REVIEW ARTICLE

ChatGPT Utility in Medical Education: A Systematic Review and Meta-Analysis

Fariba Hosseinzadegan[1], Seyedeh Masumeh Hashemi[2], Seyed Vahid Sharifi[3], Mehrnoosh Khoshnoodifar[4]

Published: 30 August 2025

Abstract

Background Learning and developing skills are crucial in science and technology. Recent advancements in communication technologies and Artificial Intelligence (AI) have introduced new tools for education, including AI chatbots that can interact with humans. Considering the importance of the subject in today's world and the impact of ChatGPT utility in medical education, the present study was conducted.

Methods A search of international databases such as PubMed, Web of Science, Scopus, Science Direct, Web of Knowledge, EBSCO, Wiley, ISI, Elsevier, Embase databases, and Google Scholar search engine was conducted based on the PRISMA 2020-27-item checklist and keywords related to the objectives of the study. The search covered the period from 2019 to February 18, 2024. The total number of articles retrieved from the database search was 668. After applying specific, predefined criteria, we selected a final set of five relevant studies from 43 potentially eligible articles. All statistical analyses were performed using STATA/MP software version 17, with a significance level of less than 0.05.

Results Five studies were selected according to the inclusion criteria. The random effects of response frequency on the positive impact of AI on medical education showed that ES = 82%, 95% CI: 51%-100%, p value < 0.001, which indicates a response frequency of 82% with a significant p value.

Conclusion To adequately prepare future medical professionals, there is an urgent need to integrate the teaching of AI into medical curricula.

Keywords Artificial intelligence, ChatGPT, Medical education

✉ Mehrnoosh Khoshnoodifar
mkhoshnoodifar@sbmu.ac.ir

1. Department of Nursing, School of Nursing and Midwifery, Tabriz University of Medical Sciences, Tabriz, Iran
2. Pediatric Intensive Care Department, Mofid Children Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
3. Faculty of Medical Education and Learning Technologies, Shahid Beheshti University of Medical Sciences, Tehran, Iran
4. Department of E-learning, Faculty of Medical Education and Learning Technologies, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Received: 03 January 2025/ Revised:11 June 2025/ Accepted: 07 May 2025

1 Introduction

In the age of information technology, Artificial Intelligence (AI) tools are increasingly integral to daily life and are transforming education and learning.[1,2] Collaborative learning and group interactions are widely used as effective, activity-based educational methods to improve skills.[3] With the emergence of AI technology and the development of systems such as ChatGPT, new possibilities for group interactions in the educational environment have emerged.[4] In addition, in today's world, where online education is widespread, the need to provide opportunities for participation and interaction in the educational environment is felt more than ever before.[5] Using ChatGPT as an intelligent partner and learning companion can allow learners to actively participate in the learning process and benefit from collaborative learning experiences and interaction with each other.[6]

AI is a branch of computer science focused on programming computers to perform human-like tasks. Its goal is to develop systems that can make human-like decisions. AI has subfields including machine learning, natural language processing, machine vision, robotics, and many others.[7] Each of these subfields focuses on specific aspects of AI. AI is the primary tool used in the development of chatbots. Machine learning algorithms and natural language processing are effective in creating AI-powered chats. These algorithms allow chatbots to converse with humans and answer various questions and requests.[8] Chatbots are software programs that can interact with humans using AI and natural language processing algorithms. This interaction may take the form of text (such as messages) or audio (using voice recognition systems).[9] Chatbots are used in various applications, including customer support, answering common questions, creating phone bots, analyzing data, and many other uses.[10] The beneficial applications of AI in healthcare have been outlined previously in areas such as drug discovery, personalized medicine, and the analysis of large datasets, in addition to its potential to improve diagnosis and clinical decision-making.[11] The capabilities of chatbots are continually improving due to advancements in AI. Advanced chatbots can answer more complex questions, analyze data, and even recognize human emotions in conversations. Chatbots are major players in the digital and internet world, serving as quick and interactive tools for human interaction, and are used in many industries and applications. As the ability of these chatbots to interact with humans improves, it is expected that their role in daily life and business will increase over time.[12]

