Review Article
Potential applications of ChatGPT in obstetrics and gynecology in Korea: a review article

YooKyung Lee, MD, PhD, So Yun Kim, DEGREE
Department of Obstetrics and Gynecology, Division of Maternal Fetal Medicine, MizMedi Hospital, Seoul, Korea

Running title:

Received: 2023.09.21.
Revised: 2023.11.08.
Accepted: 2023.11.29.

Corresponding author: YooKyung Lee, MD, PhD
Department of Obstetrics and Gynecology, Division of Maternal Fetal Medicine, MizMedi Hospital,
295 Gangseo-ro, Gangseo-gu, Seoul 07639, Korea
E-mail: asdfg2@naver.com
https://orcid.org/0009-0006-9560-7156
Abstract

The use of chatbot technology, particularly chat generative pre-trained transformer (ChatGPT) with an impressive 175 billion parameters, has garnered significant attention across various domains, including Obstetrics and Gynecology (OBGYN). This comprehensive review delves into the transformative potential of chatbots with a special focus on ChatGPT as a leading artificial intelligence (AI) technology. Moreover, ChatGPT harnesses the power of deep learning algorithms to generate responses that closely mimic human language, opening up myriad applications in medicine, research, and education. In the field of medicine, ChatGPT plays a pivotal role in diagnosis, treatment, and personalized patient education. Notably, the technology has demonstrated remarkable capabilities, surpassing human performance in OBGYN examinations, and delivering highly accurate diagnoses. However, challenges remain, including the need to verify the accuracy of the information and address the ethical considerations and limitations. In the wide scope of chatbot technology, AI systems play a vital role in healthcare processes, including documentation, diagnosis, research, and education. Although promising, the limitations and occasional inaccuracies require validation by healthcare professionals. This review also examined global chatbot adoption in healthcare, emphasizing the need for user awareness to ensure patient safety. Chatbot technology holds great promise in OBGYN and medicine, offering innovative solutions while necessitating responsible integration to ensure patient care and safety.

Keywords: Chatbot; ChatGPT; Medical field; Obstetrics and Gynecology; Artificial intelligence
Introduction

Chatbot technology has surged in popularity, expanding into various sectors such as customer services and personal virtual assistants [1]. However, the most noteworthy expansion of the aforementioned technology occurs within the specialized fields of Obstetrics and Gynecology (OBGYN), where chatbot displays considerable promise. Chatbot technology comprises two essential components: a versatile artificial intelligence (AI) system with broad capabilities and an intuitive chat interface [2]. This combination enables interactive sessions in which users initiate conversations through prompts or queries, and the chatbot responds, emulating human conversation dynamics [2].

AI is revolutionizing healthcare, promising improvements in quality, efficiency, and accessibility. A standout AI technology in this context is the chat generative pre-trained transformer (ChatGPT), a robust Large Language Model (LLM) renowned for generating high-quality text in response to various prompts. ChatGPT is invaluable for medical professionals in practice, research, and education because it has versatile applications in the medical field. Although continuing to evolve, ChatGPT has demonstrated impressive healthcare capabilities.

For instance, ChatGPT has outperformed human candidates in OBGYN examinations [3] and can generate concise lists of diagnoses that closely resemble those produced by clinicians. Despite these potential advantages, the use of ChatGPT in medicine presents several challenges and ethical considerations. One challenge is the accuracy and reliability of the information generated for all topics. Additionally, thoroughly assessing the ChatGPT output and corroborating its accuracy with
other information sources is necessary.

This study aims to delve into general AI technologies and their applications within OBGYN, exploring both their potential and associated challenges.

**General chatbot technology**

Chatbot technology is becoming increasingly prevalent, with applications across various domains, ranging from customer services to personal virtual assistants [1]. The early exploration of medical chatbots dates back to 1964-1966 when Joseph developed the ELIZA [1]. Today, advancements in computing power have empowered language models with hundreds of billions of parameters, allowing them to proficiently generate new text. A prime example of this capability coupled with access to extensive training data is ChatGPT [1].

ChatGPT, introduced by OpenAI in November 2022, marks a significant milestone in AI, serving as a LLM with an impressive 175 billion parameters. The technology harnesses deep-learning algorithms to generate human-like responses [4] and is trained on a massive corpus of text data using unsupervised learning techniques [5]. In collaboration with Microsoft, OpenAI has been a leading force in the development of progressively potent AI systems, with GPT-4 being the most advanced publicly available AI system as of March 2023 [2]. A chatbot consists of two key components: a general-purpose AI system and a chat interface [2]. Engaging with a chatbot typically
involves initiating a session by entering a query, commonly referred to as a prompt, after which the chatbot responds. This iterative exchange of prompts and responses simulates a conversation similar to that between two individuals [2].

