

Using ChatGPT in Medical Research: Current Status and Future Directions

Suebsarn Ruksakulpiwat¹, Ayanesh Kumar², Anuoluwapo Ajibade³

¹Department of Medical Nursing, Faculty of Nursing, Mahidol University, Bangkok, Thailand; ²School of Medicine, Case Western Reserve University, Cleveland, OH, USA; ³College of Art and Science, Department of Anthropology, Case Western Reserve University, Cleveland, OH, USA

Correspondence: Suebsarn Ruksakulpiwat, Department of Medical Nursing, Faculty of Nursing, Mahidol University, Bangkok, Thailand, Email suebsarn25@gmail.com

Objective: This review aims to evaluate the current evidence on the use of the Generative Pre-trained Transformer (ChatGPT) in medical research, including but not limited to treatment, diagnosis, or medication provision.

Methods: This review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. We searched Google Scholar, Web of Science, PubMed, and Medline to identify studies published between 2022 and 2023 that aimed to utilize ChatGPT in medical research. All identified references were stored in EndNote.

Results: We initially identified 114 articles, out of which six studies met the inclusion and exclusion criteria for full-text screening. Among the six studies, two focused on drug development (33.33%), two on literature review writing (33.33%), and one each on medical report improvement, provision of medical information, improving research conduct, data analysis, and personalized medicine (16.67% each).

Conclusion: ChatGPT has the potential to revolutionize medical research in various ways. However, its accuracy, originality, academic integrity, and ethical issues must be thoroughly discussed and improved before its widespread implementation in clinical research and medical practice.

Keywords: ChatGPT, artificial intelligence, medical research, review

Introduction

Artificial intelligence, or AI, has been described as a branch of computer science that focuses on creating intelligent machines that can think and act like humans.¹ AI makes decisions by learning from its environment and the information they obtain.¹ There is various type of AI, including machine learning, an algorithm that learns from data and makes a prediction.^{1,2} Another type of AI is Natural Language Processing (NLP) which uses algorithms to understand and generate human-like conversations.³ Recently, AI has been utilized in various ways, such as medical diagnosis,⁴ the Internet of Things,⁵ and artificial intelligence of things.⁶

Especially in healthcare, AI has been beginning to be applied and shows the potential to transform many aspects of patient care and administrative processes within providers or pharmaceutical organizations.² A study from Davenport & Kalakota described the potential for AI in healthcare. The author suggested that healthcare providers and life sciences companies already use several types of AI. The fundamental application categories are diagnosis and treatment recommendations, patient engagement and adherence, and administrative activities.² A recent systematic review aimed to evaluate the existing evidence of a machine learning-based classification system that stratifies patients with stroke. The result found that machine learning models have the potential to help healthcare providers with stroke diagnoses that can lead to early treatment and improve patient outcomes.⁷

Generative Pre-trained Transformer, or ChatGPT, an NLP system, is a chatbot launched by OpenAI in November 2022. It was designed to generate human-like conversations by understanding the context of a conversation and generating appropriate responses.^{8,9} ChatGPT has several features that make it a powerful NLP system. It can understand the context of a conversation and generate appropriate responses in different styles, such as formal, informal, and humorous.⁹ There is a debate about using ChatGPT in medical research, and a potential concern has been noted. For example, privacy and security. In medical research, AI systems depend on access to large amounts of private data, such as medical records. The data could be

accessed by unauthorized parties and used for nefarious purposes if the data is not adequately protected and secured.¹⁰ In addition, misuse and over-reliance is other significant concern that must be noted. Although AI systems like ChatGPT can be compelling, they could be better. Medical professionals may over-rely on AI systems and trust their decisions without adequately considering the limitations and potential errors of the technology.¹⁰

This review aims to evaluate the current evidence related to the use of ChatGPT in medical research, including its potential to provide treatment, diagnosis, and medication. Although ChatGPT's use in healthcare is still in its early stages, there is a need for sufficient prior research on the concept, research fields, and application cases. Therefore, this review seeks to provide insights into the current trends of utilizing this technology in medical research and offer suggestions for future research.

