

Original Paper

ChatGPT in Medical Education: Bibliometric and Visual Analysis

Yuning Zhang¹, MM; Xiaolu Xie², ME; Qi Xu³, MD

¹School of Basic Medical Sciences, Gannan Medical University, Ganzhou, China

²School of Medical and Information Engineering, Gannan Medical University, Ganzhou, China

³School of Public Health and Health Management, Gannan Medical University, Ganzhou, China

Corresponding Author:

Qi Xu, MD

School of Public Health and Health Management, Gannan Medical University

1 Harmony Avenue, Rongjiang New District

Ganzhou 341000

China

Phone: 86 15770701553

Email: kiki2661@163.com

Abstract

Background: ChatGPT is a generative artificial intelligence-based chatbot developed by OpenAI. Since its release in the second half of 2022, it has been widely applied across various fields. In particular, the application of ChatGPT in medical education has become a significant trend. To gain a comprehensive understanding of the research developments and trends regarding ChatGPT in medical education, we conducted an extensive review and analysis of the current state of research in this field.

Objective: This study used bibliometric and visualization analysis to explore the current state of research and development trends regarding ChatGPT in medical education.

Methods: A bibliometric analysis of 407 articles on ChatGPT in medical education published between March 2023 and June 2025 was conducted using CiteSpace, VOSviewer, and Bibliometrix (RTool of RStudio). Visualization of countries, institutions, journals, authors, keywords, and references was also conducted.

Results: This bibliometric analysis included a total of 407 studies. Research in this field began in 2023, showing a notable surge in annual publications until June 2025. The United States, China, Türkiye, the United Kingdom, and Canada produced the most publications. Networks of collaboration also formed among institutions. The University of California system was a core research institution, with 3.4% (14/407) of the publications and 0.17 betweenness centrality. *BMC Medical Education*, *Medical Teacher*, and the *Journal of Medical Internet Research* were all among the top 10 journals in terms of both publication volume and citation frequency. The most prolific author was Yavuz Selim Kiyak, who has established a stable collaboration network with Isil Irem Budakoglu and Ozlem Coskun. Author collaboration in this field is usually limited, with most academic research conducted by independent teams and little communication between teams. The most frequent keywords were “AI,” “ChatGPT,” and “medical education.” Keyword analysis further revealed “educational assessment,” “exam,” and “clinical practice” as current research hot spots. The most cited paper was “Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models,” and the paper with the strongest citation burst was “Are ChatGPT’s Knowledge and Interpretation Ability Comparable to Those of Medical Students in Korea for Taking a Parasitology Examination?: A Descriptive Study.” Both papers focus on evaluating ChatGPT’s performance in medical exams.

Conclusions: This study reveals the significant potential of ChatGPT in medical education. As the technology improves, its applications will expand into more fields. To promote the diversification and effectiveness of ChatGPT in medical education, future research should strengthen interregional collaboration and enhance research quality. These findings provide valuable insights for researchers to identify research perspectives and guide future research directions.

JMIR Med Educ 2025;11:e72356; doi: [10.2196/72356](https://doi.org/10.2196/72356)

Keywords: ChatGPT; medical education; bibliometric; VOSviewer; CiteSpace; artificial intelligence; AI

Introduction

Background

Large language models (LLMs) represent major progress in artificial intelligence (AI), especially for computational linguistics and natural language processing. These generative AI models are fundamentally based on the transformer neural network architecture [1]. Training is conducted using extensive text datasets, including books, documents, and website content. LLMs have been developed to predict subsequent words or tokens. Through this process, they learn to recognize complex language patterns, including vocabulary, grammar, semantics, and even specialized knowledge such as medicine [2].

ChatGPT, an AI chatbot developed by OpenAI [3], was launched in November 2022 as an LLM [4,5]. Built on the generative pretrained transformer architecture, it uses tens of billions of parameters trained on massive internet text datasets [6]. ChatGPT excels at understanding and generating humanlike language, conducting natural dialogues, and delivering high-quality responses to user queries [7,8]. Its advanced text processing capabilities have driven unprecedented adoption: more than 1 billion monthly users within 4 months of release, demonstrating rapid societal integration [9].

ChatGPT demonstrates significant potential across diverse fields, such as translation, text summarization, and programming assistance [10]. Its effectiveness extends to specialized domains such as medical education [11]. In preclinical education, students use ChatGPT for medical knowledge acquisition and personalized learning [12,13]. Conversely, educators are able to use ChatGPT to implement innovative teaching methodologies and cultivate interactive learning environments [14-16]. In clinical education [17], ChatGPT simulates clinical environments to help students improve clinical skills [18-21]. Furthermore, the pass rates and accuracy in medical licensing exams and professional subject tests [22,23] in countries such as the United States [24-26], China [27,28], Japan [29,30], and Italy [31] have attracted significant attention [32]. ChatGPT is regarded as a significant instrument for promoting innovation and enhancing efficiency in the domain of medical education.

Objectives

While existing literature reviews have explored ChatGPT's applications and limitations in medical education, important questions remain unanswered. These include the collaboration networks among countries, institutions, and authors; the most influential journals; and the most cited publications. This study used bibliometric analysis to map collaboration networks and thematic evolution and provide a comprehensive understanding of the development of ChatGPT in medical education.

A bibliometric analysis is a rigorous scientific method that provides researchers across various fields with comprehensive guidance and support [33]. It allows researchers to gain in-depth insights into prevailing issues, key trends, and

research limitations within their disciplines [34-36]. On the basis of recommendations from previous studies, this study proposes the following research questions (RQs):

1. Who are the most productive researchers and which are the most productive institutions and countries or regions in the field of ChatGPT in medical education?
2. What is the status of academic collaboration among researchers, countries, or regions in the field of ChatGPT in medical education?
3. Which are the most influential journals and articles in the field of ChatGPT in medical education?
4. What are the main research themes in the field of ChatGPT in medical education?
5. What are the research trends for ChatGPT in medical education?

Methods

Literature Sources and Search Strategy

The Web of Science database was chosen for this research due to its extensive coverage of more than 12,000 academic journals. When compared to other databases, including PubMed, MEDLINE, and Scopus, Web of Science offers a robust and reliable framework for bibliometric analysis [37]. After determining pertinent title keywords, a comprehensive bibliographic search was conducted online via the Web of Science database. The search was carried out in accordance with the following format:

```
((TS=(ChatGPT)) OR TS=(Chatbot*)) OR
TS=(Chat Generative Pre-trained Transformer) and
((((((((TS=(medic* educat*)) OR TS=(medic*
student*)) OR TS=(clinical clerkship*)) OR
TS=(medic* school*)) OR TS=(medic* learner*)) OR
TS=(medic* trainee*)) OR TS=(medic*clerk*)) OR
TS=(medical education)) OR TS=(medical student))
OR TS=(medical school)) OR TS=(medical student
education)) OR TS=(healthcare). NOT ALL=(retracted)
—Time: Tue Jul 01, 2025, 19:18:42 GMT+0800 (CST)
```

A total of 1817 documents were retrieved. These documents were screened according to the inclusion and exclusion criteria. The inclusion criteria were as follows: (1) original research articles and review articles related to ChatGPT in medical education and (2) English-language articles. After screening, of the 1817 retrieved articles, 1610 (88.61%) were retained. Following application of the exclusion criteria (articles unrelated to ChatGPT in medical education and duplicate articles), of the remaining 1610 articles, 1203 (74.72%) were excluded. The research topics of the 1203 excluded articles are summarized as follows: 370 (30.76%) were non-ChatGPT studies, 298 (24.77%) involved ChatGPT and patients, 263 (21.86%) involved ChatGPT and clinical treatment, 144 (11.97%) involved ChatGPT and hospitals, 52 (4.32%) involved ChatGPT and nonmedicine, 36 (2.99%) were ChatGPT non-medical education review articles, 20 (1.66%) involved ChatGPT and health care professional perspectives, 14 (1.16%) involved ChatGPT and nonmedical

interactions with students, and 5 (0.42%) involved ChatGPT and veterinary medicine. This resulted in 407 publications being selected for bibliometric analysis. A comprehensive dataset, along with the corresponding references, was subsequently extracted from the relevant publications and organized in plain-text format for future research endeavors. This process was conducted independently by 2 authors, who cross-verified their work. Any discrepancies were resolved by a senior author.

Data Collection and Statistics

The data were exported in plain-text file format using CiteSpace (version 6.3.R1; 64 bits; advanced) and R (version 4.5.0; R Foundation for Statistical Computing) with the *bibliometrix* package [38]. The data included the full record and cited references and were stored in the download format (.txt). The data extracted from the Bibliometric online platform [39] were exported in tab-delimited file format, with content and storage format identical to those described above.

CiteSpace, a bibliometric analysis software developed by Chaomei Chen, has achieved widespread use [40,41]. The software has been proven to provide feasible and reliable text mining and knowledge visualization methods. These methods have been used to explore research performance allocation and collaboration, research status and frontiers, and future trends. In this study, CiteSpace was used to detect parameters and visually analyze institution distribution, the dual-map overlay of journals, and burst detection.

Burst detection, based on the Kleinberg algorithm, uses an infinite state automaton to model document streams, thereby extracting meaningful structures [42]. These analyses can reveal themes exhibiting rapid growth over extended periods, as well as those that are inherently more transient.

VOSviewer, released in 2010 by Nees Jan van Eck and Ludo Waltman (Leiden University), is mainly used for

bibliometric network graph analysis [43]. We used VOSviewer version 1.6.20 to visualize and analyze the country distribution and collaboration, journal distribution, author distribution and collaboration, and keyword distribution.

In addition, Bibliometrix (RTool of RStudio; Posit PBC) was used to visualize the distribution of keywords over time in the form of a heat map [44]. Microsoft Excel 2024 was used to analyze the monthly publication trends of literature from March 2023 to June 2025.

Ethical Considerations

The study did not involve human participants, therefore the ethical approval was not required.

Results

Overview of Publication Status

Our search and screening efforts yielded 407 articles (Figure 1). The release of the AI chatbot ChatGPT in November 2022 was immediately followed by 2 review articles on ChatGPT published in March 2023, which provided an overview of ChatGPT in medical education and prediction of potential future applications for ChatGPT. The analysis of publication trends for this topic was conducted using Microsoft Excel 2024 and presented the number of publications in tabular form (Table 1). The number of articles showed a gradual increasing trend. In early 2023, the number of published articles per month was less than 10, and starting in September 2023, this number increased significantly. By 2024, the number of articles per month remained at more than 10, and in 2025, it was more than 20 articles per month. In May 2025, the number of articles per month reached 33. This implies that, over time, an increasing number of scholars are focusing their attention on this domain.

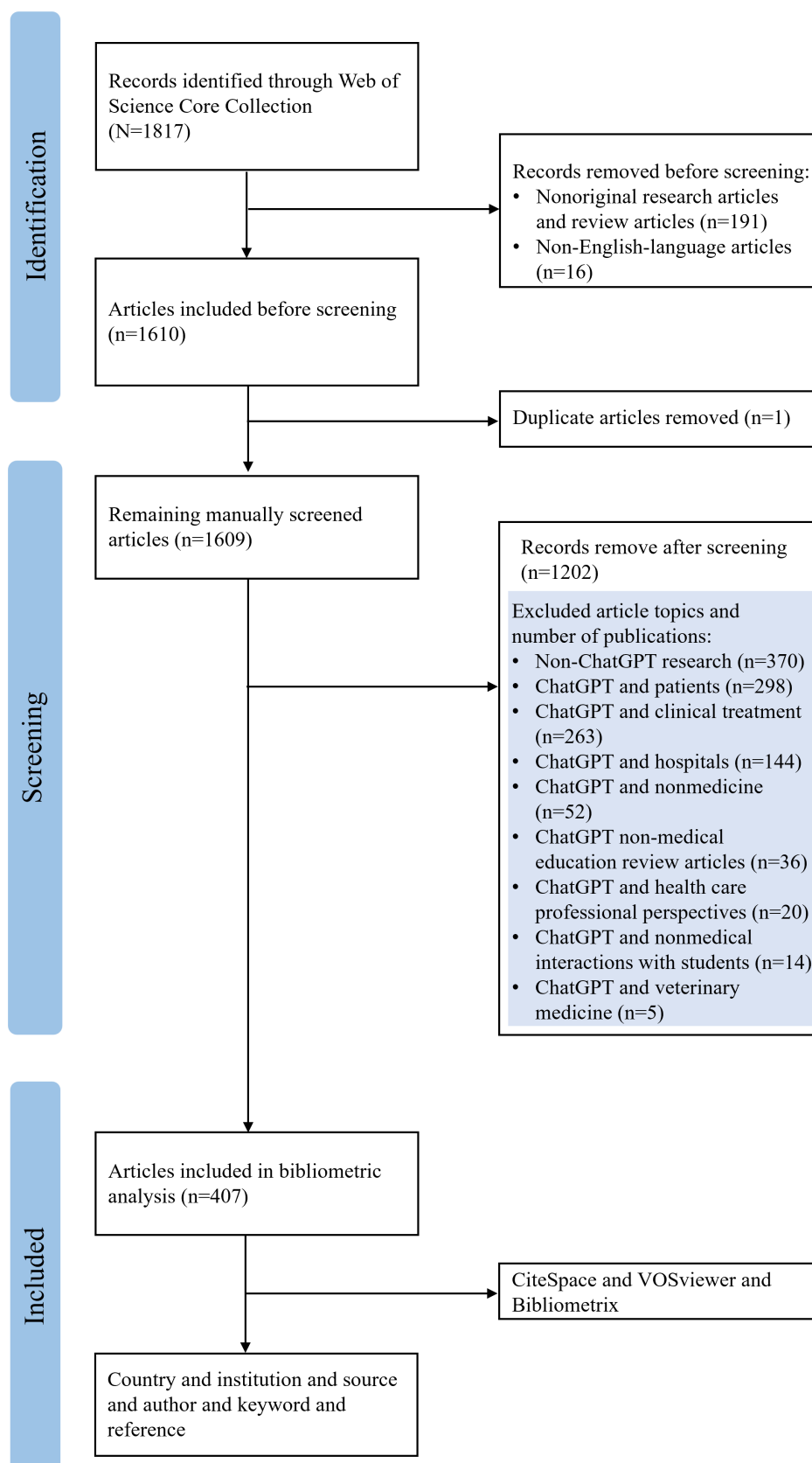
Figure 1. Flowchart of data collection and bibliometric analysis.