**ChatGPT in medical education and research**

Over the past 6 months, Microsoft Research, in collaboration with OpenAI, has explored the potential applications of the GPT-4 in healthcare and medical settings. Their objective was to gain a deep understanding of the capabilities, limitations, and potential risks to human health [2]. The specific focus areas include medical and healthcare documentation, data interoperability, diagnosis, research, and education [2]. The latest generation of chatbots exhibits promise in addressing medical documentation challenges and critical questions pertinent to differential diagnoses. However, the validation of the accuracy of these responses remains challenging. The use of chatbots for making diagnoses or recommending treatment options is of particular interest [1]. GPT-4 was not exclusively programmed for a specific "assigned task", such as image analysis or medical note interpretation. Instead, they were designed to possess general cognitive abilities aimed at assisting users in a wide array of tasks. Prompts can manifest in diverse forms, ranging from questions to directives, guiding the model to execute particular actions. These prompts can be composed of multiple human languages and may include data inputs, such as spreadsheets, technical specifications, research
papers, and mathematical equations [2]. Despite the limitations in research and healthcare applications, ChatGPT can serve as a "clinical assistant" and significantly aid research and scholarly writing [6].

Chatbots utilizing GPT have the potential to automate specific tasks and enhance the efficiency of the writing process. They extract information from electronic medical records, assist in literature searches, and provide guidance on writing styles and formatting [7]. Medical writers can then review and edit the generated text to ensure precision and clarity. Furthermore, chatbots can streamline the review and editing processes, enabling multiple reviewers to offer real-time feedback and suggestions, thereby facilitating efficient collaboration. These technologies have the potential to accelerate the speed and accuracy of document creation in medical writing [7]. Several countries are currently conducting extensive testing to determine how GPT can be used in medicine. The development of ChatGPT has significantly improved the capabilities of natural language processing models in addressing medical questions [4].

For instance, Gilson et al. [4] identified that their model achieved a score equivalent to a passing score for 3-year medical students. These findings underscore the potential of ChatGPT as an interactive medical education tool to facilitate learning [4]. Additionally, a comparison of test scores between previous and recent models, such as GPT-4 passing the 117th Japanese Medical Licensing Examination while GPT-3.5 failed, reveals the rapid evolution of GPT-4 in Japanese language processing [8].

Conducting additional research and advancing the field of dentistry to effectively utilize these
potential advantages is essential [9]. Although ChatGPT serves as a valuable tool in medical education, research, and clinical management, the technology cannot be perceived as a replacement for human expertise or knowledge. Embracing these innovations in an open mindset can effectively enhance medical education and clinical management. [10].

As various studies on ChatGPT are ongoing worldwide, crossover randomized controlled trials have recently been conducted in Canada [11]. Participants were allocated to ChatGPT (group A) or standard online tools (group B) in a 1:1 ratio using computer-generated randomization. The alternative hypothesis posits a significant difference in learning outcomes and technology usability in accessing resources and completing assignments. The results are expected to identify critical areas that require attention and help educators develop a deep understanding of AI’s impact in the educational field. By exploring the usability and efficacy of ChatGPT compared with standard online tools, this study seeks to inform educators and students about the responsible integration of AI into academic settings.

The guidelines for utilizing ChatGPT were presented in a college of business [12] highlighting its potential to assist college students, educators, and professionals with writing, communication, and learning. Moreover, the guidelines provided instructions on crafting effective prompts to ensure the responsible and ethical use of ChatGPT.

When using ChatGPT, we can obtain specific and helpful answers using specific, clear, unbiased language, and by providing context. Sequential practical application methods are presented so that they can be applied in practice.
ChatGPT in OBGYN: diagnostic advancements

Although clinical validation of AI publications in OBGYN is currently lacking, recent studies have highlighted the potential utility of ChatGPT [13]. ChatGPT has demonstrated superior performance compared with highly trained human candidates in OBGYN examinations [3]. In a study published in the American Journal of Obstetrics and Gynecology 2023, ChatGPT responses to a spectrum of questions posed by four physicians regarding OBGYN were assessed [14]. These responses, presented in quotations, were commented upon by physicians [14]. Suhag et al. [13] suggested that ChatGPT can assist in formulating differential diagnoses and guide patients and healthcare teams in managing rare prenatal conditions.