Methods

Identify Relevant Studies

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines¹¹ were used to guide the identification, screening, exclusion, and inclusion of articles in this review. Four electronic databases (Google Scholar, Web of Science, PubMed, and Medline) were searched on January 21, 2023, at 9:26 PM EST to identify articles published between 2022 and 2023 that were related to or aimed to utilize ChatGPT in medical research. The search terms “ChatGPT” AND “Chatbot” AND “Medical Research” were used. Additionally, the reference lists of the included studies were manually searched to obtain relevant studies, and all references were stored in EndNote. A flow diagram was created to present the results of the search and screening process, following the PRISMA guidelines.

Study Selection

The authors independently screened the titles and abstracts of the identified studies to determine their relevance. Subsequently, the full text of the selected articles was also assessed to ensure they met the inclusion criteria. These criteria were implemented to ensure that only studies relevant to the objective of the review were included. Similarly, exclusion criteria were used to eliminate literature unrelated to the review (Table 1).

Data Extraction

The standardized chart for data extraction (Supplementary Table 1) consisted of the following data for each study: Reference, Year, Country, Study Design, Sample size, Target population, Objective, Results (Utilize ChatGPT in medical research), Main result/key finding and Suggestions for future research.

Results

Search Results

Initially, a total of 114 articles were identified. After running Endnote X8, no duplicates were found. Subsequently, 114 articles were screened based on their title and abstract using the inclusion and exclusion criteria, resulting in six articles

Table 1 Inclusion and Exclusion Criteria

| Inclusion Criteria |
|--|
| <ul style="list-style-type: none"> • All types of articles and preprints, including quantitative studies, qualitative studies, mixed methods, conference proceedings, review articles, theoretical papers, pilot studies, protocols, dissertations, letters to the editor, opinions, viewpoints, perspectives, or editorials. • Articles related to or aiming to utilize ChatGPT in medical research, such as treatment, diagnosis, medication, etc. • Studies conducted in any setting. • Articles published in the English language. • Studies published between 2022 and 2023. |
| Exclusion Criteria |
| <ul style="list-style-type: none"> • Studies that do not address the aim of interest or involve animal subjects. • Articles that contain abstracts only. |

eligible for full-text screening. During the full-text screening phase, no articles were excluded, and all six articles were included in the final analysis. Finally, the retrieval process was outlined using the PRISMA flow chart, as shown in Figure 1.

Description of Included Studies

Table 2 displays that of the six included studies, two were published in 2023 (33.33%), and four were published in 2022 (66.67%). The studies were mainly conducted in the USA and Germany, with two studies from each country (25% each). One study was conducted in each of the following countries: the Republic of Korea, Turkey, Spain, and Switzerland, each accounting for 12.5%. The most popular study design was literature review ($n = 2$, 33.33%), followed by case study ($n = 1$, 16.67%), editorial ($n = 1$, 16.67%), and perspective ($n = 1$, 16.67%). One study had a sample size of 15 (16.67%), while the rest did not specify the target population due to the study design ($n = 5$, 83.33%). Only one study focused on radiologists (16.67%), while the majority did not have a specific target population ($n = 5$, 83.33%).

Description of the Key Finding of Included Studies

Table 3 summarizes the key findings of the included studies. Two studies (33.33%) explored the use of ChatGPT in drug development. In comparison, another two studies (33.33%) examined its application in writing literature reviews. One study (16.67%) described using ChatGPT for medical report improvement, one for providing treatment, one for providing medical information, one for improving research conduction, one for data analysis, and one for personalized medicine.