As the technologies of AI and machine learning have achieved significant progress in the field of education and learning,[13] the possibility of group interaction has improved by using ChatGPT, and innovations in collaborative learning methods have been created. Among other strengths of using ChatGPT in group interactions, it should be noted that it facilitates the sharing of resources and knowledge.[14] This technology can serve as an interface for exchanging information and resources among groups. Additionally, ChatGPT can be utilized to offer guidance for group problem-solving and assist learners in finding the most effective solutions tailored to their group's needs and everyday tasks. However, using ChatGPT in group interactions also comes with challenges. For example, the ability to analyze ChatGPT content and the ability to understand question-and-answer texts accurately may have limitations.[15] Also, over-reliance on AI technologies like ChatGPT can lead to reduced reliance on social skills and human interactions.[16] Human-machine interaction has also led to new security and ethical challenges. Among these challenges, we can mention the preservation of privacy in online interactions, ethical differences in the use of AI, and the social and psychological effects of these technologies. Studying the impact of human-machine interaction as a research topic is very important. ChatGPT does not meet the requirements to be listed as an author in scientific articles unless the ICMJE/COPE guidelines are changed or revised, according to a systematic review on the promising perspectives and valid concerns. There is an immediate need for an initiative that encompasses all parties involved in healthcare research, education, and practice. This will support the establishment of an ethical code that will direct other LLMs working in academia and healthcare to use ChatGPT responsibly.[17] Additional systematic reviews that focused on the prospects and potential drawbacks revealed that ChatGPT has specific uses in health education, such as the development of potentially useful personalized learning tools and a shift in emphasis toward problem-based learning and critical thinking. As a result, ChatGPT is ineligible for authorship listings until the ICMJE/COPE guidelines are updated and changed.[18] ChatGPT has the potential to revolutionize medical education, according to a recent systematic scoping review, but its full integration will need careful thought. The effects of AI on both teachers and students should be taken into consideration to fully utilize ChatGPT in medical education.[19] Complying with the consensus of prior research, this study attempted to investigate response frequency regarding the positive influence of AI on medical education through a meta-analysis. Previous studies have taken into account the majority of the literature review and have looked at the opportunities, challenges, and barriers.

Research in this field helps advance AI technologies, machine learning methods, the development of optimal user relationships, and the formulation of policies and laws related to these interactions.[20] Considering the

importance of the subject in today's world and the impact of ChatGPT utility in medical education, the present study was conducted.

2 Methods

Search Strategy and Information Sources

To search, international databases such as PubMed, Scopus, ISI, Embase, and the Google Scholar search engine were utilized based on the PRISMA 2020-27-item checklist [21] and keywords related to the study objectives. The search was conducted from 2019 to February 18, 2024, using standardized MeSH keywords. Additionally, relevant studies were identified by screening the reference lists of selected articles. The search strategy included terms and keywords related to the study objectives. Our study PICO framework was as follows: Population (P): Medical students or medical trainees, Intervention (I): Use of ChatGPT, Comparison (C): Traditional educational methods or no use of AI-based tools, and Outcome (O): Improvement in knowledge acquisition, learning efficiency, or academic performance. The search was conducted from 2019 to February 18, 2024, using standardized MeSH keywords. Additionally, relevant studies were identified by screening the reference lists of selected articles. The search strategy included terms such as "Education, Medical," "Artificial Intelligence," "Machine Learning," and "Deep Learning," as well as keywords like ChatGPT, chatbots, digital health, healthcare, ethics, learning, e-learning, and higher education. First, a list of titles and abstracts of all articles retrieved from the reviewed databases was prepared. This work was done independently by two researchers. Then, articles with duplicate titles were removed. Next, the abstracts of the remaining articles were reviewed to identify suitable studies for inclusion. All selected studies were saved in EndNote X8 software, and the remaining steps of the review were conducted using the software.