Table 1. Number of articles per month and cumulative number.

Month and year	Monthly publications, n	Cumulative publications, n
March 2023	2	2
April 2023	1	3
May 2023	2	5
June 2023	2	7
July 2023	2	9
August 2023	8	17
September 2023	14	31
October 2023	13	44
November 2023	7	51
December 2023	12	63
January 2024	14	77
February 2024	15	92
March 2024	9	101
April 2024	15	116
May 2024	19	135
June 2024	17	152
July 2024	15	167
August 2024	15	182
September 2024	17	199
October 2024	20	219
November 2024	16	235
December 2024	24	259
January 2025	23	282
February 2025	23	305
March 2025	26	331
April 2025	22	353
May 2025	33	386
June 2025	21	407

Analysis of National Publication Counts

Publication counts by country were used to analyze contributions in this field. According to the results, the publications originated from 66 countries. Visualizing the geographic distribution of the 66 countries using VOSviewer revealed that Asia, Europe, Africa, North America, South America, and Oceania were all represented and that the countries were

mainly concentrated in the northern hemisphere (Figure 2). In total, 38% (25/66) of the countries were in Asia, and 33% (22/66) were from Europe, the 2 continents with the highest number of countries in this study. It is noteworthy that the linkages between countries or regions were concentrated between East Asia and North America, East Asia and Europe, North America and Europe, and North America and Oceania.

Figure 2. Countries and regions involved in the research in this field. The links between the countries and regions indicate their collaborations and connections.

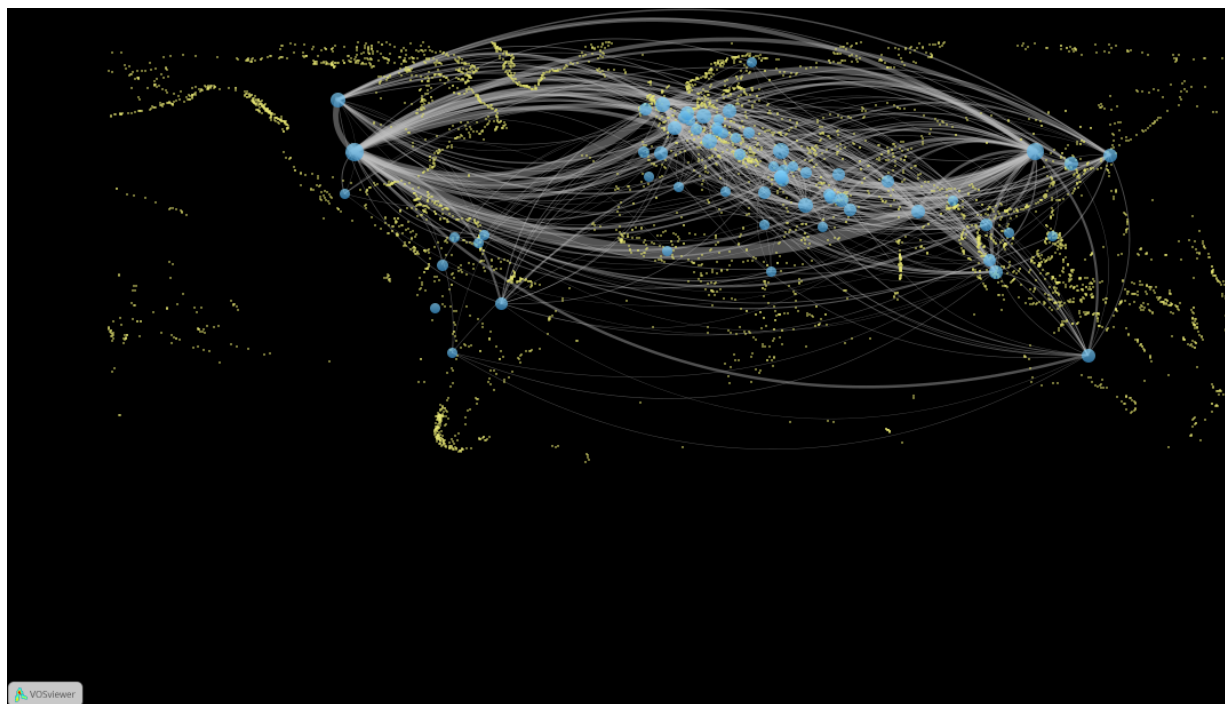


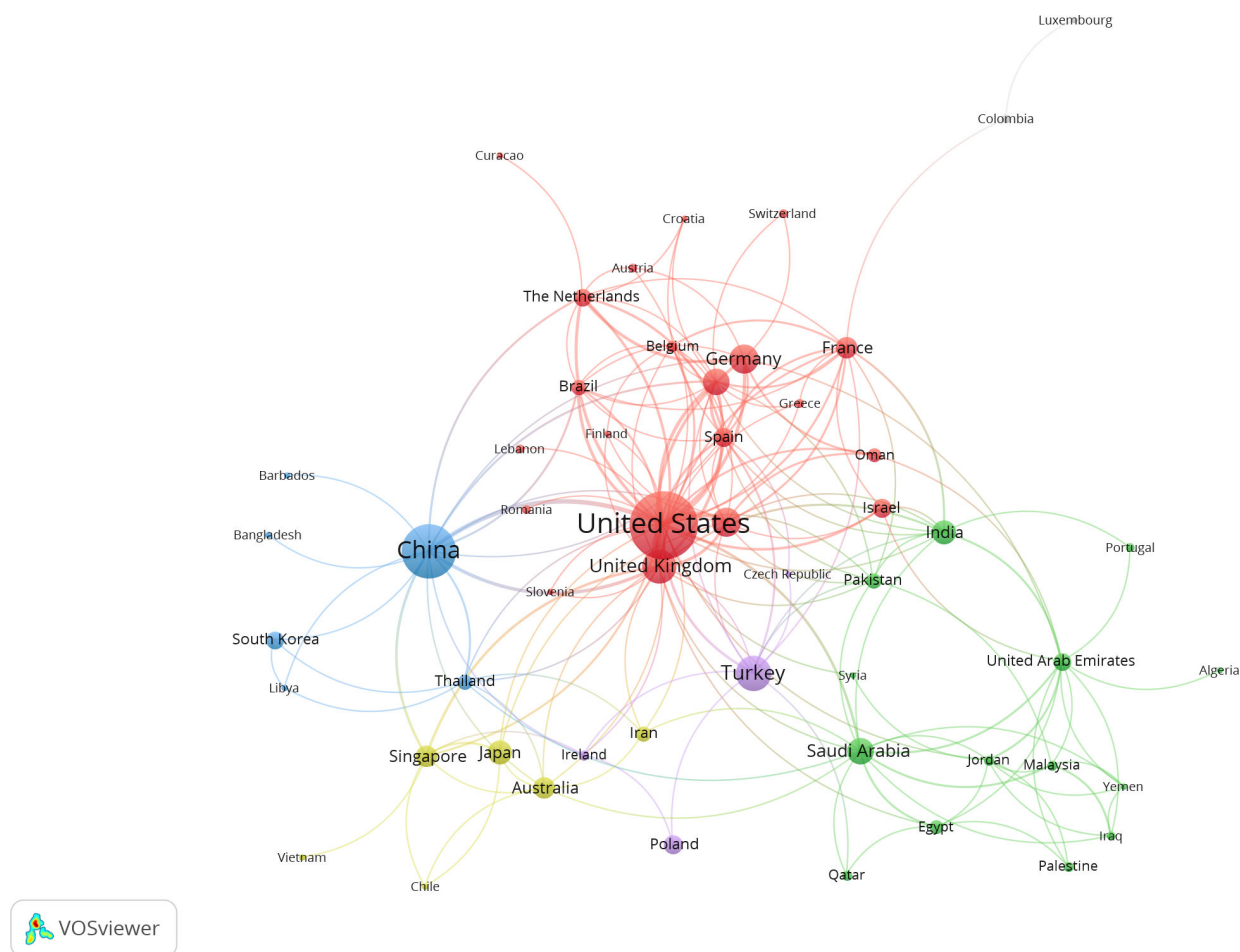
Table 2 shows the top 10 countries or regions in terms of number of publications and their corresponding citation frequency and centrality. The United States was the most prominent country with 31% (126/407) of the publications, closely followed by China, Türkiye, the United Kingdom, Canada, and Germany.

Table 2. Top 10 countries by number of publications and their number of citations and centrality (N=407).

Rank	Country	Publications, n (%)	Citations, n	Centrality
1	United States	126 (31)	3227	0.42
2	China	81 (19.9)	1026	0.19
3	Türkiye	35 (8.6)	215	0.08
4	United Kingdom	31 (7.6)	1844	0.17
5	Canada	24 (5.9)	668	0.05
6	Germany	22 (5.4)	463	0.06
7	Italy	19 (4.7)	730	0.06
8	Saudi Arabia	18 (4.4)	227	0.11
9	India	15 (3.7)	121	0.07
10	Japan	15 (3.7)	155	0.01

The results of the global collaboration network analysis show that countries and regions were roughly divided into 5 clusters in VOSviewer based on the closeness of collaboration and are indicated by different colors in Figure 3. The United States, China, Türkiye, the United Kingdom, and Canada were the top 5 countries in terms of the number of publications, and there were cooperative

relationships between them. The betweenness centrality (BC) was calculated when analyzing the national and regional collaboration networks using CiteSpace, which represents the strength of association between nodes. Among the top 10 countries, the United States, China, the United Kingdom, and Saudi Arabia were the main research centers in this field.

Figure 3. Analysis of collaborative network visualization of countries and regions in VOSviewer.

Analysis of Publication Institutions

The scientific output came from 847 institutions. CiteSpace identified 147 institutions with 346 cooperative networks (Figure 4). The most productive institutions regarding research in this field were the University of California system (14/407, 3.4% of publications), Harvard University (11/407, 2.7% of publications), National University of Singapore (11/407, 2.7% of publications), the Commonwealth System of Higher Education (11/407, 2.7% of publications),

the University of Toronto (10/407, 2.5% of publications), the University of London (8/407, 2% of publications), Gazi University (8/407, 2% of publications), the University of Pittsburgh (8/407, 2% of publications), Central South University (7/407, 1.7% of publications), and Stanford University (7/407, 1.7% of publications). Five of the top 10 institutions were from the United States. The remaining institutions were from Singapore, Canada, the United Kingdom, Türkiye, and China (Table 3).

Figure 4. Analysis of collaborative network visualization of institutions in CiteSpace.

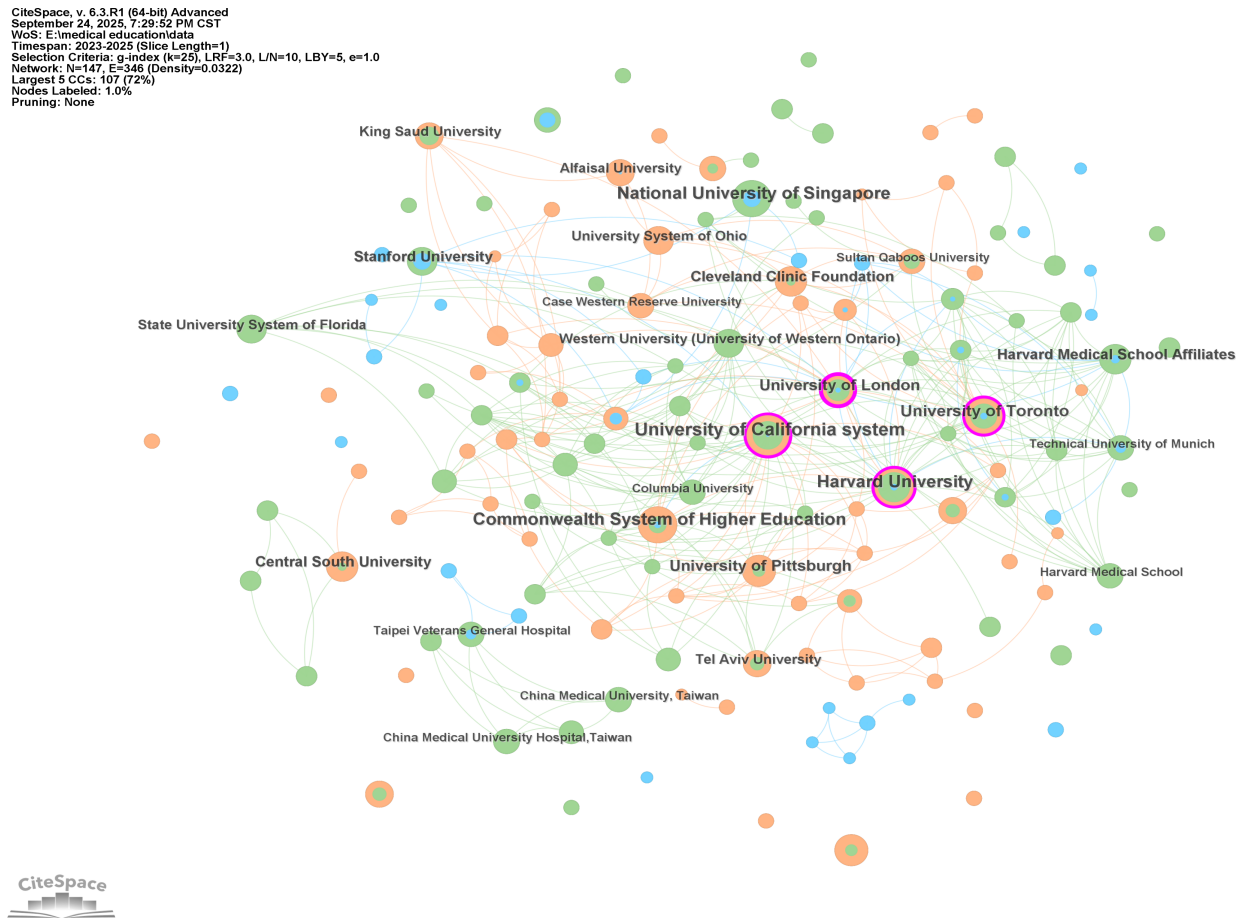


Table 3. Top 10 institutions and their centrality in CiteSpace (N=407).