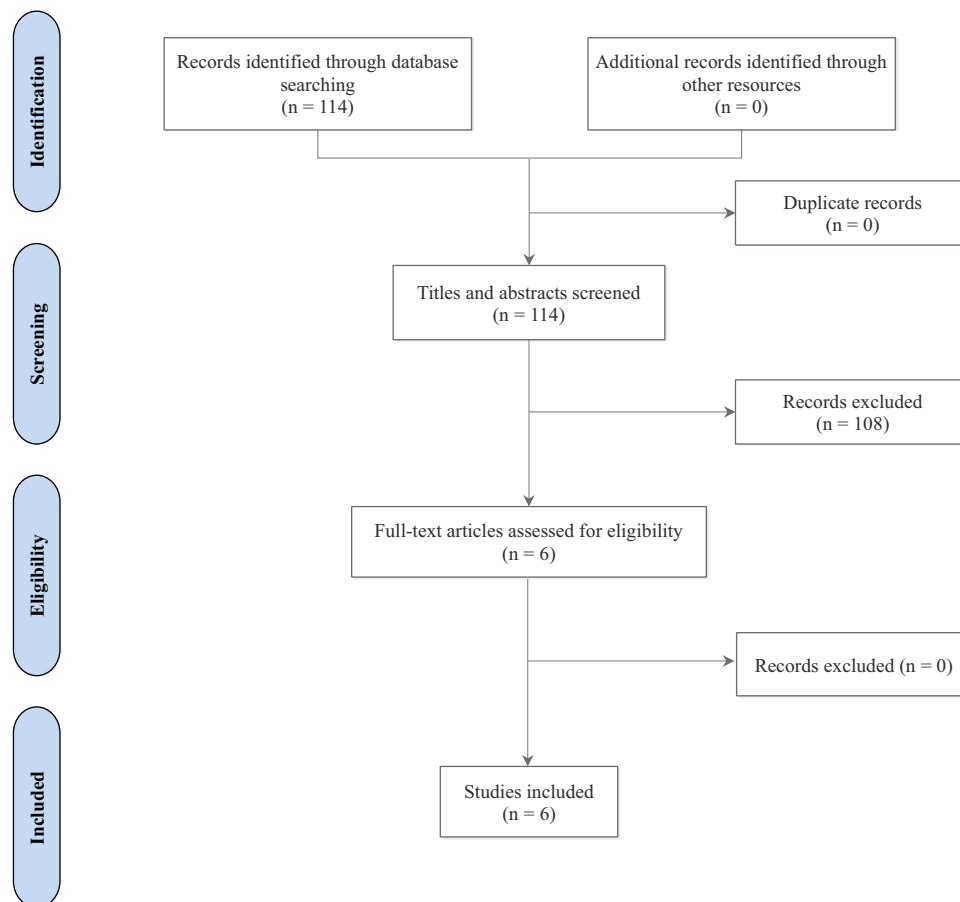


Figure 1 PRISMA flow chart.

Note: Adapted from Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Reprint—preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Phys Ther.* 2009;89(9):873–880.¹¹

Table 2 The Characteristics of the Included Studies

| Characteristic | Number* | Percentage (%) |
|--------------------------|---------|----------------|
| Publication Year | | |
| 2023 | 2 | 33.33% |
| 2022 | 4 | 66.67% |
| Country | | |
| USA | 2 | 25% |
| Germany | 2 | 25% |
| Republic of Korea | 1 | 12.5% |
| Turkey | 1 | 12.5% |
| Spain | 1 | 12.5% |
| Switzerland | 1 | 12.5% |
| Study design | | |
| Literature review | 2 | 33.33% |
| Case study | 1 | 16.67% |
| Editorial | 1 | 16.67% |
| Perspective | 1 | 16.67% |
| Not specific | 1 | 16.67% |
| Sample size | | |
| 15 | 1 | 16.67% |
| N/A | 5 | 83.33% |
| Target Population | | |
| Radiologists | 1 | 16.67% |
| N/A | 5 | 83.33% |

Note: *The number of included studies.

Abbreviation: N/A, Not applicable.

Discussion

This review summarizes the use of ChatGPT in medical research. Currently, this AI is primarily used in drug development, improving medical reports, providing treatment and medical information, writing literature reviews on health-related topics, improving research conduct, data analysis, and personalizing medicine. The discussion section details how ChatGPT was utilized in each area and how it could be implemented in future research.