Study Selection Criteria

The inclusion criteria encompassed studies discussing the utility of ChatGPT in medical education, reports on AI response frequency in medical education, as well as descriptive, cross-sectional, and experimental studies. Languages other than English, review papers, conference papers, case reports, correspondences, discussions, viewpoint papers, editorials, mini-reviews, short communication papers, book chapters, and news articles were excluded from the study.

Selection and Data Collection Process

To minimize reporting bias and data collection errors, two researchers independently extracted data using a standardized form. This form included items such as author name, category, year of publication, and study type.

Quilty Assessment

For cross-sectional studies, a modified Newcastle-Ottawa Scale recommended by the Cochrane Collaboration was utilized. A maximum score of 9 points (8 points total) could be obtained on this scale by combining the three dimensions of selection, comparability, and outcome or awareness. Scores were categorized as low (less than four), medium (five or six), or high (more than seven).[22]

Meta Analysis

The I2 statistic, used to measure inconsistency, was applied to analyze the degree of variation across studies (heterogeneity). Low levels of heterogeneity were defined as I2 = 25 – 49%; moderate levels as I2 = 50 – 74%, and high levels as I2 = 75 – 100%.[23] A random-effects model with the Restricted Maximum Likelihood (REML) method was used. All statistical analyses were performed using STATA/MP software, version 17, considering a significance level of less than 0.05.

3 Results

Study Selection

Following a review of article titles, 668 of the original search results were eliminated, along with 84 duplicate and overlapping articles. A total of 547 potentially relevant article abstracts were examined, and 504 unrelated articles were excluded. After reviewing the full texts of the remaining 43 articles, five appropriate studies were ultimately selected for inclusion in the meta-analysis (Figure 1).

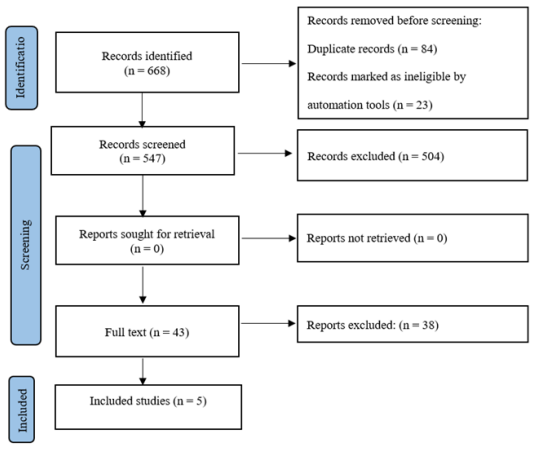

Figure 1 PRISMA 2020 Checklist

Study Characteristics

Table 1 summarizes the characteristics of these five studies. These five included studies were published between 2019 and 2024 and conducted across countries,

including Germany, Austria, Switzerland, and Turkey. Regarding their studied medical students,[24-27] one study population focused on healthcare professionals, academicians, and medical students.[28] Four studies used a cross-sectional method,[24-27] and one used the Delphi method.[28] Five studies with 4,455 participants were reviewed (Table 1).

Table 1 The characteristics of the selected articles according to the purpose of the study

Study. years	Study design	Country	Statistical population	Sample size	AI-based chat applications	ChatGPT utility	Effect of AI on variable/ outcome
Weidener et al., 2024 [27]	Cross-sectional	Germany, Austria, and Switzerland	Surveyed medical students	487	ChatGPT	Web-based survey	Students' perceptions of AI, the positive impact of AI on medical education
Roos et al., 2023 [26]	Cross-sectional	Germany	Medical students	630	ChatGPT	Media-related questions	Students' perceptions of AI, the positive impact of AI on medical education
Civaner et al., 2022 [24]	Cross-sectional	Turkey	Medical students	2,981	ChatGPT	Online survey	Students' perceptions of AI, the positive impact of AI on medical education
Çalışkan et al., 2022 [28]	Delphi method	Turkey	Healthcare professionals/ academicians; medical students	94	ChatGPT	Delphi survey	Competencies required by medical graduates, the positive impact of AI on medical education
Pinto dos Santos et al., 2019 [25]	Cross-sectional	Germany	Medical students	263	Artificial intelligence	Web-based questionnaire	Students' knowledge and attitude towards AI, and the positive impact of AI on medical education

The random-effects response frequency on the positive impact of AI on medical education showed that ES = 82%, 95% CI: 51%-100%, p < 0.001, which means that the response frequency was 82% with a significant p (Figure 2). The heterogeneity test showed Q = 0.10, p = 1.00, I2 = 0%, indicating low heterogeneity (0%) among the studies. Eighty-two percent of the 4,455 medical students reported a positive impact of AI on medical education.