Rank	Institution	Publications, n (%)	Centrality
1	University of California system	14 (3.4)	0.17
2	Harvard University	11 (2.7)	0.14
3	National University of Singapore	11 (2.7)	0
4	Commonwealth System of Higher Education	11 (2.7)	0.03
5	University of Toronto	10 (2.5)	0.1
6	University of London	8 (2)	0.22
7	Gazi University	8 (2)	0
8	University of Pittsburgh	8 (2)	0.01
9	Central South University	7 (1.7)	0
10	Stanford University	7 (1.7)	0.08

In CiteSpace, each node represents a institution, and the radius of a node increases with its contribution to research in the field, whereas the BC is proportional to the size of the purple ring around the nodes. The larger the purple circle, the larger the value of the betweenness centrality. Network visualization revealed that there were 4 central institutions: the University of California system (BC=0.17), Harvard University (BC=0.14), the University of Toronto (BC=0.10), and the University of London (BC=0.22; Figure 4). This reflects the significant bridging role of these institutions in the research on ChatGPT in medical education.

Analysis of Publication Quantity and Journal Impact

This study encompassed 407 articles published across 197 sources and journals. Table 4 lists the 10 most prolific sources and journals ranked by publication volume, along with their 2024 impact factor (IF). The top 10 journals published a total of 139 papers. Among these, *BMC Medical Education* (IF of 3.2; quartile 1; 40/407, 9.8% of publications) had the highest number of publications, followed by *Medical Teacher* (IF of 4.4; quartile 1; 32/407, 7.9% of publications), the *Journal of Medical Internet Research* (IF of 6.0; quartile 1; 11/407, 2.7% of publications), *Scientific Reports* (IF of 3.9; quartile 1; 11/407, 2.7% of publications), and *PLOS ONE*

(IF of 2.6; quartile 2; 10/407, 2.5% of publications). Eight of the top 10 journals in terms of publications were distributed in quartile 1 of the Journal Citation Reports (Figure 5). The source or journal with the highest cocitation frequency was *arXiv*, followed by *JMIR Medical Education*, *Cureus*, *Medical Teacher*, *BMC Medical Education*, and the *Journal of Medical Internet Research* (Figure 6). Seven of the top

10 sources or journals in terms of cocitation frequency were distributed in quartile 1 of the Journal Citation Reports (Table 5). It is important to note that 3 of the top 10 journals in terms of publications were also among the top 10 journals in terms of cocitation frequency: *Medical Teacher*, *BMC Medical Education*, and the *Journal of Medical Internet Research*.

Table 4. Top 10 sources by number of publications and their corresponding journal impact factor (IF; Journal Citation Reports [JCR] 2024) and JCR quartile (N=407).

Rank	Source	Publications, n (%)	IF (JCR 2024)	JCR quartile
1	<i>BMC Medical Education</i>	40 (9.8)	3.2	1
2	<i>Medical Teacher</i>	32 (7.9)	4.4	1
3	<i>Journal of Medical Internet Research</i>	11 (2.7)	6.0	1
4	<i>Scientific Reports</i>	11 (2.7)	3.9	1
5	<i>PLOS ONE</i>	10 (2.5)	2.6	2
6	<i>Frontiers in Medicine</i>	9 (2.2)	3.0	1
7	<i>Digital Health</i>	7 (1.7)	3.3	1
8	<i>Healthcare</i>	7 (1.7)	2.7	2
9	<i>Nurse Education Today</i>	6 (1.5)	4.2	1
10	<i>Postgraduate Medical Journal</i>	6 (1.5)	2.7	1

Figure 6. Analysis of collaborative network visualization of journals’ citations in VOSviewer.

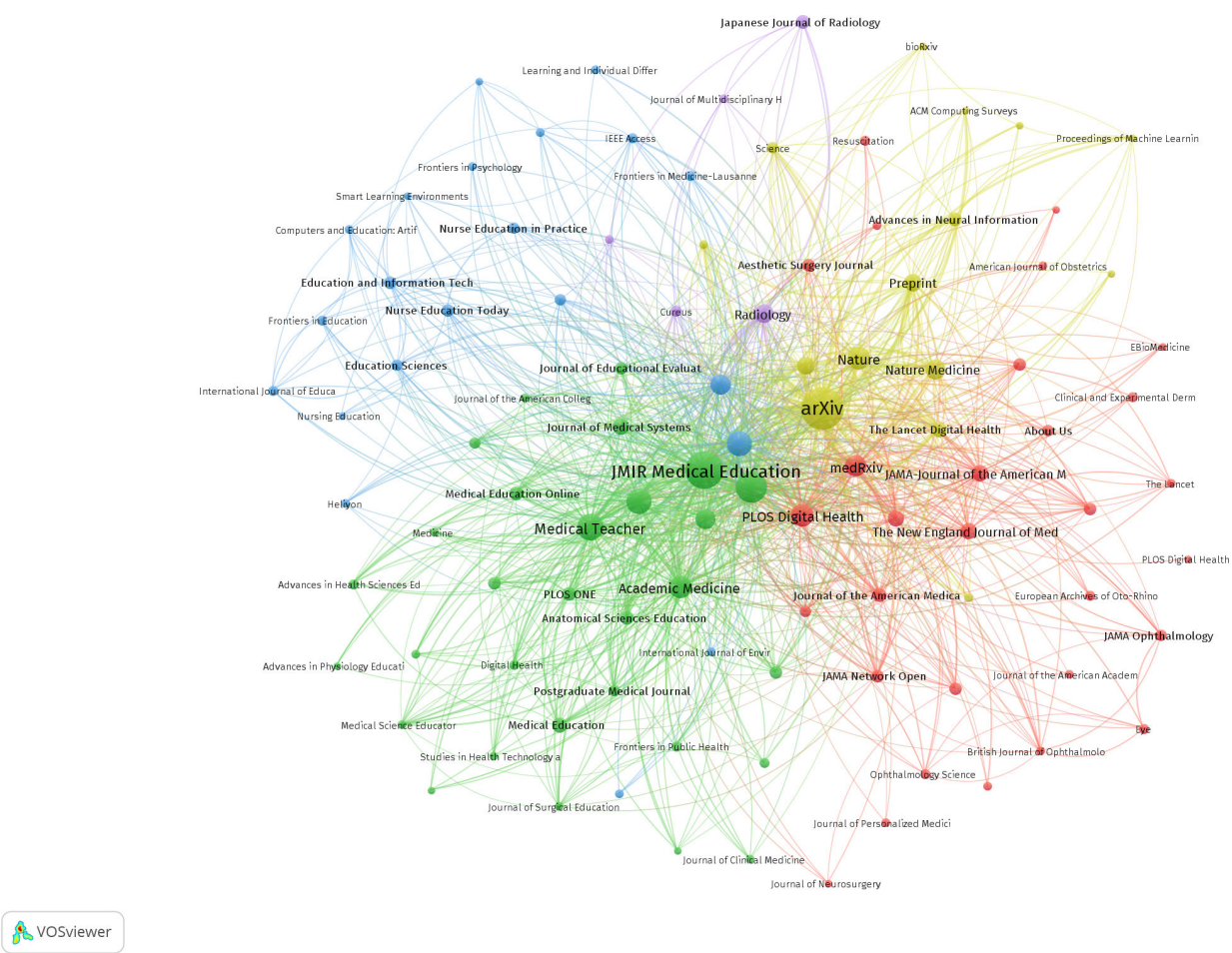


Figure 5. Analysis of collaborative network visualization of journals in VOSviewer.

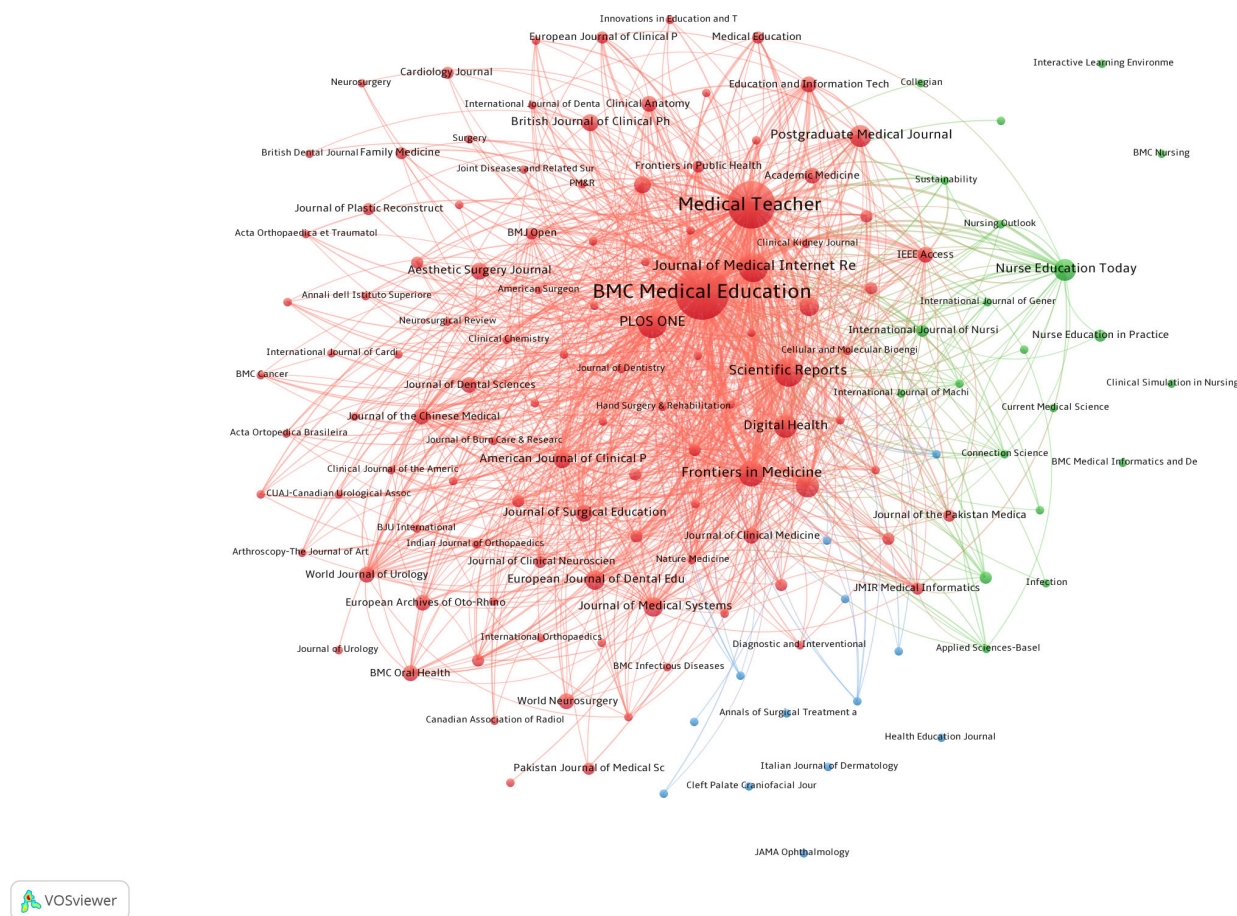


Table 5. Top 10 sources by number of cocitations and their corresponding journal impact factor (IF; Journal Citations Report [JCR] 2024) and JCR quartile.