Drug Development

Evidence shows the potential of using AI technology in new drug development. Mann used ChatGPT to discuss the role of AI in translational medicine. The result illustrated that AI algorithms could help speed up new drug development and treatment and identify the side effect and interactions.¹⁷ In addition, a study by Blanco-Gonzalez et al discussed the benefits, challenges, and drawbacks of AI in drug discovery. It has shown that AI techniques, such as machine learning, can accelerate and improve drug development processes by enabling more efficient and accurate analysis of large amounts of data.¹² This is consistent with previous studies suggesting that AI has been found to improve the efficiency of the drug development process and successfully

Table 3 Key Findings of Included Studies

| Main Result/Key Finding | References | | | | | | The Total Included Studies (n, %) |
|-------------------------------|------------|------|------|------|------|------|-----------------------------------|
| | [12] | [13] | [14] | [15] | [16] | [17] | |
| Drug Development | X | | | | | X | 2 (33.33%) |
| Medical Report Improvement | | X | | | | | 1 (16.67%) |
| Providing Treatment | | | X | | | | 1 (16.67%) |
| Providing Medical Information | | | X | | | | 1 (16.67%) |
| Writing Literature Review | X | | | X | | | 2 (33.33%) |
| Improve Research Conduction | | | | | X | | 1 (16.67%) |
| Data Analysis | | | | | | X | 1 (16.67%) |
| Personalize Medicine | | | | | | X | 1 (16.67%) |

predict the efficacy of drug compounds with high accuracy.^{18,19} Although AI is expected to significantly contribute to developing new medications in the following few years, limitations, including ethical issues, must be considered. Moreover, AI should be cautiously employed in science until it can be entrusted to produce reliable and accurate information. Also, it is crucial to carefully evaluate the information provided by AI tools such as ChatGPT and validate it using reliable sources.

Medical Report Improvement

Although ChatGPT is not explicitly developed for simplifying medical reports, it has surprisingly performed well. One of our included studies investigated the phenomenon that ChatGPT may be used to simplify radiology reports. This case study included 15 radiologists to provide the agreement in radiology reports generated by ChatGPT.¹³ Overall, radiologists agreed that the simplified reports generated by ChatGPT were complete but had some errors, such as missed findings or unspecific locations. Accordingly, it may cause imprecision in the medical context, leading to an error in patient treatment and clinical decision-making.¹³ A previous study discussing using AI for quality improvement in radiology also suggested that AI can provide more support, such as ensuring medical reports are accurate, readable, and helpful to patients and healthcare providers.²⁰ However, future research is still needed to validate the findings and further explore the possibilities of this new technology in the medical domain. In this case, ChatGPT may involve a simplified radiology report autogenerated alongside an original report, proofread by a radiologist, and corrected where needed.

Providing Treatment and Medical Information

The study by Kim determined the medical information and treatment options that ChatGPT can provide for shoulder impingement syndrome (SIS).¹⁴ The result reveals that ChatGPT output provides essential and uncontroversial treatment options for SIS. However, more details and specific treatment options still need to be obtained through a medical professional. Thus, it will be difficult for non-experts to determine treatment options based on minimal depth in the information provided by ChatGPT since effective treatment varies greatly depending on an individual's condition or degree of SIS.¹⁴ Regarding medical information, ChatGPT considers exercise to be one of the critical factors in recovering from SIS and provides several exercise examples. However, the description of the posture and action of the pendulum exercise was inaccurate.¹⁴ As Davenport & Kalakota suggested, AI will have a vital role in future healthcare offerings, particularly in treatment applications. However, accuracy and ethical issues must be addressed, making it challenging.² Similarly, ChatGPT can provide general and basic-level medical information. Nevertheless, future technological advancements in the medical field must pave the way to develop this AI to be more accurate and detailed in medical and treatment information.

Writing a Literature Review on a Health-Related Topic

Two included studies used ChatGPT to conduct the literature review on the medical-related topic. The literature review by Blanco-Gonzalez et al discussed the benefits, challenges, and drawbacks of AI in drug discovery and proposed possible strategies and approaches for overcoming barriers. This review was generated with the assistance of ChatGPT and verified by human authors.¹² Furthermore, Aydin & Karaarslan used ChatGPT to create a literature review of the theme of the application of digital twins in healthcare.¹⁵ The result shows that the texts written by study authors tested a low level of plagiarism compared to ChatGPT. On the other hand, the paraphrased abstracts created by ChatGPT had very high levels of plagiarism. Also, ChatGPT does not produce original texts after paraphrasing.¹⁵

A previous study aimed to determine whether ChatGPT can write a good boolean query for a systematic literature review search. The authors recommended that AI can follow complex instructions and generate exact queries, making it a valuable innovation for researchers conducting systematic reviews.²¹ However, human editing and verification are still necessary to minimize plagiarism and errors. Additionally, utilizing ChatGPT to conduct research, such as a literature review, presents challenges due to originality and academic integrity issues. Therefore, researchers should carefully consider the policies of their research institutions and journals before using AI in their research writing.

