <https://doi.org/10.1097/aog.0000000000005563>
6. Abraham TH. (Physio)logical circuits: the intellectual origins of the McCulloch-Pitts neural networks. *J Hist Behav Sci* 2002;38:3-25.
7. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. 1943. *Bull Math Biol* 1990;52:99-115.
8. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature* 1986;323:533-536.
9. Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. *Neural Comput* 2006;18:1527-1554.
10. Kolokythas A, Dahan MH. Is Artificial Intelligence (AI) currently able to provide evidence-based scientific responses on methods that can improve the outcomes of embryo transfers?: no. *JBRA Assist Reprod* 2024;28:629-638.
11. Mapari SA, Shrivastava D, Dave A, et al. Revolutionizing Maternal health: the role of artificial intelligence in enhancing care and accessibility. *Cureus* 2024;16:e69555.
12. Lee Y, Kim SY. Potential applications of ChatGPT in obstetrics and gynecology in Korea: a review article. *Obstet Gynecol Sci* 2024;67:153-159.
13. Bedi S, Liu Y, Orr-Ewing L, et al. Testing and evaluation of health care applications of large language models: a systematic review. *JAMA* 2025;333:319-328.
14. Thirunavukarasu AJ, Ting DS, Elangovan K, Gutierrez L, Tan TE, Ting DS. Large language models in medicine. *Nat Med* 2023;29:1930-1940.
15. Reicher L, Lutsker G, Michaan N, Grisaru D, Laskov I. Exploring the role of artificial intelligence, large language models: Comparing patient-focused information and clinical decision support capabilities to the gynecologic oncology guidelines. *Int J Gynaecol Obstet* 2025;168:419-427.
16. Wartman SA, Combs CD. Medical education must move from the information age to the age of artificial intelligence. *Acad Med* 2018;93:1107-1109.
17. Jung JS. Current status and future direction of artificial intelligence in healthcare and medical education. *Korean Med Educ Rev* 2020;22:99-114.
18. Institute of Medicine, Board on Health Care Services, Committee on the Health Professions Education Summit, Knebel E, Greiner AC. Health professions education: a bridge to quality. National Academies Press; 2003.
19. Safranek CW, Sidamon-Eristoff AE, Gilson A, Chartash D. The role of large language models in medical education: applications and implications. *JMIR Med Educ* 2023;9:e50945.
20. Cook DA. Creating virtual patients using large language models: scalable, global, and low cost. *Med Teach* 2025;47:40-42.
21. Holderried F, Stegemann-Philipps C, Herrmann-Werner A, et al. A language model-powered simulated patient with automated feedback for history taking: prospective study. *JMIR Med Educ* 2024;10:e59213.
22. Goh E, Gallo R, Hom J, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. *JAMA Netw Open* 2024;7:e2440969.
23. Lucas HC, Upperman JS, Robinson JR. A systematic review of large language models and their implications in medical education. *Med Educ* 2024;58:1276-1285.
24. Mondal H, De R, Mondal S, Juhi A. A large language model in solving primary healthcare issues: a potential implication for remote healthcare and medical education. *J Educ Health Promot* 2024;13:362.
25. Shah K, Xu AY, Sharma Y, et al. Large language model prompting techniques for advancement in clinical medicine. *J Clin Med* 2024;13:5101.
26. Feehan M, Owen LA, McKinnon IM, DeAngelis MM. Artificial intelligence, heuristic biases, and the optimization of health outcomes: cautionary optimism. *J Clin Med* 2021;10:5284.
27. Chaurasia A, Curry G, Zhao Y, et al. Use of artificial intelligence in obstetric and gynaecological diagnostics: a protocol for a systematic review and meta-analysis. *BMJ Open* 2024;14:e082287.
28. Thakur GK, Thakur A, Kulkarni S, Khan N, Khan S. Deep learning approaches for medical image analysis and diagnosis. *Cureus* 2024;16:e59507.
29. Amann J, Blasimme A, Vayena E, Frey D, Madai VI; Precise4Q consortium. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak* 2020;20:310.
30. Komorowski M, Celi LA. Will artificial intelligence contribute to overuse in healthcare? *Crit Care Med* 2017;45:912-913.
31. Smith MJ, Axler R, Bean S, Rudzicz F, Shaw J. Four equity considerations for the use of artificial intelligence in public health. *Bull World Health Organ* 2020;98:290-292.