Rank	Source	Number of cocitations	IF (JCR 2024)	JCR quartile
1	<i>arXiv</i>	692	None	None
2	<i>JMIR Medical Education</i>	542	3.2	1
3	<i>Cureus Journal of Medical Science</i>	357	1.3	2
4	<i>Medical Teacher</i>	252	4.4	1
5	<i>BMC Medical Education</i>	237	3.2	1
6	<i>Journal of Medical Internet Research</i>	216	6.0	1
7	<i>PLOS Digital Health</i>	191	7.7	1
8	<i>Nature</i>	188	48.5	1
9	<i>Academic Medicine</i>	187	5.2	1
10	<i>medRxiv</i>	164	None	None

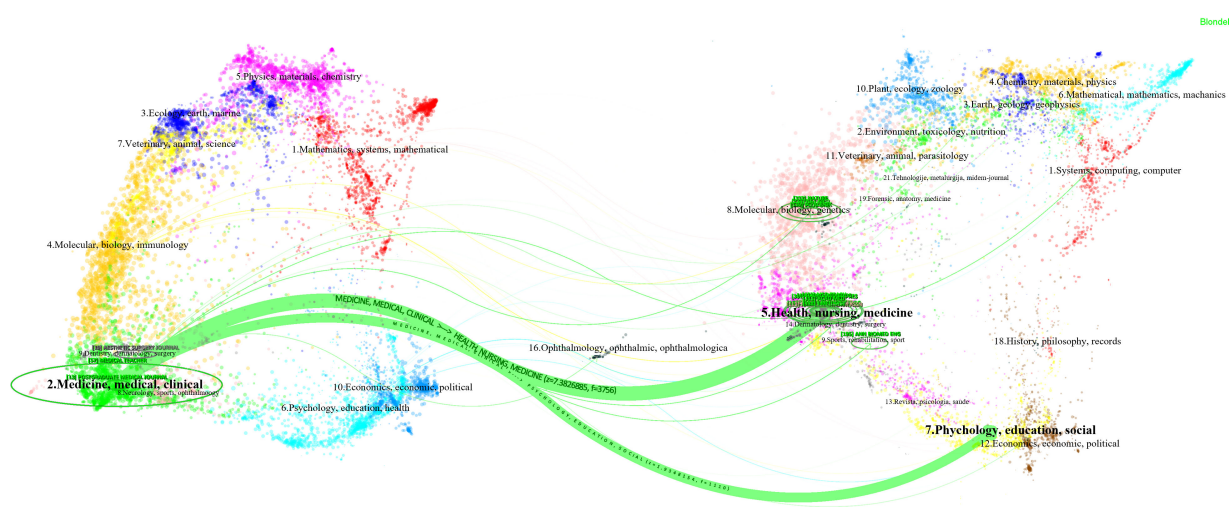
The visualization in VOSviewer showed the journals that have published articles on ChatGPT in medical education and the relationship between them. On the basis of the similarity between journals, they were divided into 3 categories: the red cluster focused on educational, technical, and basic research (eg, *BMC Medical Education* and *Medical Teacher*); the green cluster focused on the discipline of nursing and extended to nursing education, clinical simulation training, and health care informatics (eg, *Nurse Education Today* and *International Journal of Nursing Studies*); and the blue cluster focused on specialized clinical practice.

particularly in surgery, obstetrics and gynecology, orthopedics, and ophthalmology (eg, *American Journal of Obstetrics and Gynecology* and *Cleft Palate Craniofacial Journal*).

Journals were among the most common sources for publishing research results. Figure 7 presents a dual-map overlay of all academic journals, illustrating the citation paths of various subject areas. This dual-map overlay consists of 2 base maps: one on the left for citing journals and one on the right for cited journals [45]. The disciplines represented by the citing journals are indicated by the labels on the left side

of the dual map, whereas the disciplines of the cited journals are shown on the right [46].

Figure 7. The dual-map overlay of journals.



Notably, 2 primary citation trajectories can be identified on the map. They are from health, nursing, and medicine to medicine, medical, and clinical (green) and from psychology, education, and social to medicine, medical, and clinical (green). In addition, the leading-edge research results were mainly distributed in 2 (medicine, medical, and clinical). This suggests that research on ChatGPT in medical education is particularly active within these disciplines. In contrast, the knowledge base that frontier researchers rely on primarily stems from 5 (health, nursing, and medicine) and 7 (psychology, education, and social). Citing journals and cited journals were both from the field of medicine. This implies that the application of ChatGPT in fields where medical education intersects with other disciplines is still limited.

Analysis of the Author Collaboration Network Graph

Analyzing the coauthorship network of this study helped identify potential collaborators and authoritative figures in

the field. The author with the highest number of publications was Yavuz Selim Kiyak (Table 6). He has formed a stable core collaboration group with Isil Irem Budakoglu and Ozlem Coskun, who are from the same institution (Figure 8). All 3 authors ranked among the top 10 in terms of publication volume. However, among the top 10 most prolific authors, the remaining 7 have each formed independent teams. The analysis of author collaboration suggests that most academic research was conducted in independent teams without cross-team communication. Therefore, large interinstitutional collaboration networks have yet to be established.

Table 6. Top 10 authors by number of publications and their institutions and total link strength (N=407).

Rank	Author	Publications, n (%)	Institution	Total link strength
1	Yavuz Selim Kiyak	8 (2)	Gazi University (Türkiye)	1372
2	Ken Masters	5 (1.2)	Sultan Qaboos University (Oman)	945
3	Isil Irem Budakoglu	4 (1)	Gazi University (Türkiye)	1081
4	Ozlem Coskun	4 (1)	Gazi University (Türkiye)	837
5	Chia-Hung Kao	4 (1)	China Medical University (China)	641
6	Michael Alfertshofer	3 (0.7)	Ludwig Maximilian University of Munich (Germany)	774
7	Shuji Awano	3 (0.7)	Kyushu Dental University (Japan)	748
8	Olena Bolgova	3 (0.7)	Alfaisal University (Saudi Arabia)	735
9	Tzeng-Ji Chen	3 (0.7)	Taipei Veterans General Hospital, Hsinchu Branch (China)	1041
10	Wisit Cheungpasitporn	3 (0.7)	Mayo Clinic (United States)	1087

Figure 8. Collaborative network visualization of authors in VOSviewer.



Co-citation refers to the situation in which different authors are cited by the same article. These authors then form a co-citation relationship. The increase in co-citation counts indicates a greater degree of similarity among different authors' research, with the analysis itself reflecting the research strength of the respective authors. Table 7 lists the

top 10 authors in terms of cocitation frequency. The most frequently cocited authors were Tiffany H Kung (n=176), Aidan Gilson (n=141), Malik Sallam (n=109), and Arun James Thirunavukarasu (n=58). It is noteworthy that Kung is highly influential in the field of research on ChatGPT in medical education.

Table 7. Top 10 authors by number of citations and their institutions and total link strength.

Rank	Author	Number of citations	Institution	Total link strength
1	Tiffany H Kung	176	Harvard Medical School (United States)	1945
2	Aidan Gilson	141	Yale School of Medicine (United States)	1577
3	Malik Sallam	109	The University of Jordan (Jordan)	1264
4	Arun James Thirunavukarasu	58	University of Cambridge (United Kingdom)	773
5	Gunther Eysenbach	53	JMIR Publications (Canada)	645
6	Hyunsu Lee	49	Keimyung University (South Korea)	775
7	Karan Singhal	49	Google Research (United States)	826
8	Yavuz Selim Kiyak	42	Gazi University (Türkiye)	389
9	Andrew Mihalache	42	University of Western Ontario (Canada)	549
10	Rehan Ahmed Khan	40	Riphah International University (Pakistan)	502

Keyword Analysis of Global Research

To provide an overview of the primary content of the articles, it is possible to use keywords to analyze the frontiers of research on ChatGPT in medical education. Table 8 lists the

top 20 keywords by frequency. The most frequent keyword was “artificial intelligence (AI),” followed by “ChatGPT,” “medical education,” “large language models,” and “generative AI.” The keyword co-occurrence network was visualized

using VOSviewer, where the connecting lines between different keywords indicate that they have co-occurrence relationships (Figure 9). The keywords that make up this network were categorized into 4 clusters. The keywords in the red cluster in Figure 9 were related to the foundational elements of medical education, core disciplines, and ethical issues, such as “academic writing,” “radiology,” “ethics-medical,” and “machine learning.” The keywords in the green cluster focused on assessment methods, exam systems, and clinical decision-making processes in medical education, such

as “medical exam,” “clinical decision-making,” and “teaching and learning.” The keywords in the blue cluster focused on specific applications, challenges, and practical effects of generative AI in medical practice, medical education, and specialties, such as “clinical practice,” “nursing education,” and “clinical skills.” The keywords in the yellow cluster were related to the evaluation of different generative AI models and multiple research methods and practices, such as “google bard,” “meta-analysis,” and “diagnosis.”

Table 8. Top 20 keywords with the highest frequency of occurrence and their corresponding total link strength.

Rank	Keyword	Number of occurrences	Total link strength
1	“AI”	225	715
2	“ChatGPT”	216	676
3	“Medical education”	107	336
4	“Large language models”	101	338
5	“Generative AI”	29	108
6	“Education”	28	102
7	“Chatbot”	26	86
8	“Natural language processing”	22	110
9	“Medical student”	20	56
10	“Machine learning”	18	74
11	“Healthcare”	12	62
12	“Ethics”	9	48
13	“Gemini”	9	39
14	“Clinical decision-making”	9	26
15	“Medical exam”	9	48
16	“Bard”	8	40
17	“Nursing”	8	31
18	“Assessment”	8	40
19	“Clinical reasoning”	8	20
20	“Exam”	8	34

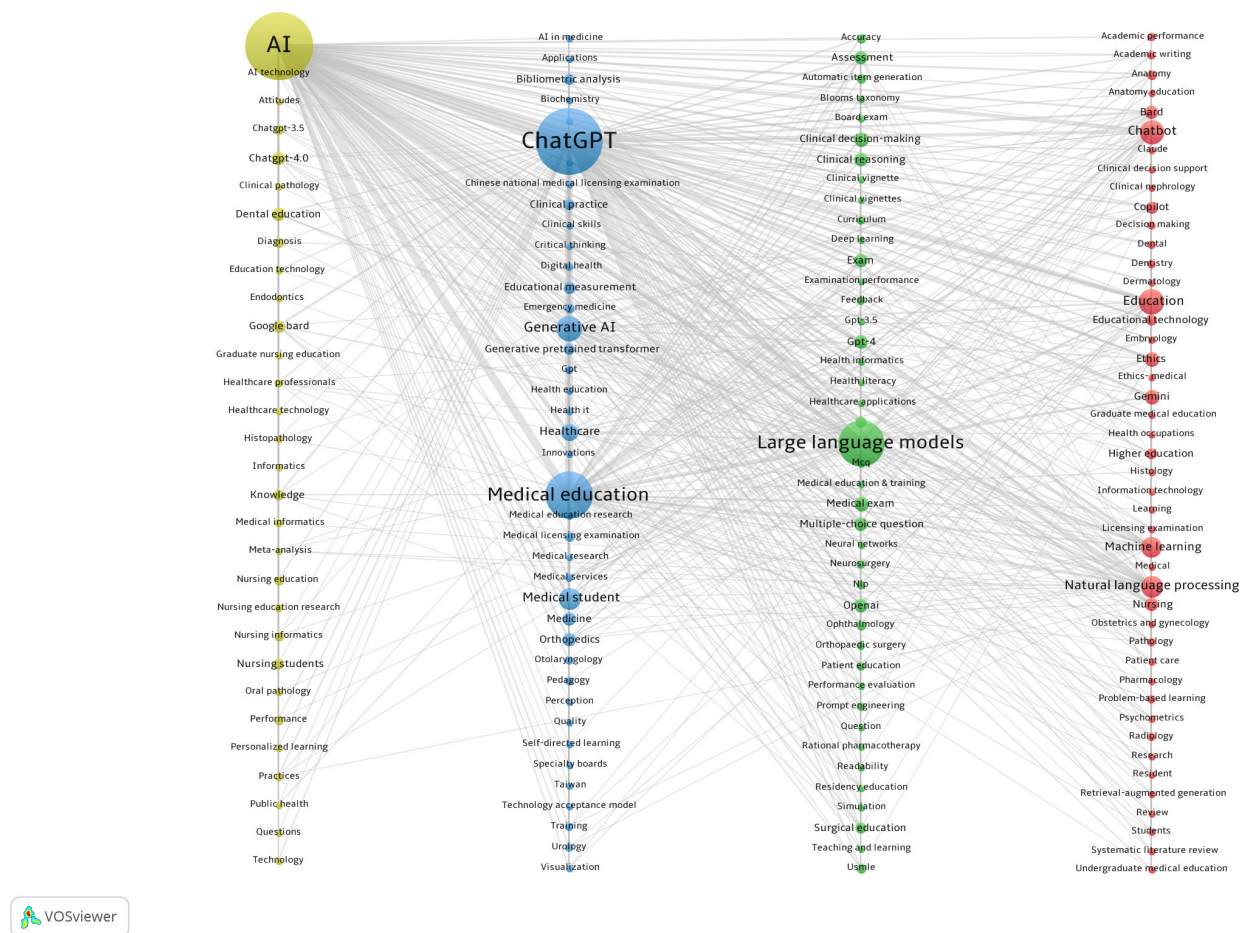
Figure 9. The co-occurrence of keywords in VOSviewer.

Figure 10 shows the monthly prevalence of the keywords from March 2023 to June 2025. Keywords such as “educational evaluation” and “medical disciplines” were research hot spots in 2023. Studies on GPT-4 continued throughout 2024. By 2025, research expanded to other LLMs such as Google Bard, Gemini, and Copilot. Research on medical

exams continued from 2024 to 2025. Figure 11 shows the cumulative frequency of keywords between March 2023 and June 2025. Although research on clinical decision-making began in April 2023, related studies increased nearly in 2025. Ethics, starting in early 2024, rapidly became a research hot spot in a relatively short period.

Figure 10. Monthly distribution heat map of keywords in Bibliometrix.