Research Conduction, Data Analysis, and Personalized Medicine

Two studies included in the review report the potential of using ChatGPT in conducting research, data analysis, and personalized medicine.^{16,17} Cahan & Treutlein illustrated how ChatGPT could assist practitioners across the broader stem cell research discipline, mainly by saving them time to conduct more research.¹⁶ The author stated that ChatGPT could aid stem cell research in three ways: 1) the ability to process and analyze a large amount of cell-related data, 2) the optimization of stem cell culture conditions to allow for more efficient and controlled growth of stem cells, and 3) the creation of detailed models of stem cell behavior, helping researchers better understand how these cells respond to stimuli and how they can be manipulated for different purposes.¹⁶

Likewise, Mann discussed the role of ChatGPT in translational medicine. The researcher suggested the following future directions for AI in translational medicine: 1) AI can be used in big data analysis, such as electronic medical records and genomic data, to help identify factors causing disease and predict patient outcomes; 2) AI can help develop personalized medicine that is tailored to the specific needs and characteristics of individual patients.¹⁷ However, using ChatGPT in translational medicine presents challenges, such as: 1) algorithms can be biased and incomplete if the training data is biased and incomplete, potentially leading to inaccurate results that could impact patient care and outcomes; 2) AI may not fully understand the biological mechanisms of the human body system, resulting in limitations in providing factors contributing to disease and treatment; and 3) ethical implications of using AI, such as discrimination against certain populations or prioritizing specific groups, could lead to future healthcare inequalities and impact human life.¹⁷

There are limitations to our review that need to be noted. Firstly, our search was limited to articles published in English between 2022 and 2023. Therefore, research published in other languages or outside of this timeframe may have been omitted, which could limit the generalizability and raise issues regarding the validity of the finding. Additionally, we focused specifically on the use of ChatGPT in medical research and did not include other potential applications of this technology in related fields. Finally, while we searched a range of databases, it is possible that some relevant studies were not captured by our search strategy.

Conclusion and Suggestions for Future Research

Over the recent years, ChatGPT has demonstrated its role in advancing the medical field, from supporting translational medicine and drug development with detailed and accurate data analysis to complementing medical practice and patient experience with improved medical reporting, diagnostics, and treatment plans. However, further work is required to enhance accuracy, originality, bias, and misuse and overcome issues pertaining to academic integrity, privacy, and ethics prior to the extensive application of this tool within research and clinical practice.

It is important to note that ChatGPT was not explicitly developed for application in research and medicine. As such, it lacks a much-needed depth in scientific and medical knowledge underlying mechanisms of disease and treatment. Yet, it has performed surprisingly well in providing basic-level support in research and clinical settings. This is a testament to its potential to revolutionize medicine and healthcare if further technological advancement of this model is made in conjunction with the medical field. For instance, future work must focus on training AI to develop an extensive understanding of the biological and medical sciences to improve the accuracy and depth of its analysis, diagnostics, and treatment plan generation. In addition, appropriate ethical guidelines, limitations, and authorizations must be placed on its use to prevent unauthorized use and protect privacy. Moreover, similar studies conducted in the future with improved models applied within research and clinical settings are required to track its progress. Finally, surveys of physicians' and patients' experiences and opinions may also aid in evaluating ChatGPT's role in paving the way for a more seamless provider and patient care experience, respectively.

