



# The Challenges of ChatGPT in Healthcare Scientific Writing

SeyedAhmad SeyedAlinaghi<sup>1</sup>, Faeze Abbaspour<sup>2</sup> and Esmaeil Mehraeen<sup>3,\*</sup>

<sup>1</sup>Iranian Research Center for HIV/AIDS, Iranian Institute for Reduction of High Risk Behaviors, Tehran University of Medical Sciences, Tehran, Iran

<sup>2</sup>School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

<sup>3</sup>Department of Health Information Technology, Khalkhal University of Medical Sciences, Khalkhal, Iran

\*Corresponding author: Department of Health Information Technology, Khalkhal University of Medical Sciences, P. O. Box: 5681761351, Khalkhal, Iran. Tel: +98-4532426801, Fax: +98-4532422305, Email: es.mehraeen@gmail.com

Received 2023 October 09; Revised 2023 November 11; Accepted 2023 November 15.

**Keywords:** Healthcare, ChatGPT, Writing

## Dear Editor,

Due to the increase in artificial intelligence (AI) applications in various fields of healthcare (1), ChatGPT, as one of the branches of AI, is used for writing scientific texts. In addition to its capabilities and advantages, ChatGPT also faces certain challenges and limitations. The authors aimed to provide evidence in this letter.

ChatGPT, powered by the generative pre-trained transformer 3.5 architecture, was introduced by OpenAI on November 30, 2022 (2). The advent of OpenAI's large language models began with GPT-1 in 2017, followed by GPT-2 and GPT-3, and finally the well-known ChatGPT (3). ChatGPT's main purpose was to produce human-like text responding to user inputs. Using a vast dataset gathered from various sources, ChatGPT acquired extensive knowledge spanning a wide range of topics up until its last update in September 2021. Despite chatbots and conversational agents being utilized in the medical field, it was only after the introduction of ChatGPT that the potential of chatbots was fully revealed (4).

ChatGPT is a valuable tool for various tasks in the medical field, such as extracting information from medical records, aiding in literature searches by finding academic papers and extracting key findings from them, emphasizing areas of uncertainty, and providing suggestions for structure (5). Moreover, its capabilities are evident as it can achieve acceptable scores on the United States medical licensing exam (USMLE) (6). Additionally, ChatGPT can generate abstracts that, in certain cases, resemble those produced by humans, making it a valuable resource for medical professionals and researchers (7).

ChatGPT, similar to other chatbots, depends on machine learning, particularly deep learning and natural

language processing, to produce texts that closely resemble human writing when generating responses to user inputs. While specific characteristics like a lack of creativity, style, and originality may suggest that a paper was written by a chatbot, it's essential to note that not all AI-generated content is easily distinguishable from human-authored articles (8, 9). As AI-generated content becomes more advanced, it might become progressively challenging for individuals to differentiate between text produced by AI and text authored by humans (10). It is noted that advancements in technology are currently enabling the conversion of AI-generated text into human writing. Therefore, in the near future, it may become increasingly difficult for humans to distinguish between AI-generated and human-written text.

Gao et al. conducted a study where they presented a set of abstracts, 50 original and 50 generated by ChatGPT based on original abstracts, to reviewers and AI output detectors. After evaluation, the AI output detector effectively distinguished the generated abstracts, with high scores indicating AI generation, from the original abstracts with low probabilities of being AI-generated. However, when human reviewers were blinded to the source, they correctly identified 68% of the generated abstracts but mistakenly identified 14% of the original abstracts as AI-generated (9). The results of another related study concluded that detecting AI-based articles is challenging, and it often requires a combination of techniques, including language analysis, plagiarism detection, and using AI-powered tools (11).

ChatGPT's reliability and independent capability in scientific writing (SW) are questionable. One of the most concerning issues is that it tends to generate references

that may not be legitimate. In a study conducted by Eysenbach, an interview was conducted with ChatGPT, where a series of questions were asked in the field of medicine, medical practice, and medical research, and the primary concern was that ChatGPT tends to invent references (12). In another study, 178 references listed on ChatGPT were analyzed. The analysis revealed that 69 of these references did not have a digital object identifier (DOI), and 28 references were inaccessible via Google search and lacked a DOI (13).

One limitation of ChatGPT is its tendency to generate responses that appear credible but are incorrect, which is called "artificial hallucination". Artificial hallucination is the term used to describe the phenomenon where a machine, like a chatbot, generates sensory experiences that appear realistic despite lacking any connection to real-world input. Language models like ChatGPT can generate impressive and appropriate responses, but they may sometimes generate content that is entirely made up and inaccurate (14).

Another challenge is plagiarism. ChatGPT's training involves a vast amount of text data, and it generates responses based on statistical probabilities and examples it has seen during its training. If a user inputs a question that is similar to content available on the internet, ChatGPT may produce a response that closely resembles existing content. ChatGPT rephrases or paraphrases information, and it may reproduce content from different sources and unintentionally result in plagiarism. As a result, there is a possibility that ChatGPT might use similar phrases to the text it has learned (15). In a study conducted by Khalil and Er, 50 topics were given to ChatGPT to generate 500-word articles. Afterward, 2 plagiarism detection software were used to measure the percentage of plagiarism. The results showed that out of 50 articles, 40 of them had originality, with the generated text by ChatGPT having 20% or less similarity (16).

ChatGPT can be a valuable tool in SW. It can assist in various aspects, such as selecting appropriate study topics, providing key terms, introducing databases, and summarizing articles. These tasks can be time-consuming, but with the use of AI, high-quality articles can be produced in a shorter time frame. In addition, ChatGPT can be beneficial for non-native English speakers in terms of finding appropriate vocabulary and assisting with grammar and sentence structure. With its high translation capabilities, it can also translate written text into English (15). While ChatGPT can assist in improving the content and structure of articles, it cannot be replaced by our deep understanding of SW. Text generated by ChatGPT should be carefully reviewed and analyzed to ensure its accuracy and reliability (7). Considering that

ChatGPT may provide a mixture of accurate and fabricated information, revising the policies and practices for evaluating scientific manuscripts submitted to journals and medical conferences to improve scientific standards can be helpful (14).

Regarding the issue of plagiarism resulting from the use of ChatGPT, several solutions can be implemented. We can complement the text generated by ChatGPT with information from other sources. Properly citing the mentioned content in the article using accurate citations is also essential. Using plagiarism detection tools can be helpful, but ultimately, nothing can replace the importance of thorough review and the author (15, 17).

## Footnotes

**Authors' Contribution:** All authors contributed equally to this article.

**Conflict of Interests:** The authors have no conflict of interest.

**Funding/Support:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## References

1. Mohammadi S, SeyedAlinaghi SA, Heydari M, Pashaei Z, Pashaei Z, Mirzapour P, et al. Artificial Intelligence in COVID-19 Management: A Systematic Review. *J Comput Sci.* 2023;19(5):554-68. <https://doi.org/10.3844/jcssp.2023.554.568>.
2. Cheng HW. Challenges and Limitations of ChatGPT and Artificial Intelligence for Scientific Research: A Perspective from Organic Materials. *AI.* 2023;4(2):401-5. <https://doi.org/10.3390/ai4020021>.
3. Shen Y, Heacock L, Elias J, Hentel KD, Reig B, Shih G, et al. ChatGPT and Other Large Language Models Are Double-edged Swords. *Radiology.* 2023;307(2):e230163. [PubMed ID: 36700838]. <https://doi.org/10.1148/radiol.230163>.
4. Biswas S. ChatGPT and the Future of Medical Writing. *Radiology.* 2023;307(2):e223312. [PubMed ID: 36728748]. <https://doi.org/10.1148/radiol.223312>.
5. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepano C, et al. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digit Health.* 2023;2(2):e0000198. [PubMed ID: 36812645]. [PubMed Central ID: PMC9931230]. <https://doi.org/10.1371/journal.pdig.0000198>.
6. Flores-Cohaila JA, Garcia-Vicente A, Vizcarra-Jimenez SF, De la Cruz-Galan JP, Gutierrez-Arratia JD, Quiroga Torres BG, et al. Performance of ChatGPT on the Peruvian National Licensing Medical Examination: Cross-Sectional Study. *JMIR Med Educ.* 2023;9:e48039. [PubMed ID: 37768724]. [PubMed Central ID: PMC10570896]. <https://doi.org/10.2196/48039>.
7. Salvagno M, Taccone FS, Gerli AG. Can artificial intelligence help for scientific writing? *Crit Care.* 2023;27(1):75. <https://doi.org/10.1186/s13054-023-04380-2>.
8. van Dis EAM, Bollen J, Zuidema W, van Rooij R, Bockting CL. ChatGPT: five priorities for research. *Nature.* 2023;614(7947):224-6. [PubMed ID: 36737653]. <https://doi.org/10.1038/d41586-023-00288-7>.

9. Gao CA, Howard FM, Markov NS, Dyer EC, Ramesh S, Luo Y, et al. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digit Med.* 2023;**6**(1):75. [PubMed ID: 37100871]. [PubMed Central ID: PMC10133283]. <https://doi.org/10.1038/s41746-023-00819-6>.
10. Dergaa I, Chamari K, Zmijewski P, Ben Saad H. From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biol Sport.* 2023;**40**(2):615–22. [PubMed ID: 37077800]. [PubMed Central ID: PMC10108763]. <https://doi.org/10.5114/biolsport.2023.125623>.
11. Elkhataf AM, Elsaid K, Almeer S. Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *Int J Educ Integr.* 2023;**19**(1):17. <https://doi.org/10.1007/s40979-023-00140-5>.
12. Eysenbach G. The Role of ChatGPT, Generative Language Models, and Artificial Intelligence in Medical Education: A Conversation With ChatGPT and a Call for Papers. *JMIR Med Educ.* 2023;**9**:e46885. [PubMed ID: 36863937]. [PubMed Central ID: PMC10028514]. <https://doi.org/10.2196/46885>.
13. Athaluri SA, Manthena SV, Kesapragada V, Yarlagadda V, Dave T, Duddumpudi RTS. Exploring the Boundaries of Reality: Investigating the Phenomenon of Artificial Intelligence Hallucination in Scientific Writing Through ChatGPT References. *Cureus.* 2023;**15**(4):e37432. [PubMed ID: 37182055]. [PubMed Central ID: PMC10173677]. <https://doi.org/10.7759/cureus.37432>.
14. Alkaissi H, McFarlane SI. Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. *Cureus.* 2023;**15**(2):e35179. <https://doi.org/10.7759/cureus.35179>.
15. Huang J, Tan M. The role of ChatGPT in scientific communication: writing better scientific review articles. *Am J Cancer Res.* 2023;**13**(4):1148–54. [PubMed ID: 37168339]. [PubMed Central ID: PMC10164801].
16. Khalil M, Er E. Will ChatGPT get you caught? Rethinking of plagiarism detection. Preprint. *EdArXiv*. Posted online February 8, 2023. <https://doi.org/10.35542/osf.io/fnh48>.
17. Meyer JG, Urbanowicz RJ, Martin PCN, O'Connor K, Li R, Peng PC, et al. ChatGPT and large language models in academia: opportunities and challenges. *BioData Min.* 2023;**16**(1):20. [PubMed ID: 37443040]. [PubMed Central ID: PMC10339472]. <https://doi.org/10.1186/s13040-023-00339-9>.