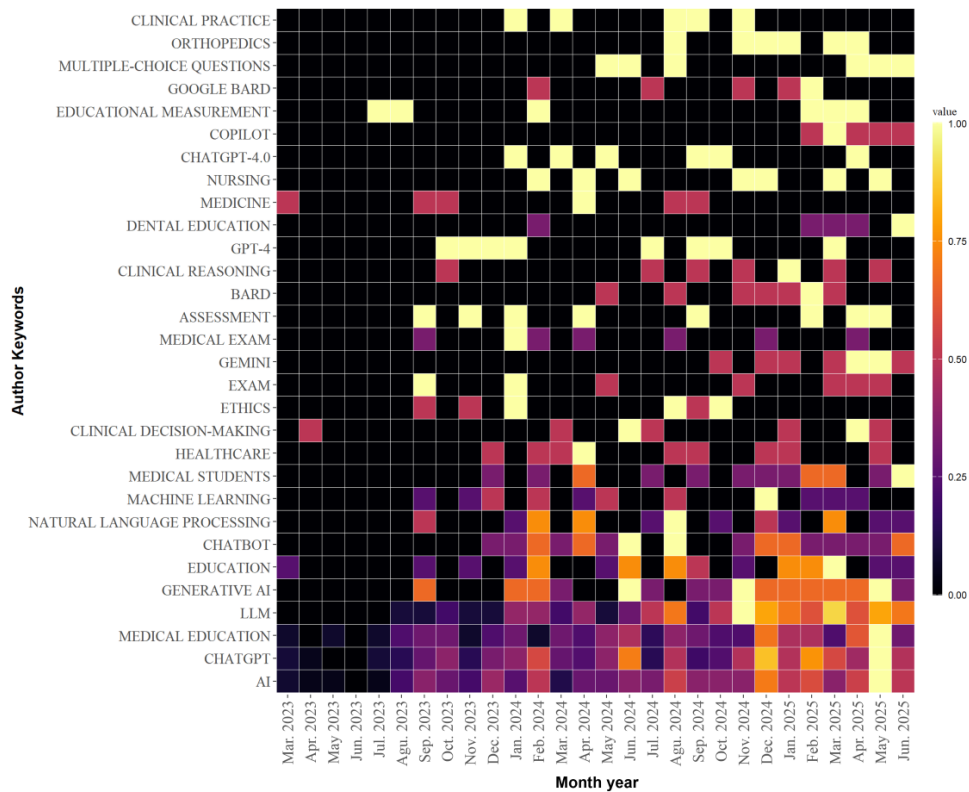
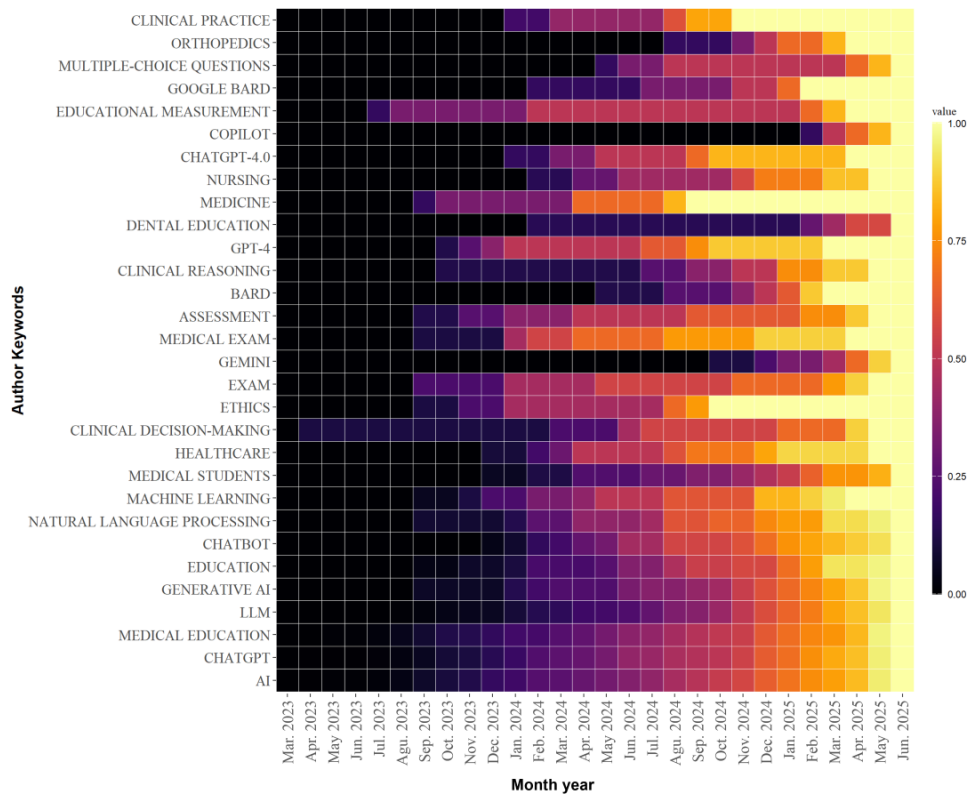


Figure 11. Cumulative distribution heat map of keywords in Bibliometrix.



Characteristics of Cited Research Articles

Table 9 lists the top 10 articles in terms of citation frequency. The most frequently cited article was “Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical

Education Using Large Language Models” [47] (n=169). The second most cited article was “How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment” [48] (n=125). Both articles were about ChatGPT’s participation in

the United States Medical Licensing Examination (USMLE). They evaluated ChatGPT’s performance on the USMLE, reflecting strong researcher interest in AI’s exam capabilities during the 2023 to 2025 study period.

Table 9. Top 10 most cited references.

Rank	Article title	Source	Authors	Year	Number of citations
1	“Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models”	<i>PLOS Digital Health</i>	Kung et al [47]	2023	169
2	“How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment”	<i>JMIR Medical Education</i>	Gilson et al [48]	2023	125
3	“ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns”	<i>Healthcare (Basel)</i>	Malik Sallam [49]	2023	63
4	“The Rise of ChatGPT: Exploring Its Potential in Medical Education”	<i>Anatomical Sciences Education</i>	Hyunsu Lee [50]	2024	51
5	“The Role of ChatGPT, Generative Language Models, and Artificial Intelligence in Medical Education: A Conversation With ChatGPT and a Call for Papers”	<i>JMIR Medical Education</i>	Gunther Eysenbach [19]	2023	51
6	“Large Language Models in Medicine”	<i>Nature Medicine</i>	Thirunavukarasu et al [51]	2023	44
7	“ChatGPT - Reshaping Medical Education and Clinical Management”	<i>Pakistan Journal of Medical Sciences</i>	Khan et al [52]	2023	39
8	“Artificial Hallucinations in ChatGPT: Implications in Scientific Writing”	<i>Cureus Journal of Medical Science</i>	Alkaissi et al [53]	2023	36
9	“ChatGPT in Medicine: An Overview of Its Applications, Advantages, Limitations, Future Prospects, and Ethical Considerations”	<i>Frontiers in Artificial Intelligence</i>	Dave et al [54]	2023	35
10	“Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine”	<i>New England Journal of Medicine</i>	Lee et al [55]	2023	34

Table 10 shows the top 20 references with the strongest citation bursts. The first citation burst occurred in 2023. This was a study comparing ChatGPT and Korean medical students on a parasitology exam. An article published in 2023, titled “Chat Generative Pretrained Transformer Fails the Multiple-Choice American College of Gastroenterology Self-Assessment Test” [56], experienced a citation burst in 2024, with the burst lasting until 2025. Researchers have continued to study ChatGPT’s ability to pass medical exams; the types of exams have ranged from the USMLE to basic subject exams. Furthermore, researchers have begun to

evaluate ChatGPT’s performance on different types of exam questions.

Table 10. Top 20 references with the strongest citation bursts.

Title	First author	Source	IF2024	JCR	Publication type	Publication year	Strength	Begin	End	2023-2025
Are ChatGPT's Knowledge and Interpretation Ability Comparable to Those of Medical Students in Korea for Taking a Parasitology Examination?: A Descriptive Study	Sun Huh [57]	Journal of Educational Evaluation for Health Professions	3.7	Q1	Article	2023	3.97	2023	2023	
Will ChatGPT Transform Healthcare?	No authors listed	Nature Medicine	50	Q1	Editorial	2023	3.91	2023	2023	
ChatGPT Passing USMLE Shines a Spotlight on the Flaws of Medical Education	Amarachi B Mbakwe [58]	PLOS Digital Health	7.7	Q1	Editorial	2023	3.79	2023	2023	
ChatGPT: the Future of Discharge Summaries?	Sajan B Patel [59]	The Lancet Digital Health	24.1	Q1	Comment	2023	3.54	2023	2023	
ChatGPT for Clinical Vignette Generation, Revision, and Evaluation	James RA Benoit [60]	medRxiv	None	None	Article	2023	3.35	2023	2023	
Abstracts Written by ChatGPT Fool Scientists	Holly Else [61]	Nature	48.5	Q1	Article	2023	3.09	2023	2023	
Benefits, Limits, and Risks of GPT-4 as an AI Chatbot for Medicine	Peter Lee [55]	The New England Journal of Medicine	78.5	Q1	Review	2023	3.01	2023	2023	
Tools Such as ChatGPT Threaten Transparent Science; Here Are Our Ground Rules for Their Use	No authors listed	Nature	48.5	Q1	Editorial	2023	2.23	2023	2023	
Could AI Help You to Write Your Next Paper?	Matthew Hutson [62]	Nature	48.5	Q1	Review	2022	2.23	2023	2023	
Appropriateness of Cardiovascular Disease Prevention Recommendations Obtained From a Popular Online Chat-Based Artificial Intelligence Model	Ashish Sarraju [63]	JAMA-Journal of the American Medical Association	55	Q1	Article	2023	2.23	2023	2023	

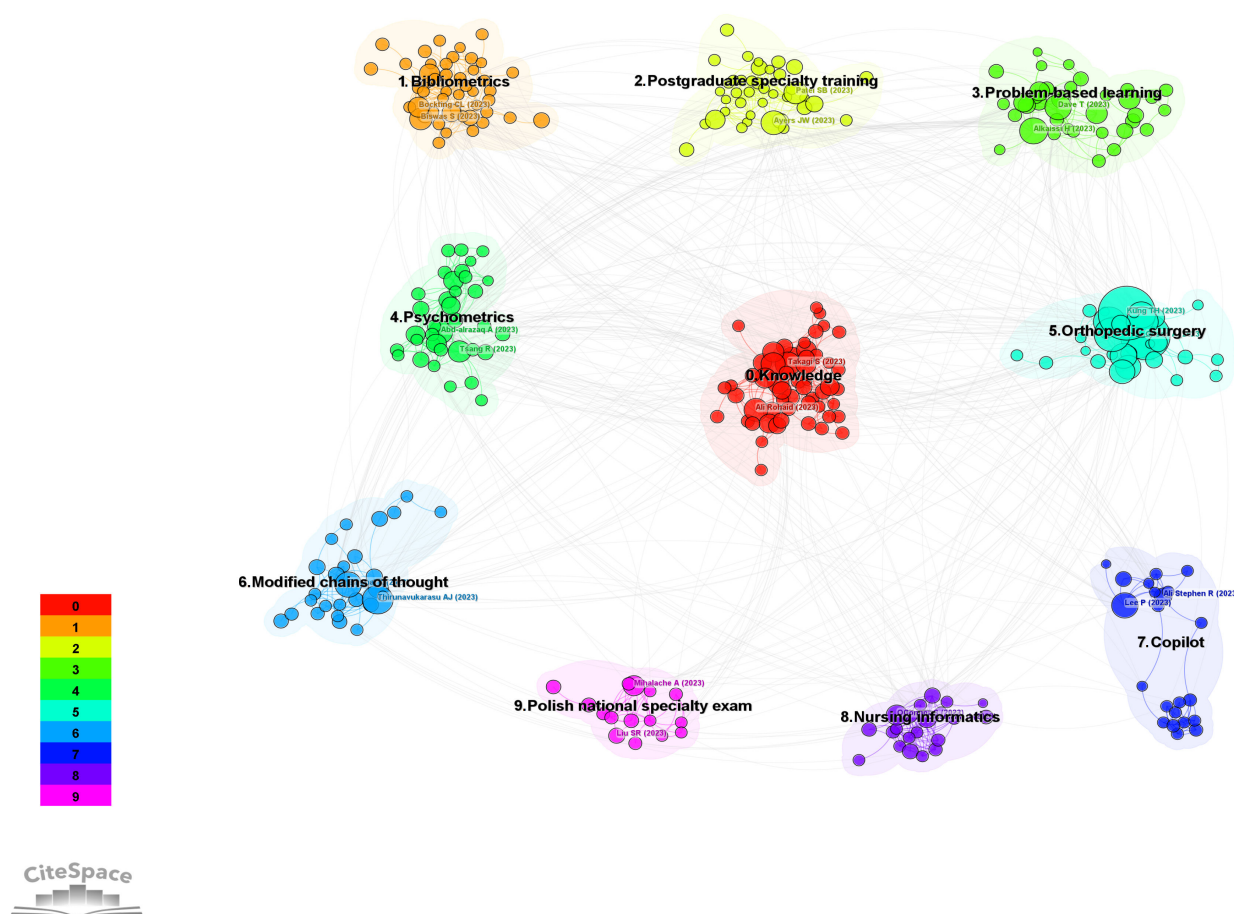
Title	First author	Source	IF2024	JCR	Publication type	Publication year	Strength	Begin	End	2023-2025
Medical Education Trends for Future Physicians in the Era of Advanced Technology and Artificial Intelligence: An Integrative Review	Eui-Ryoung Han [64]	BMC Medical Education	3.2	Q1	Review	2019	2.23	2023	2023	<div><div></div><div></div><div></div></div>
The Exciting Potential for ChatGPT in Obstetrics and Gynecology	Amos Grünebaum [65]	American Journal of Obstetrics and Gynecology	8.4	Q1	Article	2023	2.23	2023	2023	<div><div></div><div></div><div></div></div>
ChatGPT: Not All Languages Are Equal	Mohamed L Seghier [66]	Nature	48.5	Q1	Comment	2023	2.23	2023	2023	<div><div></div><div></div><div></div></div>
GPT Takes the Bar Exam	Michael James Bonmarito [67]	arXiv	None	None	Article	2022	2.23	2023	2023	<div><div></div><div></div><div></div></div>
Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models	Tiffany H Kung [47]	PLOS Digital Health	7.7	Q1	Article	2023	2.13	2023	2023	<div><div></div><div></div><div></div></div>
ChatGPT: Five Priorities for Research	Eva AM van Dis [8]	Nature	48.5	Q1	Comment	2023	2.13	2023	2023	<div><div></div><div></div><div></div></div>
Chat Generative Pretrained Transformer Fails the Multiple-Choice American College of Gastroenterology Self-Assessment Test	Kelly Suchman [56]	American Journal of Gastroenterology	7.6	Q1	Article	2023	1.8	2024	2025	<div><div></div><div></div><div></div></div>
Capabilities of GPT-4 on Medical Challenge Problems	Harsha Nori [68]	arXiv	None	None	Article	2023	1.67	2023	2023	<div><div></div><div></div><div></div></div>
Domain-Specific Language Model Pretraining for Biomedical Natural Language Processing	Yu Gu [69]	ACM Transactions on Computing for Healthcare	8	Q1	Article	2022	1.67	2023	2023	<div><div></div><div></div><div></div></div>
Natural Language Processing: State of the Art, Current Trends and Challenges	Diksha Khurana [70]	Multimedia Tools and Applications	3	Q3	Review	2023	1.67	2023	2023	<div><div></div><div></div><div></div></div>