ChatGPT also has practical potential for enhancing patient care and treatment outcomes by providing medical information and facilitating communication between patients and healthcare providers. Academically, ChatGPT can advance understanding, identify new research questions, and improve data analysis and interpretation accuracy. Nevertheless, there are potential risks, such as the dissemination of inaccurate information and ethical concerns about informed consent, privacy, and data security. Therefore, researchers and healthcare providers must consider these factors when implementing ChatGPT-based interventions.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Deng J, Lin Y. The benefits and challenges of ChatGPT: an overview. *Front Comput Intelligent Syst.* 2022;2(2):81–83.
2. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthcare J.* 2019;6(2):94–98.
3. Deng L, Liu Y. *Deep Learning in Natural Language Processing.* Springer; 2018.
4. Kamdar J, Jeba Praba J, George JJ. Artificial intelligence in medical diagnosis: methods, algorithms and applications. In: *Machine Learning with Health Care Perspective.* Springer; 2020:27–37.
5. Rose K, Eldridge S, Chapin L. The internet of things: an overview. *Internet Soc.* 2015;80:1–50.
6. Yu K, Guo Z, Shen Y, Wang W, Lin JC-W, Sato T. Secure artificial intelligence of things for implicit group recommendations. *IEEE Internet Things J.* 2021;4(4):2698–2707.
7. Ruksakulpiwat S, Thongking W, Zhou W, et al. Machine learning-based patient classification system for adults with stroke: a systematic review. *Chronic Illn.* 2021;19(1):26–39.
8. King MR. The future of ai in medicine: a perspective from a chatbot. *Ann Biomed Eng.* 2022;51(2):291–295.
9. Deng J, Lin Y. The benefits and challenges of ChatGPT: an overview. *Front Comput Intelligent Syst.* 2023;2(2):81–83.
10. King MR. *The Future of AI in Medicine: A Perspective from a Chatbot.* Springer; 2022:1–5.
11. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Reprint—preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Phys Ther.* 2009;89(9):873–880.
12. Blanco-Gonzalez A, Cabezon A, Seco-Gonzalez A, et al. The role of ai in drug discovery: challenges, opportunities, and strategies. *arXiv.* 2022. doi:10.48550/arXiv.2212.08104
13. Jeblick K, Schachtner B, Dexl J, et al. ChatGPT makes medicine easy to swallow: an exploratory case study on simplified radiology reports. *arXiv.* 2022. doi:10.48550/arXiv.2212.14882
14. J-h K. Search for medical information and treatment options for musculoskeletal disorders through an artificial intelligence chatbot: focusing on shoulder impingement syndrome. *medRxiv.* 2022. doi:10.1101/2022.12.16.22283512
15. Aydın Ö, Karaarslan E. OpenAI ChatGPT generated literature review: digital twin in healthcare. Available at SSRN 4308687; 2022.
16. Cahan P, Treutlein B. A conversation with ChatGPT on the role of computational systems biology in stem cell research. *Stem Cell Rep.* 2023;18(1):1–2.
17. Mann DL. Artificial intelligence discusses the role of artificial intelligence in translational medicine: a JACC: basic to translational science interview with ChatGPT. *Basic Transl Sci.* 2023;8(2):221–223.
18. Mak K-K, Pichika MR. Artificial intelligence in drug development: present status and future prospects. *Drug Discov Today.* 2019;24(3):773–780.
19. Liang G, Fan W, Luo H, Zhu X. The emerging roles of artificial intelligence in cancer drug development and precision therapy. *Biomed Pharmacother.* 2020;128:110255.
20. Loehfelm TW. Artificial intelligence for quality improvement in radiology. *Radiol Clin North Am.* 2021;59(6):1053–1062.
21. Wang S, Scells H, Koopman B, Zuccon G. Can chatgpt write a good boolean query for systematic review literature search? *arXiv.* 2023. doi:10.48550/arXiv.2302.03495

Journal of Multidisciplinary Healthcare

Dovepress

Publish your work in this journal

The Journal of Multidisciplinary Healthcare is an international, peer-reviewed open-access journal that aims to represent and publish research in healthcare areas delivered by practitioners of different disciplines. This includes studies and reviews conducted by multidisciplinary teams as well as research which evaluates the results or conduct of such teams or healthcare processes in general. The journal covers a very wide range of areas and welcomes submissions from practitioners at all levels, from all over the world. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-inflammation-research-journal>