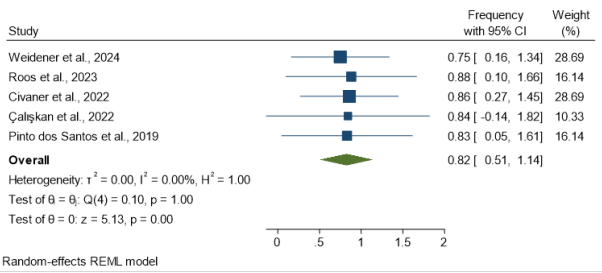

Figure 2 The forest plot illustrates the positive impact of AI on medical education

4 Discussion

Using ChatGPT in group interactions among medical students can lead to significant improvements in critical thinking, problem-solving, social skills, and cooperation. This technology can also increase motivation and creativity, encouraging active and deeper learning.[29]

The current meta-analysis reveals an 82% positive effect of ChatGPT on learning in medical education; however, several challenges and issues remain. One of these challenges is ChatGPT's ability to analyze and understand academic content. Although ChatGPT is capable of answering questions, it may not correctly analyze some complex or specialized content and might provide ineffective answers.[30]

To improve the use of ChatGPT in group interactions and collaborative learning, solutions such as enhancing its analytical capabilities can be implemented. Enhancing these capabilities can lead to a more accurate understanding of questions and more effective responses. Additionally, developing algorithms for diagnosing and evaluating group interactions, as well as determining each participant's level of engagement, can be helpful. Enhancing ChatGPT's multilingual capabilities can also be beneficial in group interactions that involve different languages. With the ability to provide answers and guidance in various languages, ChatGPT can facilitate cross-cultural and cross-linguistic interactions and provide an opportunity for collaborative learning and knowledge exchange.[31]

Combining ChatGPT with other technologies such as digital publications, online learning platforms, and group collaboration tools can improve and expand group interactions in educational environments. By integrating ChatGPT with these technologies, it is possible to improve group interactions and facilitate collaborative activities. To increase the confidence of medical students in ChatGPT and encourage active participation in the learning process, it is essential to provide appropriate feedback and motivation.[32]

There are various ways to implement AI in clinical practice, and clinicians and health informaticians require formal training to adopt the most effective approaches. Health informaticians and physician champions who design and develop AI-based protocols need to have a good understanding of complex algorithms, data quality assessment methodologies, probabilistic forecasting, and comparative model assessment to work with engineers and develop reliable AI applications. Moreover, clinicians who use AI applications should become familiar with potential challenges. AI-based relational time pattern analysis has replaced simple threshold-based diagnostic rules. Current medical education and health informatics curricula still do not provide sufficient training to understand AI communications or the necessary skill sets to develop AI systems that can detect and analyze relational time patterns. The ability to interpret errors in AI algorithms and devise strategies to correct them requires specialized training. Therefore, medical and health informatics education must emphasize algorithm-based platforms and include relevant data analytics and AI-related topics in their curricula. Furthermore, computer science and health informatics programs should consist of healthcare–focused digital skills training.[33]

5 Conclusion

The present meta-analysis indicates that ChatGPT improves learning in medical training by 82%. Through ChatGPT, medical students can engage in interactive and cooperative group activities. Human-machine interaction plays a crucial role in this direction. In the era of AI and information technology, lifelong learning and education have emerged as critical competencies for both social and professional spheres. However, while they present challenges such as security and ethical issues, they also facilitate the development of learning opportunities and access to new educational content. For this reason, more research in this area is needed to provide conclusive evidence in this regard. The future of AI for improving healthcare is exciting, and we are likely to see more use cases. Medical schools should find ways to offer machine learning training modules. A forward-thinking and straightforward initiative would be to provide introductory machine learning courses as part of medical education at accredited institutions.

