

Early applications of ChatGPT in medical practice, education and research

Author: Sam Sedaghat^A

ABSTRACT

ChatGPT, which can automatically generate written responses to queries using internet sources, soon went viral after its release at the end of 2022. The performance of ChatGPT on medical exams shows results near the passing threshold, making it comparable to third-year medical students. It can also write academic abstracts or reviews at an acceptable level. However, it is not clear how ChatGPT deals with harmful content, misinformation or plagiarism; therefore, authors using ChatGPT professionally for academic writing should be cautious. ChatGPT also has the potential to facilitate the interaction between healthcare providers and patients in various ways. However, sophisticated tasks such as understanding the human anatomy are still a limitation of ChatGPT. ChatGPT can simplify radiological reports, but the possibility of incorrect statements and missing medical information remain. Although ChatGPT has the potential to change medical practice, education and research, further improvements of this application are needed for regular use in medicine.

KEYWORDS: ChatGPT, artificial intelligence, academic writing, medical education, healthcare

DOI: 10.7861/clinmed.2023-0078

Introduction

ChatGPT (OpenAI, San Francisco, CA, USA) is an AI chatbot which was introduced in November 2022¹ and soon after went viral. ChatGPT has the ability to respond to various kinds of queries, automatically generating responses using internet sources. People across different fields, generations and continents started using ChatGPT,² leading to a continuous increase in its popularity. Medicine is a field in which simplifying artificial intelligence (AI)-based technologies are highly important. It is obvious that applications such as ChatGPT have the potential to change medicine, with uses ranging from the automated extraction of electronic medical records³ to the development of sophisticated treatment plans. This article presents an overview of the early applications of ChatGPT in medicine.

Author: ^Aconsultant radiologist, University Hospital of Heidelberg, Heidelberg, Germany

ChatGPT in medical education

A recent study evaluated the performance of ChatGPT on the United States Medical Licensing Examination (USMLE). The study revealed that ChatGPT passed all three exams (Step 1, Step 2 CK, Step 3) near the passing threshold without any previous training.⁴ On the other hand, Gilson *et al* state that there is a significant decrease in performance with an increased difficulty of the questions. However, the authors compare the performance of ChatGPT to a third-year medical student.⁵ Antaki *et al* tested ChatGPT for use in two multiple-choice question banks for the Ophthalmic Knowledge Assessment Program (OKAP) exam. The authors found similar results, with ChatGPT achieving 55.8% and 42.7% accuracy in the exams.⁶ Another study from Korea aimed at directly correlating the knowledge of Chat-GPT to that of medical students on the topic of parasitology. The authors revealed that the performance of ChatGPT was lower than medical students and concluded that ChatGPT's ability is not yet at an entirely acceptable level.⁷

Academic writing

Gao *et al* tested ChatGPT's ability to write academic abstracts. They included 50 abstracts from five high-impact medical journals and asked ChatGPT to produce research abstracts using provided titles and journal requirements. The authors concluded that all ChatGPT-derived abstracts were acceptably written, but only 8% of them respected the formatting requirements of the journals. 68% of the generated abstracts by ChatGPT were correctly identified by the reviewers due to the 'vaguer' and more 'formulaic' type of writing. An AI output detector showed similar results in detecting the ChatGPT-derived abstracts.⁸

In a recent article by Guo *et al*,⁹ the main attributes of ChatGPT's writing style are identified. The authors mention that ChatGPT writes in an organised manner and prefers a straightforward concept in the questions. Its answers are long and detailed, it shows less harmful information, and refuses to answer when it has no information about topics. However, it might 'fabricate facts' to give an answer, which should make users cautious regarding the professional use of ChatGPT. The authors observed that the main difference between ChatGPT's writing style and human writing is that the human writer is more subjective, colloquial and emotional, giving human abstracts a personal note.⁹ OpenAI warns that ChatGPT could sometimes 'responds to harmful instructions'.^{10,11}