The process of clustering analysis was conducted based on the relevance between documents, with the result that the literature was divided into 9 categories (Figure 12), each of which was identified with a different color. The category with the highest number of publications was cluster 0. The term “knowledge” was a frequently occurring keyword in this category, indicating a concentration of studies evaluating ChatGPT’s medical knowledge. This finding is related to the

research themes of the highly cited articles and burst articles. The cluster evolution shows that cluster 0 (knowledge) originated from clusters 2 (postgraduate specialty training) and 3 (problem-based learning), later developing into cluster 5 (orthopedic surgery). This progression reflects a research shift from foundational knowledge toward clinical applications. It is noteworthy that, after Microsoft released Copilot in May 2023, studies expanded to cluster 7 (Copilot).

Figure 12. Clustering of references based on similarity.

CiteSpace, v. 6.3.R1 (64-bit) Advanced
September 25, 2025, 3:08:59 PM CST
WoS: E:\medical education\data
Timespan: 2023-2025 (Slice Length=1)
Selection Criteria: g-index (k=25), LRF=3.0, L/N=10, LB=5, e=1.0
Network: N=344, E=1379 (Density=0.0234)
Largest 5 CCs: 336 (97%)
Nodes Labeled: 1.0%
Pruning: Pathfinder
Modularity Q=0.5931
Weighted Mean Silhouette S=0.8091
Harmonic Mean(Q, S)=0.6845



Discussion

Principal Findings

We analyzed Web of Science literature on ChatGPT in medical education using VOSviewer and CiteSpace. The bibliometric results showed that researchers from 66 countries have participated in this field. The United States was the most prolific contributor, with 31% (126/407) of the published papers included in this study. Furthermore, the University of California system was ranked first with 3.4% (14/407) of the publications, reflecting the sustained investment of the United States in this area. Networks of countries and institutions have

been established. This confirms active global communication and cooperation on the application of ChatGPT in medical education. Current research collaborations facilitate deeper international exchange in this field.

Notable contributors in this field include Yavuz Selim Kiyak, Ken Masters, Isil Irem Budakoglu, and Ozlem Coskun, all of whom are from academic institutions. Notably, 2 of the top 10 authors with the most publications are from hospitals or clinics. This indicates that clinicians are increasingly paying attention to the application of ChatGPT in clinical education. The results of the journal analysis showed that both citing and cited journals were from the field of medicine, and there was no emerging trend toward

interdisciplinary research. The results of keyword analysis showed a shift in research focus from broader topics such as medical education assessment and medical exams to specific clinical disciplines such as dentistry and nephrology.

Comparison to the Literature

Our research findings are consistent with the existing literature. Research on ChatGPT in medical education primarily focuses on “medical knowledge,” “educational assessment,” “clinical decision-making,” and “exam performance” [50,71,72], with increasing attention to ethics concerns [73]. Previous studies have documented the increase in publications in the field, as well as the central role played by the United States [74]. Analyzing collaboration networks and journals may provide researchers in this field with a comprehensive understanding of institutional collaboration and journal publication information. Moreover, through thematic analysis, the evolution of research topics and recent research hot spots in this field were revealed.

Implications of the Findings

The findings of this study have a number of implications. The considerable increase in research on ChatGPT in medicine indicates its extensive integration into medical education and clinical practice. The strong international collaboration suggests that research outcomes related to ChatGPT in medical education are being shared worldwide. Authors conduct research in multiple small and isolated groups. This means that researchers in this field tend to work in independent teams and lack communication across different teams. The results of the dual-map overlay of the journals showed that both the citing and cited journals were from the field of medicine, indicating that research on ChatGPT in medical education has not yet been integrated with other disciplines, failing to form a cross-disciplinary trend.

Research themes evolved from preclinical education to clinical practice simulation, confirmed via keyword and citation analysis. As a learning tool, ChatGPT has passed medical licensing exams in the United States [75,76], India [77], the United Kingdom [78], and South Korea [79]. Its applications have expanded from basic exams to specialty tests, including the American Orthopaedic In-Training Examination [80], the Membership of the Royal Colleges of Physicians of the United Kingdom exam [81], and the Chinese Critical Care Examination [82]. Researchers test ChatGPT’s medical knowledge through multiple exams. Nowadays, assessing the feasibility of ChatGPT as a learning tool is a research hot spot. Concurrently with the release of LLMs such as Microsoft Copilot, Google Gemini, and China’s DeepSeek, studies have begun to compare different models’ exam performance [83-85]. However, as LLMs keep evolving, we urgently need rigorous evidence of their reliability in medical testing. This proof remains essential before medical students fully adopt these learning tools.

In clinical teaching, ChatGPT has been used to emulate a range of clinical scenarios, including undiagnosed diabetes, kidney injury, and ophthalmic diseases [18-20]. This

creates an interactive clinical reasoning environment for students, enhancing engagement during learning. However, it is important to note that ChatGPT sometimes generates inaccurate or fabricated information [86,87]. Medical educators and students need a clear understanding of ChatGPT’s capabilities and limitations across medical specialties to effectively use AI tools for teaching and learning.

ChatGPT’s integration into medical education raises ethical issues, highlighted by our keyword analysis. Researchers are concerned that it could unintentionally reveal a patient’s personal information [88]. However, little research has been conducted on ChatGPT in medical ethics education (eg, medical ethics courses) and its educational impact [89]. While many studies have evaluated ChatGPT’s performance in medical licensing exams across various countries, research on its ability to address medical ethics issues remains limited. The 2024 study by Danehy et al [90] showed that GPT-3.5 and ChatGPT-4 performed worse on ethics questions than on medical knowledge questions. This suggests that ChatGPT’s training emphasizes medical knowledge over medical ethics. This training bias may be a potential trigger for ethical controversies involving ChatGPT in clinical practice.

Study Strengths and Limitations

This study has both strengths and limitations. To our knowledge, this is the first study to use bibliometric analysis to study the use of ChatGPT in medical education rather than general medicine. Furthermore, the visualization of quantitative results provides a comprehensive understanding of the current status of publications, research hot spots, and development trends related to ChatGPT in medical education.

Despite best efforts to include all the relevant terms and terminology in the literature search, some relevant papers may have been omitted. The search was confined to Web of Science, and only research articles written in English were included, with articles in other languages not being considered. In addition, due to the ongoing nature of the research, recent high-quality studies may not have been included.

Subsequently, the discussion focus on providing strong evidence to demonstrate the feasibility of ChatGPT as a learning tool, evaluating ChatGPT’s medical ethics awareness in medical education, and offering evidence to support the application of ChatGPT in medical ethics.

Conclusions

In conclusion, this bibliometric analysis of ChatGPT in medical education reveals characteristics such as rapid publication growth, concentrated contributions from leading countries and institutions, decentralized author networks, and evolving thematic focuses. It will be crucial to enhance institution collaboration and cross-team partnerships in the future. This will promote the application potential of ChatGPT in various fields of medical education. Improving the effectiveness of ChatGPT is expected to provide educators and students with a more efficient medical teaching and learning process.

Acknowledgments

This study was supported by the Jiangxi Province Education Science 14th Five-Year Plan Project (grant 22QN048).

Authors' Contributions

YZ designed the study, collected and analyzed the data, and wrote the manuscript. XX assisted with data collection and analysis. QX contributed to the methodology, provided writing and editing support, and supervised the project.

Conflicts of Interest

None declared.

References

1. Floridi L, Chiriatti M. GPT-3: its nature, scope, limits, and consequences. *Minds Mach.* Nov 2020;30:681-694. [doi: [10.1007/s11023-020-09548-1](https://doi.org/10.1007/s11023-020-09548-1)]
2. Bedi S, Liu Y, Orr-Ewing L, et al. Testing and evaluation of health care applications of large language models: a systematic review. *JAMA.* Jan 28, 2025;333(4):319-328. [doi: [10.1001/jama.2024.21700](https://doi.org/10.1001/jama.2024.21700)] [Medline: [39405325](https://pubmed.ncbi.nlm.nih.gov/39405325/)]
3. ChatGPT. OpenAI. URL: <https://chat.openai.com/> [Accessed 2025-09-26]
4. Ali K, Barhom N, Tamimi F, Duggal M. ChatGPT-a double-edged sword for healthcare education? Implications for assessments of dental students. *Eur J Dent Educ.* Feb 2024;28(1):206-211. [doi: [10.1111/eje.12937](https://doi.org/10.1111/eje.12937)] [Medline: [37550893](https://pubmed.ncbi.nlm.nih.gov/37550893/)]
5. Tian S, Jin Q, Yeganova L, et al. Opportunities and challenges for ChatGPT and large language models in biomedicine and health. *Brief Bioinform.* Nov 22, 2023;25(1):bbad493. [doi: [10.1093/bib/bbad493](https://doi.org/10.1093/bib/bbad493)]
6. Sharma S, Pajai S, Prasad R, et al. A critical review of ChatGPT as a potential substitute for diabetes educators. *Cureus.* May 1, 2023;15(5):e38380. [doi: [10.7759/cureus.38380](https://doi.org/10.7759/cureus.38380)] [Medline: [37265899](https://pubmed.ncbi.nlm.nih.gov/37265899/)]
7. Jansen BJ, Jung SG, Salminen J. Employing large language models in survey research. *Nat Lang Proc J.* Sep 2023;4:100020. [doi: [10.1016/j.nlp.2023.100020](https://doi.org/10.1016/j.nlp.2023.100020)]
8. van Dis EA, Bollen J, Zuidema W, van Rooij R, Bockting CL. ChatGPT: five priorities for research. *Nature.* Feb 2023;614(7947):224-226. [doi: [10.1038/d41586-023-00288-7](https://doi.org/10.1038/d41586-023-00288-7)] [Medline: [36737653](https://pubmed.ncbi.nlm.nih.gov/36737653/)]
9. Haleem A, Javaid M, Singh RP. An era of ChatGPT as a significant futuristic support tool: a study on features, abilities, and challenges. *BenchCouncil Trans Benchmarks Stand Eval.* Oct 2022;2(4):100089. [doi: [10.1016/j.tbench.2023.100089](https://doi.org/10.1016/j.tbench.2023.100089)]
10. Ignjatović A, Stevanović L. Efficacy and limitations of ChatGPT as a biostatistical problem-solving tool in medical education in Serbia: a descriptive study. *J Educ Eval Health Prof.* 2023;20:28. [doi: [10.3352/jeehp.2023.20.28](https://doi.org/10.3352/jeehp.2023.20.28)] [Medline: [37840252](https://pubmed.ncbi.nlm.nih.gov/37840252/)]
11. Abujaber AA, Abd-Alrazaq A, Al-Qudimat AR, Nashwan AJ. A strengths, weaknesses, opportunities, and threats (SWOT) analysis of ChatGPT integration in nursing education: a narrative review. *Cureus.* Nov 11, 2023;15(11):e48643. [doi: [10.7759/cureus.48643](https://doi.org/10.7759/cureus.48643)] [Medline: [38090452](https://pubmed.ncbi.nlm.nih.gov/38090452/)]
12. Liu J, Liu F, Fang J, Liu S. The application of chat generative pre-trained transformer in nursing education. *Nurs Outlook.* 2023;71(6):102064. [doi: [10.1016/j.outlook.2023.102064](https://doi.org/10.1016/j.outlook.2023.102064)] [Medline: [37879261](https://pubmed.ncbi.nlm.nih.gov/37879261/)]
13. Wu Y, Zheng Y, Feng B, Yang Y, Kang K, Zhao A. Embracing ChatGPT for medical education: exploring its impact on doctors and medical students. *JMIR Med Educ.* Apr 10, 2024;10:e52483. [doi: [10.2196/52483](https://doi.org/10.2196/52483)] [Medline: [38598263](https://pubmed.ncbi.nlm.nih.gov/38598263/)]
14. Jeyaraman M, K SP, Jeyaraman N, Nallakumarasamy A, Yadav S, Bondili SK. ChatGPT in medical education and research: a boon or a bane? *Cureus.* Aug 29, 2023;15(8):e44316. [doi: [10.7759/cureus.44316](https://doi.org/10.7759/cureus.44316)] [Medline: [37779749](https://pubmed.ncbi.nlm.nih.gov/37779749/)]
15. Wang C, Li S, Lin N, et al. Application of large language models in medical training evaluation-using ChatGPT as a standardized patient: multimetric assessment. *J Med Internet Res.* Jan 1, 2025;27:e59435. [doi: [10.2196/59435](https://doi.org/10.2196/59435)] [Medline: [39742453](https://pubmed.ncbi.nlm.nih.gov/39742453/)]
16. Wu Z, Li S, Zhao X. The application of ChatGPT in medical education: prospects and challenges. *Int J Surg.* Jan 1, 2025;111(1):1652-1653. [doi: [10.1097/JS9.0000000000001887](https://doi.org/10.1097/JS9.0000000000001887)] [Medline: [38935099](https://pubmed.ncbi.nlm.nih.gov/38935099/)]
17. Scherr R, Halaseh FF, Spina A, Andalib S, Rivera R. ChatGPT interactive medical simulations for early clinical education: case study. *JMIR Med Educ.* Nov 10, 2023;9:e49877. [doi: [10.2196/49877](https://doi.org/10.2196/49877)] [Medline: [37948112](https://pubmed.ncbi.nlm.nih.gov/37948112/)]
18. Chatterjee S, Bhattacharya M, Pal S, Lee SS, Chakraborty C. ChatGPT and large language models in orthopedics: from education and surgery to research. *J Exp Orthop.* Dec 1, 2023;10(1):128. [doi: [10.1186/s40634-023-00700-1](https://doi.org/10.1186/s40634-023-00700-1)] [Medline: [38038796](https://pubmed.ncbi.nlm.nih.gov/38038796/)]
19. 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.* Mar 6, 2023;9:e46885. [doi: [10.2196/46885](https://doi.org/10.2196/46885)] [Medline: [36863937](https://pubmed.ncbi.nlm.nih.gov/36863937/)]