Declarations

Acknowledgments
Not applicable.

Authors' Contributions
Fariba Hosseinzadegan and Mehrnoosh Khoshnoodifar designed the study. Fariba Hosseinzadegan, Seyedeh Masumeh Hashemi, and Seyed Vahid Sharifi conducted the literature search. Fariba Hosseinzadegan, Seyedeh Masumeh Hashemi, Seyed Vahid Sharifi, and Mehrnoosh Khoshnoodifar performed data analysis. Manuscript preparation was carried out by Fariba Hosseinzadegan, Seyedeh Masumeh Hashemi, Seyed Vahid Sharifi, and Mehrnoosh Khoshnoodifar, while Fariba Hosseinzadegan and Mehrnoosh Khoshnoodifar did manuscript editing.

Availability of Data and Materials
Upon a reasonable request, the corresponding author can provide the datasets analyzed in this study.

Conflict of Interest
The authors declare that they have no conflict of interest.

Consent for Publication
All authors have read and approved the final manuscript and have provided their consent for publication.

Funding
This study had no financial sponsorship.

Ethical Considerations
The current study is a systematic review and does not have a code of ethics.

References

1. Chen L, Chen P, Lin Z. Artificial intelligence in education: A review. IEEE Access. 2020; 8: 75264-75278.
2. Mansouri V, Gholizadeh S, Hosseinpoor S. Impact of artificial intelligence on medical entomology research. Health Science Monitor. 2024; 3(2): 113-119.
3. Bilad MR, Yaqin LN, Zubaidah S. Recent progress in the use of artificial intelligence tools in education. Jurnal Penelitian dan Pengkajian Ilmu Pendidikan: e-Saintika. 2023; 7(3): 279-314.
4. Venkateswaran P, Ayasrah FTM, Nomula VK, Paramasivan P, Anand P, Bogeshwaran K. Applications of Artificial Intelligence Tools in Higher Education. In: Data-Driven Decision Making for Long-Term Business Success. IGI Global; 2024. p. 124-136.
5. Pedro F, Subosa M, Rivas A, Valverde P. Pedro F, Subosa M, Rivas A, Valverde P. Artificial intelligence in education: challenges and opportunities for sustainable development. Paris: UNESCO; 2019.
6. Delgado HOK, de Azevedo Fay A, Sebastiany MJ, Silva ADC. Artificial intelligence adaptive learning tools: the teaching of English in focus. BELT-Brazilian English Language Teaching Journal. 2020; 11(2): e38749-e38749.
7. Fitria TN. Artificial Intelligence (AI) In Education: Using AI Tools for Teaching and Learning Process. Prosiding Seminar Nasional & Call for Paper STIE AAS; 2021:134-147.
8. Ruiz-Rojas LI, Acosta-Vargas P, De-Moreta-Llovet J, Gonzalez-Rodriguez M. Empowering education with generative artificial intelligence tools: Approach with an instructional design matrix. Sustainability. 2023; 15(15): 11524.
9. Kim N-Y, Cha Y, Kim H-S. Future english learning: Chatbots and artificial intelligence. Multimedia-Assisted Language Learning. 2019; 22(3):32.
10. Lin Y, Yu Z. A bibliometric analysis of artificial intelligence chatbots in educational contexts. Interactive Technology and Smart Education. 2024; 21(2): 189-213.
11. Johnson KB, Wei WQ, Weeraratne D, Frisse ME, Misulis K, Rhee K, et al. Precision medicine, AI, and the future of personalized health care. Clin Transl Sci. 2021; 14(1): 86-93.

12. Skrebeca J, Kalniete P, Goldbergs J, Pitkevica L, Tihomirova D, Romanovs A. Modern development trends of chatbots using artificial intelligence (ai). 2021 62nd International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS); 2021: IEEE.