A recent study evaluated ChatGPT's ability to generate a literature review on the concept of the 'digital twin' in healthcare, asking it to paraphrase selected literature from 2020 to 2022. Although the results were promising, the iThenticate plagiarism detection tool identified many plagiarism matches.¹²

Interaction with patients and radiological reporting

Thurzo *et al* reviewed AI-based applications, including ChatGPT, in the dental field. They concluded that ChatGPT could facilitate the interaction between healthcare providers and patients in various ways, from analysing patient messages to personalising the communication between healthcare professionals and patients. However, they found that ChatGPT has limitations in relation to sophisticated tasks such as understanding the human anatomy.¹³ Nov *et al* tested ChatGPT against healthcare providers' responses to patients. They found that ChatGPT had a similar rate of correct answers compared to the providers.¹⁴ In a case study conducted by Jeblick *et al*,¹⁵ radiologists had the task of evaluating the quality of simplified radiology reports generated with ChatGPT. The results showed that the reports were 'correct, complete, and not potentially harmful to patients'. However, incorrect statements and missing medical information that could potentially have led to harmful conclusions were also detected. Although this case study comprises small sample numbers, the authors emphasise the great potential of ChatGPT in radiology while also mentioning the need for further improvements.¹⁵

Conclusion

ChatGPT seems to fulfil a long-held desire to simplify medical practice, education and research. However, ChatGPT is still a very novel and early-stage application which needs further improvements to be widely usable in medicine. Although ChatGPT is a highly sophisticated application, which could change medical practice, research and education substantially, the last instance should remain human judgment. ■

References

- Liebrezn M, Schleifer R, Buadze A, Bhugra D, Smith A. Generating scholarly content with ChatGPT: ethical challenges for medical publishing. *Lancet Digit Health* 2023;5:e105–6.
- Aydin Ö, Karaarslan E. *Is ChatGPT leading generative AI? What is beyond expectations?* Social Science Research Network, 2023. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4341500.
- Biswas S. ChatGPT and the future of medical writing. *Radiology* 2023;307:e223312.
- 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* 2023;2:e0000198.
- Gilson A, Safranek C, Huang T *et al*. How does ChatGPT perform on the medical licensing exams? The implications of large language models for medical education and knowledge assessment. *MedRxiv* 2022;2022.12.23.22283901.
- 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. *medRxiv* 2023;2023.01.22.23284882.
- 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* 2023;20:1.
- Gao CA, Howard FM, Markov NS *et al*. Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *bioRxiv* 2022;2022.12.23.521610.
- Guo B, Zhang X, Wang Z *et al*. How close is ChatGPT to human experts? Comparison Corpus, Evaluation, and Detection. *arXiv* 2023;2301.07597.
- Kitamura FC. ChatGPT is shaping the future of medical writing but still requires human judgment. *Radiology* 2023;230171.
- OpenAI TB. *Chatgpt: Optimising language models for dialogue*. OpenAI, 2022.
- Aydin Ö, Karaarslan E. *OpenAI ChatGPT generated literature review: digital twin in healthcare*. Social Science Research Network, 2023. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4308687
- Thurzo A, Strunga M, Urban R, Surovková J, Afrashtehfar KI. Impact of artificial intelligence on dental education: a review and guide for curriculum update. *Educ Sci* 2023;13:150.
- Nov O, Singh N, Mann DM. Putting ChatGPT's medical advice to the (Turing) test. *medRxiv* 2023;2023.01.23.23284735.
- Jeblick K, Schachtner B, Dextl J *et al*. ChatGPT makes medicine easy to swallow: an exploratory case study on simplified radiology reports. *arXiv* 2022;2212.14882.

**Address for correspondence: Dr Sam Sedaghat MD, Department of Diagnostic and Interventional Radiology, University Hospital Heidelberg, Im Neuenheimer Feld 420, 69120 Heidelberg, Germany
Email: samsedaghat1@gmail.com**