This innovative approach may benefit perinatologists, geneticists, and neonatologists. ChatGPT-4 has generated abbreviated lists of diagnoses that closely align with those generated by clinicians, demonstrating improved differential diagnostic capabilities compared to the Human Phenotype Ontology phenomizer and online Mendelian inheritance in man. However, ChatGPT-4 may have achieved the correct diagnosis because of the fetal and newborn phenotype that is indicative of a known genetic syndrome [13]. Nevertheless, acknowledging that chatbots such as ChatGPT have limitations, including misinformation, frequency bias, exclusion of rare diagnoses, and variability in responses based on phrasing is crucial [13]. Martin et al. explored patient inquiries in ChatGPT regarding elevated blood pressure during pregnancy, showcasing its potential to provide patients with answers to their queries.
Santo and Joviano-Santos [15] investigated the use of ChatGPT, a LLM, for guidance during unexpected labor. They discovered that ChatGPT could be a valuable tool for providing laypeople with simple and easy-to-understand information in emergency labor situations. Due to the simplicity and easy access, ChatGPT can help laypeople clarify any questions or concerns they may have and guide them through the labor process. However, it is important to note that ChatGPT does not replace medical advice or assistance.

Allahqoli et al. [16] reported a series of cases of OBGYN that were presented to ChatGPT, and their diagnostic and management performances were compared with those of a gynecologist. The results of 30 cases were provided, including early pregnancy, general gynecology, emergency gynecology obstetrics, peripartum care, obstetric emergencies, family planning, and sexual health, adapted from the book titled ‘100 cases in Obstetrics and Gynecology’. The cases were submitted to ChatGPT and the following questions were entered at the end of the introduction of each case: differential diagnosis, cause of the condition, interpretation of these findings, further investigation, and management. The accuracy of ChatGPT in diagnosing and managing cases was 90% (27 of 30 cases). The diagnosis and management offered by ChatGPT are generally articulate, well-informed, and free from a significant degree of error or misinformation. Even in instances of an incorrect diagnosis, the responses provided by ChatGPT included a logical explanation of the case and the information contained in the question stem. In instances where ChatGPT initially failed to determine the correct diagnosis in three cases, its performance was improved through additional explanations. Upon reassessment, its secondary diagnosis precisely aligned with the gynecologist’s diagnosis.
Arslan [17] proposed the use of ChatGPT for personalized obesity treatment. This technology can assess a patient’s medical history, physical characteristics, and lifestyle to offer tailored advice on topics such as nutrition plans, exercise programs, and psychological support. ChatGPT also follows a patient’s progress over time and amends recommendations accordingly. This personalized approach could lead to more effective weight management and a reduction in associated health risks. Another potential use of ChatGPT is in the development of predictive models for obesity-related diseases such as diabetes and cardiovascular diseases.

**ChatGPT in medical imaging**

Recently, in the field of radiology, ChatGPT has been evaluated as a support tool for decision-making in the diagnosis of breast tumors [18]. Ten consecutive patients presented to the breast tumor board. The chatbot was asked to recommend management and the results generated by ChatGPT were compared with the final recommendations of the tumor board. Recommendations were independently graded by two senior radiologists. In seven of 10 cases (70%), ChatGPT’s recommendations were similar to those of the tumor board.

Jeblick et al. [19] reported that most radiologists expressed the opinion that simplified reports were accurate and complete, with no potential harm to patients. The initial results of this study suggest that ChatGPT has great potential to improve patient-centered care in radiology.
AI advancements in Korean healthcare

AI technology has been used since 2019 to assist medical professionals in addressing patient anxiety, confusion, and questions. The integration of AI into patient care is paramount, necessitating medical students to develop the ability to discern accurate information [20]. In Korea, parasitology examinations were administered to both ChatGPT and medical students to compare their knowledge and interpretation abilities [21]. The performance of ChatGPT on this examination did not align with that of the medical students [21]. In the field of surgery in Korea, GPT-4 has demonstrated a remarkable ability to comprehend complex surgical clinical information, achieving an accuracy rate of 76.4% on the Korean general surgery board exam [22]. Recent developments in Korea include the potential adoption of a newly developed database-based "Mobile Application Chatbot" for perinatal women's and their partners’ obstetric and mental health care. Text mining techniques and contextual usability testing were used to provide a convenient and pleasant user experience [23].