20. Antaki F, Touma S, Milad D, El-Khoury J, Duval R. Evaluating the performance of ChatGPT in ophthalmology: an analysis of its successes and shortcomings. *Ophthalmol Sci*. May 5, 2023;3(4):100324. [doi: [10.1016/j.xops.2023.100324](https://doi.org/10.1016/j.xops.2023.100324)] [Medline: [37334036](https://pubmed.ncbi.nlm.nih.gov/37334036/)]
21. Gonzalez-Garcia A, Bermejo-Martinez D, Lopez-Alonso AI, Trevisson-Redondo B, Martín-Vázquez C, Perez-Gonzalez S. Impact of ChatGPT usage on nursing students education: a cross-sectional study. *Heliyon*. Dec 31, 2024;11(1):e41559. [doi: [10.1016/j.heliyon.2024.e41559](https://doi.org/10.1016/j.heliyon.2024.e41559)] [Medline: [39850430](https://pubmed.ncbi.nlm.nih.gov/39850430/)]
22. Gencer A, Aydin S. Can ChatGPT pass the thoracic surgery exam? *Am J Med Sci*. Oct 2023;366(4):291-295. [doi: [10.1016/j.amjms.2023.08.001](https://doi.org/10.1016/j.amjms.2023.08.001)] [Medline: [37549788](https://pubmed.ncbi.nlm.nih.gov/37549788/)]
23. Soulage CO, Van Coppenolle F, Guebre-Egziabher F. The conversational AI “ChatGPT” outperforms medical students on a physiology university examination. *Adv Physiol Educ*. Dec 1, 2024;48(4):677-684. [doi: [10.1152/advan.00181.2023](https://doi.org/10.1152/advan.00181.2023)] [Medline: [38991037](https://pubmed.ncbi.nlm.nih.gov/38991037/)]
24. Benítez TM, Xu Y, Boudreau JD, et al. Harnessing the potential of large language models in medical education: promise and pitfalls. *J Am Med Inform Assoc*. Feb 16, 2024;31(3):776-783. [doi: [10.1093/jamia/ocad252](https://doi.org/10.1093/jamia/ocad252)] [Medline: [38269644](https://pubmed.ncbi.nlm.nih.gov/38269644/)]
25. Kim TW. Application of artificial intelligence chatbots, including ChatGPT, in education, scholarly work, programming, and content generation and its prospects: a narrative review. *J Educ Eval Health Prof*. 2023;20:38. [doi: [10.3352/jeehp.2023.20.38](https://doi.org/10.3352/jeehp.2023.20.38)] [Medline: [38148495](https://pubmed.ncbi.nlm.nih.gov/38148495/)]
26. Sabry Abdel-Messih M, Kamel Boulos MN. ChatGPT in clinical toxicology. *JMIR Med Educ*. Mar 8, 2023;9:e46876. [doi: [10.2196/46876](https://doi.org/10.2196/46876)] [Medline: [36867743](https://pubmed.ncbi.nlm.nih.gov/36867743/)]
27. Ming S, Guo Q, Cheng W, Lei B. Influence of model evolution and system roles on ChatGPT’s performance in Chinese medical licensing exams: comparative study. *JMIR Med Educ*. Aug 13, 2024;10:e52784. [doi: [10.2196/52784](https://doi.org/10.2196/52784)] [Medline: [39140269](https://pubmed.ncbi.nlm.nih.gov/39140269/)]
28. Huang CH, Hsiao HJ, Yeh PC, Wu KC, Kao CH. Performance of ChatGPT on stage 1 of the Taiwanese medical licensing exam. *Digit Health*. Feb 16, 2024;10:20552076241233144. [doi: [10.1177/20552076241233144](https://doi.org/10.1177/20552076241233144)] [Medline: [38371244](https://pubmed.ncbi.nlm.nih.gov/38371244/)]
29. Ishida K, Hanada E. Potential of ChatGPT to pass the Japanese medical and healthcare professional national licenses: a literature review. *Cureus*. Aug 6, 2024;16(8):e66324. [doi: [10.7759/cureus.66324](https://doi.org/10.7759/cureus.66324)] [Medline: [39247019](https://pubmed.ncbi.nlm.nih.gov/39247019/)]
30. Kawahara T, Sumi Y. GPT-4/4V’s performance on the Japanese National Medical Licensing Examination. *Med Teach*. Mar 2025;47(3):450-457. [doi: [10.1080/0142159X.2024.2342545](https://doi.org/10.1080/0142159X.2024.2342545)] [Medline: [38648547](https://pubmed.ncbi.nlm.nih.gov/38648547/)]
31. Scaioli G, Lo Moro G, Conrado F, Rosset L, Bert F, Siliquini R. Exploring the potential of ChatGPT for clinical reasoning and decision-making: a cross-sectional study on the Italian Medical Residency Exam. *Ann Ist Super Sanita*. 2023;59(4):267-270. [doi: [10.4415/ANN_23_04_05](https://doi.org/10.4415/ANN_23_04_05)] [Medline: [38088393](https://pubmed.ncbi.nlm.nih.gov/38088393/)]
32. Torres-Zegarra BC, Rios-Garcia W, Ñaña-Cordova AM, et al. Performance of ChatGPT, Bard, Claude, and Bing on the Peruvian National Licensing Medical Examination: a cross-sectional study. *J Educ Eval Health Prof*. 2023;20:30. [doi: [10.3352/jeehp.2023.20.30](https://doi.org/10.3352/jeehp.2023.20.30)] [Medline: [37981579](https://pubmed.ncbi.nlm.nih.gov/37981579/)]
33. Arruda H, Silva ER, Lessa M, Proença Jr D, Bartholo R. VOSviewer and Bibliometrix. *J Med Libr Assoc*. Jul 1, 2022;110(3):392-395. [doi: [10.5195/jmla.2022.1434](https://doi.org/10.5195/jmla.2022.1434)] [Medline: [36589296](https://pubmed.ncbi.nlm.nih.gov/36589296/)]
34. Zhao X, Nan D, Chen C, Zhang S, Che S, Kim JH. Bibliometric study on environmental, social, and governance research using CiteSpace. *Front Environ Sci*. 2023;10. [doi: [10.3389/fenvs.2022.1087493](https://doi.org/10.3389/fenvs.2022.1087493)]
35. Zhou F, Zhang T, Jin Y, et al. Worldwide tinnitus research: a bibliometric analysis of the published literature between 2001 and 2020. *Front Neurol*. Jan 31, 2022;13:828299. [doi: [10.3389/fneur.2022.828299](https://doi.org/10.3389/fneur.2022.828299)] [Medline: [35173675](https://pubmed.ncbi.nlm.nih.gov/35173675/)]
36. Zhou F, Zhang T, Jin Y, et al. Unveiling the knowledge domain and emerging trends of olfactory dysfunction with depression or anxiety: a bibliometrics study. *Front Neurosci*. Sep 8, 2022;16:959936. [doi: [10.3389/fnins.2022.959936](https://doi.org/10.3389/fnins.2022.959936)] [Medline: [36161166](https://pubmed.ncbi.nlm.nih.gov/36161166/)]
37. Zhou Q, Pei J, Poon J, et al. Worldwide research trends on aristolochic acids (1957-2017): suggestions for researchers. *PLoS ONE*. May 2, 2019;14(5):e0216135. [doi: [10.1371/journal.pone.0216135](https://doi.org/10.1371/journal.pone.0216135)] [Medline: [31048858](https://pubmed.ncbi.nlm.nih.gov/31048858/)]
38. Bibilometrix. URL: <https://www.bibliometrix.org/home/> [Accessed 2025-09-29]
39. Bibliometric. URL: <https://bibliometric.com/> [Accessed 2025-09-29]
40. Synnestvedt MB, Chen C, Holmes JH. CiteSpace II: visualization and knowledge discovery in bibliographic databases. *AMIA Annu Symp Proc*. 2005;2005:724-728. [Medline: [16779135](https://pubmed.ncbi.nlm.nih.gov/16779135/)]
41. Chen C. Searching for intellectual turning points: progressive knowledge domain visualization. *Proc Natl Acad Sci U S A*. Apr 6, 2004;101 Suppl 1(Suppl 1):5303-5310. [doi: [10.1073/pnas.0307513100](https://doi.org/10.1073/pnas.0307513100)] [Medline: [14724295](https://pubmed.ncbi.nlm.nih.gov/14724295/)]
42. Xu S, Xu D, Wen L, et al. Integrating unified medical language system and Kleinberg’s burst detection algorithm into research topics of medications for post-traumatic stress disorder. *Drug Des Devel Ther*. Sep 24, 2020;14:3899-3913. [doi: [10.2147/DDDT.S270379](https://doi.org/10.2147/DDDT.S270379)] [Medline: [33061296](https://pubmed.ncbi.nlm.nih.gov/33061296/)]

43. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*. Aug 2010;84(2):523-538. [doi: [10.1007/s11192-009-0146-3](https://doi.org/10.1007/s11192-009-0146-3)] [Medline: [20585380](#)]
44. Aria M, Cuccurullo C. bibliometrix: an R-tool for comprehensive science mapping analysis. *J Informetr*. Nov 2017;11(4):959-975. [doi: [10.1016/j.joi.2017.08.007](https://doi.org/10.1016/j.joi.2017.08.007)]
45. Hou J, Yang X, Chen C. Emerging trends and new developments in information science: a document co-citation analysis (2009–2016). *Scientometrics*. May 1, 2018;115(2):869-892. [doi: [10.1007/s11192-018-2695-9](https://doi.org/10.1007/s11192-018-2695-9)]
46. Li Q, Long R, Chen H, Chen F, Wang J. Visualized analysis of global green buildings: development, barriers and future directions. *J Clean Prod*. Feb 1, 2020;245:118775. [doi: [10.1016/j.jclepro.2019.118775](https://doi.org/10.1016/j.jclepro.2019.118775)]
47. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health*. Sep 2023;2(2):e0000198. [doi: [10.1371/journal.pdig.0000198](https://doi.org/10.1371/journal.pdig.0000198)] [Medline: [36812645](#)]
48. Gilson A, Safranek CW, Huang T, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ*. Feb 8, 2023;9:e45312. [doi: [10.2196/45312](https://doi.org/10.2196/45312)] [Medline: [36753318](#)]
49. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare (Basel)*. Mar 19, 2023;11(6):887. [doi: [10.3390/healthcare11060887](https://doi.org/10.3390/healthcare11060887)] [Medline: [36981544](#)]
50. Lee H. The rise of ChatGPT: exploring its potential in medical education. *Anat Sci Educ*. 2024;17(5):926-931. [doi: [10.1002/ase.2270](https://doi.org/10.1002/ase.2270)] [Medline: [36916887](#)]
51. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med*. Aug 2023;29(8):1930-1940. [doi: [10.1038/s41591-023-02448-8](https://doi.org/10.1038/s41591-023-02448-8)] [Medline: [37460753](#)]
52. Khan RA, Jawaid M, Khan AR, Sajjad M. ChatGPT - Reshaping medical education and clinical management. *Pak J Med Sci*. ;39(2). [doi: [10.12669/pjms.39.2.7653](https://doi.org/10.12669/pjms.39.2.7653)]
53. Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*. Feb 2023;15(2):e35179. [doi: [10.7759/cureus.35179](https://doi.org/10.7759/cureus.35179)] [Medline: [36811129](#)]
54. Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Front Artif Intell*. 2023;6:1169595. [doi: [10.3389/frai.2023.1169595](https://doi.org/10.3389/frai.2023.1169595)] [Medline: [37215063](#)]
55. Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *N Engl J Med*. Mar 30, 2023;388(13):1233-9. [doi: [10.1056/NEJMs2214184](https://doi.org/10.1056/NEJMs2214184)] [Medline: [36988602](#)]
56. Suchman K, Garg S, Trindade AJ. Chat generative pretrained transformer fails the multiple-choice American College of Gastroenterology Self-Assessment Test. *Am J Gastroenterol*. Dec 1, 2023;118(12):2280-2282. [doi: [10.14309/ajg.0000000000002320](https://doi.org/10.14309/ajg.0000000000002320)] [Medline: [37212584](#)]
57. Huh S. Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: a descriptive study. *J Educ Eval Health Prof*. ;20:1. [doi: [10.3352/jeehp.2023.20.1](https://doi.org/10.3352/jeehp.2023.20.1)]
58. Mbakwe AB, Lourentzou I, Celi LA, Mechanic OJ, Dagan A. ChatGPT passing USMLE shines a spotlight on the flaws of medical education. *PLOS Digit Health*. Feb 2023;2(2):e0000205. [doi: [10.1371/journal.pdig.0000205](https://doi.org/10.1371/journal.pdig.0000205)] [Medline: [36812618](#)]
59. Patel SB, Lam K. ChatGPT: the future of discharge summaries? *Lancet Digit Health*. Mar 2023;5(3):e107-8. [doi: [10.1016/S2589-7500\(23\)00021-3](https://doi.org/10.1016/S2589-7500(23)00021-3)] [Medline: [36754724](#)]
60. Benoit JRA. ChatGPT for clinical vignette generation, revision, and evaluation. *Medical Education*. Preprint posted online on Feb 8, 2023. [doi: [10.1101/2023.02.04.23285478](https://doi.org/10.1101/2023.02.04.23285478)]
61. Else H. Abstracts written by ChatGPT fool scientists. *Nature New Biol*. Jan 19, 2023;613(7944):423. [doi: [10.1038/d41586-023-00056-7](https://doi.org/10.1038/d41586-023-00056-7)]
62. Hutson M. Could AI help you to write your next paper? *Nature New Biol*. Nov 3, 2022;611(7934):192-3. [doi: [10.1038/d41586-022-03479-w](https://doi.org/10.1038/d41586-022-03479-w)]
63. Sarraju A, Bruemmer D, Van Iterson E, Cho L, Rodriguez F, Laffin L. Appropriateness of cardiovascular disease prevention recommendations obtained from a popular online chat-based artificial intelligence model. *JAMA*. Mar 14, 2023;329(10):842-4. [doi: [10.1001/jama.2023.1044](https://doi.org/10.1001/jama.2023.1044)] [Medline: [36735264](#)]
64. Han ER, Yeo S, Kim MJ, Lee YH, Park KH, Roh H. Medical education trends for future physicians in the era of advanced technology and artificial intelligence: an integrative review. *BMC Med Educ*. Dec 2019;19(1). [doi: [10.1186/s12909-019-1891-5](https://doi.org/10.1186/s12909-019-1891-5)]
65. Grünebaum A, Chervenak J, Pollet SL, Katz A, Chervenak FA. The exciting potential for ChatGPT in obstetrics and gynecology. *Am J Obstet Gynecol*. Jun 2023;228(6):696-705. [doi: [10.1016/j.ajog.2023.03.009](https://doi.org/10.1016/j.ajog.2023.03.009)]

66. Seghier ML. ChatGPT: not all languages are equal. *Nature New Biol.* Mar 9, 2023;615(7951):216. [doi: [10.1038/d41586-023-00680-3](https://doi.org/10.1038/d41586-023-00680-3)]
67. Bommarito MJ, Katz DM. GPT takes the bar exam. *SSRN Journal.* Dec 29, 2022. [doi: [10.2139/ssrn.4314839](https://doi.org/10.2139/ssrn.4314839)]
68. Capabilities of GPT-4 on medical challenge problems. *arXiv.* Preprint posted online on Apr 12, 2023. [doi: [10.48550/arXiv.2303.13375](https://doi.org/10.48550/arXiv.2303.13375)]
69. Gu Y, Tinn R, Cheng H, et al. Domain-specific language model pretraining for biomedical natural language processing. *ACM Trans Comput Healthcare.* Jan 31, 2022;3(1):1-23. [doi: [10.1145/3458754](https://doi.org/10.1145/3458754)]
70. Khurana D, Koli A, Khatter K, Singh S. Natural language processing: state of the art, current trends and challenges. *Multimed Tools Appl.* Jan 2023;82(3):3713-44. [doi: [10.1007/s11042-022-13428-4](https://doi.org/10.1007/s11042-022-13428-4)]
71. Boscardin CK, Gin B, Golde PB, Hauer KE. ChatGPT and generative artificial intelligence for medical education: potential impact and opportunity. *Acad Med.* Jan 1, 2024;99(1):22-27. [doi: [10.1097/ACM.0000000000005439](https://doi.org/10.1097/ACM.0000000000005439)] [Medline: [37651677](https://pubmed.ncbi.nlm.nih.gov/37651677/)]
72. Tan S, Xin X, Wu D. ChatGPT in medicine: prospects and challenges: a review article. *Int J Surg.* Jun 1, 2024;110(6):3701-3706. [doi: [10.1097/JS9.0000000000001312](https://doi.org/10.1097/JS9.0000000000001312)] [Medline: [38502861](https://pubmed.ncbi.nlm.nih.gov/38502861/)]
73. Cheng Y, Zhu L. A review of ChatGPT in medical education: exploring advantages and limitations. *Int J Surg.* Jul 1, 2025;111(7):4586-4602. [doi: [10.1097/JS9.0000000000002505](https://doi.org/10.1097/JS9.0000000000002505)] [Medline: [40465793](https://pubmed.ncbi.nlm.nih.gov/40465793/)]
74. Wu J, Ma Y, Wang J, Xiao M. The application of ChatGPT in medicine: a scoping review and bibliometric analysis. *J Multidiscip Healthc.* Apr 18, 2024;17:1681-1692. [doi: [10.2147/JMDH.S463128](https://doi.org/10.2147/JMDH.S463128)] [Medline: [38650670](https://pubmed.ncbi.nlm.nih.gov/38650670/)]
75. Bicknell BT, Butler D, Whalen S, et al. ChatGPT-4 omni performance in USMLE disciplines and clinical skills: comparative analysis. *JMIR Med Educ.* Nov 6, 2024;10:e63430. [doi: [10.2196/63430](https://doi.org/10.2196/63430)] [Medline: [39504445](https://pubmed.ncbi.nlm.nih.gov/39504445/)]
76. Alfertshofer M, Knoedler S, Hoch CC, et al. Analyzing question characteristics influencing ChatGPT's performance in 3000 USMLE®-style questions. *Med Sci Educ.* Sep 28, 2024;35(1):257-267. [doi: [10.1007/s40670-024-02176-9](https://doi.org/10.1007/s40670-024-02176-9)] [Medline: [40144074](https://pubmed.ncbi.nlm.nih.gov/40144074/)]
77. Surapaneni KM. Assessing the performance of ChatGPT in medical biochemistry using clinical case vignettes: observational study. *JMIR Med Educ.* Nov 7, 2023;9:e47191. [doi: [10.2196/47191](https://doi.org/10.2196/47191)] [Medline: [37934568](https://pubmed.ncbi.nlm.nih.gov/37934568/)]
78. Lai UH, Wu KS, Hsu TY, Kan JK. Evaluating the performance of ChatGPT-4 on the United Kingdom Medical Licensing Assessment. *Front Med (Lausanne).* Sep 19, 2023;10:1240915. [doi: [10.3389/fmed.2023.1240915](https://doi.org/10.3389/fmed.2023.1240915)] [Medline: [37795422](https://pubmed.ncbi.nlm.nih.gov/37795422/)]
79. Oh N, Choi GS, Lee WY. ChatGPT goes to the operating room: evaluating GPT-4 performance and its potential in surgical education and training in the era of large language models. *Ann Surg Treat Res.* May 2023;104(5):269-273. [doi: [10.4174/astr.2023.104.5.269](https://doi.org/10.4174/astr.2023.104.5.269)] [Medline: [37179699](https://pubmed.ncbi.nlm.nih.gov/37179699/)]
80. Kung JE, Marshall C, Gauthier C, Gonzalez TA, Jackson JB 3rd. Evaluating ChatGPT performance on the orthopaedic in-training examination. *JB JS Open Access.* 2023;8(3):e23.00056. [doi: [10.2106/JBJS.OA.23.00056](https://doi.org/10.2106/JBJS.OA.23.00056)] [Medline: [37693092](https://pubmed.ncbi.nlm.nih.gov/37693092/)]
81. Maitland A, Fowkes R, Maitland S. Can ChatGPT pass the MRCP (UK) written examinations? Analysis of performance and errors using a clinical decision-reasoning framework. *BMJ Open.* Mar 15, 2024;14(3):e080558. [doi: [10.1136/bmjopen-2023-080558](https://doi.org/10.1136/bmjopen-2023-080558)] [Medline: [38490655](https://pubmed.ncbi.nlm.nih.gov/38490655/)]
82. Wang X, Tang J, Feng Y, Tang C, Wang X. Can ChatGPT-4 perform as a competent physician based on the Chinese critical care examination? *J Crit Care.* Apr 2025;86:155010. [doi: [10.1016/j.jcrc.2024.155010](https://doi.org/10.1016/j.jcrc.2024.155010)] [Medline: [40023616](https://pubmed.ncbi.nlm.nih.gov/40023616/)]
83. Thesen T, Tuan RL, Blumer J, Lee MW. LLM-based generation of USMLE-style questions with ASPET/AMSPC knowledge objectives: all RAGs and no riches. *Br J Clin Pharmacol.* Jun 8, 2025. [doi: [10.1002/bcp.70119](https://doi.org/10.1002/bcp.70119)] [Medline: [40483567](https://pubmed.ncbi.nlm.nih.gov/40483567/)]
84. Camarata T, McCoy L, Rosenberg R, Temprine Grellinger KR, Brettschnieder K, Berman J. LLM-generated multiple choice practice quizzes for preclinical medical students. *Adv Physiol Educ.* Sep 1, 2025;49(3):758-763. [doi: [10.1152/advan.00106.2024](https://doi.org/10.1152/advan.00106.2024)] [Medline: [40516963](https://pubmed.ncbi.nlm.nih.gov/40516963/)]
85. Yang H, Li M, Zhou H, et al. Large language model synergy for ensemble learning in medical question answering: design and evaluation study. *J Med Internet Res.* Jul 14, 2025;27:e70080. [doi: [10.2196/70080](https://doi.org/10.2196/70080)] [Medline: [40658884](https://pubmed.ncbi.nlm.nih.gov/40658884/)]
86. Barrington NM, Gupta N, Musmar B, et al. A bibliometric analysis of the rise of ChatGPT in medical research. *Med Sci (Basel).* Sep 17, 2023;11(3):61. [doi: [10.3390/medsci11030061](https://doi.org/10.3390/medsci11030061)] [Medline: [37755165](https://pubmed.ncbi.nlm.nih.gov/37755165/)]
87. Ang TL, Choolani M, See KC, Poh KK. The rise of artificial intelligence: addressing the impact of large language models such as ChatGPT on scientific publications. *Singapore Med J.* Apr 2023;64(4):219-221. [doi: [10.4103/singaporemedj.SMJ-2023-055](https://doi.org/10.4103/singaporemedj.SMJ-2023-055)] [Medline: [37006087](https://pubmed.ncbi.nlm.nih.gov/37006087/)]
88. Naik N, Hameed BM, Shetty DK, et al. Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? *Front Surg.* Mar 14, 2022;9:862322. [doi: [10.3389/fsurg.2022.862322](https://doi.org/10.3389/fsurg.2022.862322)] [Medline: [35360424](https://pubmed.ncbi.nlm.nih.gov/35360424/)]

89. Weidener L, Fischer M. Teaching AI ethics in medical education: a scoping review of current literature and practices. *Perspect Med Educ*. Oct 16, 2023;12(1):399-410. [doi: [10.5334/pme.954](https://doi.org/10.5334/pme.954)] [Medline: [37868075](https://pubmed.ncbi.nlm.nih.gov/37868075/)]
90. Danehy T, Hecht J, Kentis S, Schechter CB, Jariwala SP. ChatGPT performs worse on USMLE-style ethics questions compared to medical knowledge questions. *Appl Clin Inform*. Oct 2024;15(5):1049-1055. [doi: [10.1055/a-2405-0138](https://doi.org/10.1055/a-2405-0138)] [Medline: [39209308](https://pubmed.ncbi.nlm.nih.gov/39209308/)]

Abbreviations

AI: artificial intelligence
BC: betweenness centrality
IF: impact factor
LLM: large language model
RQ: research question
USMLE: United States Medical Licensing Examination

Edited by Blake Lesselroth; peer-reviewed by Chandrashekar Br, Dongyan Nan, Gerard Gill, Mohan Krishna Ghanta, Weihua Yang; submitted 08.02.2025; final revised version received 11.08.2025; accepted 09.09.2025; published 07.10.2025

Please cite as:

Zhang Y, Xie X, Xu Q
ChatGPT in Medical Education: Bibliometric and Visual Analysis
JMIR Med Educ 2025;11:e72356
URL: <https://mededu.jmir.org/2025/1/e72356>
doi: [10.2196/72356](https://doi.org/10.2196/72356)

©Yuning Zhang, Xiaolu Xie, Qi Xu. Originally published in JMIR Medical Education (<https://mededu.jmir.org>), 07.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Education, is properly cited. The complete bibliographic information, a link to the original publication on <https://mededu.jmir.org/>, as well as this copyright and license information must be included.