13. Birenbaum M. The Chatbots' Challenge to Education: Disruption or Destruction? Educ Sci. 2023; 13(7): 711.

14. Gill SS, Xu M, Patros P, Wu H, Kaur R, Kaur K, et al. Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots. Internet of Things and Cyber-Physical Systems. 2024; 4: 19-23.

15. Ngo TTA. The perception by university students of the use of ChatGPT in education. International Journal of Emerging Technologies in Learning (Online). 2023; 18(17): 4.

16. Kostick-Quenet K, Lang BH, Smith J, Hurley M, Blumenthal-Barby J. Trust criteria for artificial intelligence in health: normative and epistemic considerations. J Med Ethics. 2024; 50(8): 544-551.

17. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. Healthcare; 2023: 11(6):887.

18. Sallam M. The utility of ChatGPT as an example of large language models in healthcare education, research and practice: Systematic review on the future perspectives and potential limitations. MedRxiv. 2023; 23286155.

19. Xu X, Chen Y, Miao J. Opportunities, challenges, and future directions of large language models, including ChatGPT in medical education: a systematic scoping review. J Educ Eval Health Prof. 2024; 21: 6.

20. Tlili A, Shehata B, Adarkwah MA, Bozkurt A, Hickey DT, Huang R, et al. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learn Environ. 2023; 10(1): 15.

21. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Int J Surg. 2021; 88: 105906.

22. Rodríguez AJ, Nunes VdS, Mastronardi CA, Neeman T, Paz-Filho GJ. Association between circulating adipocytokine concentrations and microvascular complications in patients with type 2 diabetes mellitus: A systematic review and meta-analysis of controlled cross-sectional studies. J Diabetes Complications. 2016; 30(2): 357-367.

23. Sokouti M, Shafiee-Kandjani AR, Sokouti M, Sokouti B. A meta-analysis of systematic reviews and meta-analyses to evaluate the psychological consequences of COVID-19. BMC Psychol. 2023; 11(1): 279.

24. Civaner MM, Uncu Y, Bulut F, Chalil EG, Tatli A. Artificial intelligence in medical education: a cross-sectional needs assessment. BMC Med Educ. 2022; 22(1): 772.

25. Pinto dos Santos D, Giese D, Brodehl S, Chon S-H, Staab W, Kleinert R, et al. Medical students' attitude towards artificial intelligence: a multicentre survey. Eur Radiol. 2019; 29: 1640-1646.

26. Roos J, Kasapovic A, Jansen T, Kaczmarczyk R. Artificial intelligence in medical education: Comparative analysis of ChatGPT, Bing, and medical students in Germany. JMIR Med Educ. 2023; 9(1): e46482.

27. Weidener L, Fischer M. Artificial Intelligence in Medicine: Cross-Sectional Study Among Medical Students on Application, Education, and Ethical Aspects. JMIR Med Educ. 2024; 10(1): e51247.

28. Çalışkan SA, Demir K, Karaca O. Artificial intelligence in medical education curriculum: An e-Delphi study for competencies. PLoS One. 2022; 17(7): e0271872.

29. Boscardin CK, Gin B, Golde PB, Hauer KE. ChatGPT and generative artificial intelligence for medical education: potential impact and opportunity. Acad Med. 2024; 99(1): 22-27.

30. Bond M, Khosravi H, De Laat M, Bergdahl N, Negrea V, Oxley E, et al. A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour. Int J Educ Technol High Educ. 2024; 21(1): 4.

31. Díaz-Guio DA, Henao J, Pantoja A, Arango MA, Díaz-Gómez AS, Gómez AC. Artificial intelligence, applications and challenges in simulation-based education. Colombian Journal of Anestesiology. 2024; 52(1).

32. Eysenbach G. The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers. JMIR Med Educ. 2023; 9(1): e46885.

33. Sapci AH, Sapci HA. Artificial intelligence education and tools for medical and health informatics students: systematic review. JMIR Med Educ. 2020; 6(1): e19285.