Implications for practice and limitations

As of January 2023, ChatGPT garnered an estimated 100 million active monthly users, making it the fastest-growing consumer application in history [24]. The rapid diffusion of generative AI tools, such as ChatGPT, has significant implications for practices and policy [24]. Although ChatGPT often
provides accurate responses, occasional lapses in understanding the context of questions may lead to misleading information. Variations in the responses based on prompts and updates can result in inappropriate answers for certain users [14]. Healthcare professionals using ChatGPT must be aware of the limitations of the application to avoid potential harm to patients. Furthermore, exercising caution is essential as ChatGPT may occasionally produce incorrect responses, a phenomenon known as "hallucinations". Fortunately, GPT-4 is proficient in identifying such mistakes, thereby enhancing its reliability [2].

ChatGPT can serve as a valuable tool for healthcare professionals to refine differential diagnoses while upholding patient privacy and confidentiality. Clinicians must continually educate themselves about cutting-edge technologies such as ChatGPT and explore ways to tailor their implementation in healthcare settings for optimal outcomes [13]. Chatbots hold promise as essential tools in medical practice; however, their effective use requires responsible and informed adoption. Both the users and tools undergo a learning process as they adapt to each other [1].

ChatGPT will be leveraged more extensively in the following areas: real-time monitoring and predictive analytics, precision medicine and personalized treatment, telemedicine and remote healthcare, and integration with existing healthcare systems [25].

Chatbot technology is revolutionizing various aspects of the medical field from improving documentation to aiding diagnosis and research. Although these AI systems offer tremendous potential, they have limitations that must be acknowledged and managed to ensure patient safety and a high quality of care. As the medical community continues to explore and integrate these
technologies, maintaining a cautious and adaptive approach to harness their benefits while mitigating the risks associated with their use is essential.

**Conclusion**

In the continually evolving healthcare landscape, the integration of chatbot technology into the fields of OB/GYN displays great promise [13]. Although clinical validation in this field is ongoing, recent studies have underscored the potential of ChatGPT in various ways [3].

One promising application of ChatGPT in medical imaging is to support radiologists in interpreting images. ChatGPT can assist in generating reports, identifying abnormalities, and suggesting further diagnostic tests. For instance, ChatGPT can generate an initial report for a patient using a chest radiograph, which a radiologist can then review. The application can also help identify potential imaging abnormalities that radiologists can subsequently verify.

Another valuable application of ChatGPT in pathology is the generation of pathological reports. ChatGPT can summarize the findings of the pathologist, including diagnosis, prognosis, and treatment recommendations, thereby reducing workload and enhancing efficiency in pathology departments.

Additionally, ChatGPT can contribute to education and training in medical imaging and pathology by creating interactive learning modules and offering feedback to students regarding
their work. However, challenges exist that must be addressed before their widespread adoption in the aforementioned fields. ChatGPT may occasionally produce incorrect or misleading information and currently struggles with nuances in medical language.

This AI-driven conversational agent has not only surpassed highly trained human counterparts in OBGYN examinations [3] but has also marked a significant advancement in integrating AI into specialized medical domains. The application of ChatGPT in OBGYN goes beyond examination improvement and extends to the development of differential diagnoses and guidance for rare prenatal conditions [13]. This innovative approach holds promise for various specialists, including perinatologists, geneticists, and neonatologists [13]. ChatGPT-4's ability to generate concise lists of diagnoses closely aligned with those of clinicians indicates progress in addressing challenges related to misinformation, frequency bias, and variations in responses based on phrasing [13].

Furthermore, the utility of ChatGPT extends to patient engagement and knowledge dissemination. Investigations of patient inquiries regarding high blood pressure during pregnancy highlight its potential as a resource for providing answers to patient queries [26]. This functionality fosters informed patient decision-making and enhances the healthcare experience in OBGYN. Although challenges related to misinformation, frequency bias, and rare diagnoses exist, they should not deter further exploration of the applications of ChatGPT. In conclusion, the initial stages of ChatGPT integration into OBGYN offer a promising glimpse into the transformative potential of AI in specialized medical domains [3,13,26]. As research and development continue, medical professionals must remain vigilant and adaptable, harnessing AI's capabilities of AI while upholding
the highest patient care and safety standards [13].

Conflict of interest

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

Ethical approval

The study is not applicable to Institutional Review Board.

Patient consent

Patient consent was not required in this review article.

Funding information

None.
Reference


12. Atlas S. ChatGPT for higher education and professional development: a guide to conversational AI [Internet]. College of Business Faculty Publications; c2023 [cited YEAR Mon Day]. Available from:


