

Letter to the Editor

# Reassessing AI in Medicine: Exploring the Capabilities of AI in Academic Abstract Synthesis

Zijian Wang<sup>1</sup>, MS; Chunyang Zhou<sup>2</sup>, MS

<sup>1</sup>Shandong University of Traditional Chinese Medicine, College of Traditional Chinese Medicine, Jinan, China

<sup>2</sup>Department of Radiation Oncology, Qilu Hospital (Qingdao), Cheeloo College of Medicine, Shandong University, Qingdao, China

**Corresponding Author:**

Chunyang Zhou, MS

Department of Radiation Oncology, Qilu Hospital (Qingdao)

Cheeloo College of Medicine

Shandong University

758 Hefei Road, Qingdao

Qingdao, 266000

China

Phone: 86 18561813085

Email: [chunyangzhou29@163.com](mailto:chunyangzhou29@163.com)

**Related Articles:**

Comment on: <http://www.jmir.org/2023/1/e51229/>

Comment in: <http://mhealth.jmir.org/2024/1/e65123/>

## Abstract

(*J Med Internet Res* 2024;26:e55920) doi: [10.2196/55920](https://doi.org/10.2196/55920)

**KEYWORDS**

ChatGPT; AI-generated scientific content; plagiarism; AI; artificial intelligence; NLP; natural language processing; LLM; language model; text; textual; generation; generative; extract; extraction; scientific research; academic research; publication; abstract; comparative analysis; reviewer bias

In their study, Cheng and colleagues [1] delved into a comparison between academic abstracts generated by ChatGPT and those written by humans. Their aim was to evaluate the applicability of an artificial intelligence (AI) model, specifically ChatGPT 3.5, in generating abstracts for basic preclinical research. This investigation, involving 30 articles from renowned journals, revealed that the quality of ChatGPT-generated abstracts generally falls short compared to human-written counterparts, particularly in unstructured formats, occasionally leading to erroneous conclusions. The study also concluded that such AI-generated abstracts can be identified by humans. However, several aspects of this study invite further discussion.

First, the selection process for the experts involved in the study was not explicitly detailed, potentially leading to subjective biases. These biases could arise from the experts' personal preferences, academic or cultural backgrounds, or their familiarity with the subject matter. Such biases might distort the results, causing the quality of AI-generated abstracts to be overestimated or underestimated.

Second, the study used ChatGPT 3.5 as the primary research model. However, there are concerns that ChatGPT 3.5 might not be ideally suited for processing lengthy texts, potentially leading to misunderstandings or the provision of inaccurate information [2]. Contrastingly, its accuracy might surpass that of humans when dealing with short and concise texts [3,4]. We propose that employing ChatGPT in a "layered progressive" manner for text generation could address this issue. This method involves dividing an article into smaller sections, having ChatGPT summarize each section individually, and then compiling these summaries into a cohesive whole. Such an approach is likely to yield better results than generating a summary from the entire text.

Lastly, the study concluded that blinded reviewers could accurately identify ChatGPT-generated abstracts with an accuracy of 93%. Does this conclusion actually hold practical significance? Does the realization that an abstract is AI generated influence the reviewers' judgment of its content? In scientific research, the primary goal of writing a paper is to present research results to the readers. If AI assistance in refining

an article does not alter the accuracy of the research and aids in clearer communication of research content, such practice should be deemed acceptable. A paper generated by AI and reviewed by humans, conveying high-quality research findings, deserves acceptance without bias toward its AI-generated origin.

Therefore, what needs to be avoided is the complete reliance on AI language models for labor-intensive tasks, as

demonstrated in the study, where an entire article is entrusted to AI. This could lead to a lack of critical thinking and the production of low-quality or erroneous articles by human researchers. We must recognize the importance of AI language models as assistants in our writing endeavors. They can inspire our ideas and help refine our text (as researchers often lack the literary flair of writers). If the authors of the article consider the above suggestions, the conclusions of their study might change.

## Conflicts of Interest

None declared.

## References

1. Cheng S, Tsai S, Bai Y, Ko C, Hsu C, Yang F, et al. Comparisons of quality, correctness, and similarity between ChatGPT-generated and human-written abstracts for basic research: cross-sectional study. *J Med Internet Res*. Dec 25, 2023;25:e51229. [FREE Full text] [doi: [10.2196/51229](https://doi.org/10.2196/51229)] [Medline: [38145486](https://pubmed.ncbi.nlm.nih.gov/38145486/)]
2. Meyer JG, Urbanowicz RJ, Martin PCN, O'Connor K, Li R, Peng P, et al. ChatGPT and large language models in academia: opportunities and challenges. *BioData Min*. Jul 13, 2023;16(1):20. [FREE Full text] [doi: [10.1186/s13040-023-00339-9](https://doi.org/10.1186/s13040-023-00339-9)] [Medline: [37443040](https://pubmed.ncbi.nlm.nih.gov/37443040/)]
3. Herbold S, Hautli-Janisz A, Heuer U, Kikteva Z, Trautsch A. A large-scale comparison of human-written versus ChatGPT-generated essays. *Sci Rep*. Oct 30, 2023;13(1):18617. [FREE Full text] [doi: [10.1038/s41598-023-45644-9](https://doi.org/10.1038/s41598-023-45644-9)] [Medline: [37903836](https://pubmed.ncbi.nlm.nih.gov/37903836/)]
4. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med*. Jun 01, 2023;183(6):589-596. [FREE Full text] [doi: [10.1001/jamainternmed.2023.1838](https://doi.org/10.1001/jamainternmed.2023.1838)] [Medline: [37115527](https://pubmed.ncbi.nlm.nih.gov/37115527/)]

## Abbreviations

**AI:** artificial intelligence

*Edited by T Leung, L Beri; this is a non-peer-reviewed article. Submitted 29.12.23; accepted 27.09.24; published 23.12.24.*

*Please cite as:*

Wang Z, Zhou C

*Reassessing AI in Medicine: Exploring the Capabilities of AI in Academic Abstract Synthesis*

*J Med Internet Res* 2024;26:e55920

URL: <https://www.jmir.org/2024/1/e55920>

doi: [10.2196/55920](https://doi.org/10.2196/55920)

PMID:

©Zijian Wang, Chunyang Zhou. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 23.12.2024. 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 the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.