LETTER TO THE EDITOR

Peeking inside GPT-4 for medical research and practice

Partha Pratim Ray

Department of Computer Applications, Sikkim University, Gangtok, Sikkim, India

DEAR EDITOR,

I am writing to express my critique of the article titled "ChatGPT surges ahead: GPT-4 has arrived in the arena of medical research" by Ying-Mei Wang and Tzeng-Ji Chen, published in the *Journal of the Chinese Medical Association*.[1] While the article provides an overview of the performance and potential applications of ChatGPT and GPT-4 in the medical field, I believe there are certain aspects that warrant further consideration and clarification.[2]

First, the article emphasizes the remarkable performance of ChatGPT in various medical examinations, citing studies that indicate its ability to approach or surpass passing thresholds without prior training. While these findings are intriguing, it is important to acknowledge the limitations of these studies, such as the specific examination domains and question formats utilized. Medical examinations, particularly those designed for professionals and specialists, often require complex reasoning, critical thinking, and context-specific knowledge.[3,4] It is crucial to evaluate the extent to which ChatGPT can handle such nuanced tasks beyond simple question-answering, as highlighted in the study on Taiwan's family medicine board examination.

Furthermore, the article's discussion of GPT-4's capabilities and potential impact on medical research and healthcare is rather limited. The article briefly mentions that GPT-4 is more creative, powerful, and reliable but fails to delve into the specific advancements and improvements that GPT-4 offers over ChatGPT. It would have been beneficial to explore the enhanced features of GPT-4, such as improved contextual understanding, advanced reasoning abilities, or domain-specific fine-tuning. Without a comprehensive analysis of GPT-4, the article leaves readers with incomplete information regarding the latest developments in AI language models for medical research.[5,6]

In addition to the limited discussion on GPT-4, the article briefly mentions the surge in publications related to ChatGPT and GPT-4 in medical journals. However, it does not delve into the quality and reliability of these publications. Given the relatively recent introduction of these models, it is important to scrutinize the methodology, peer review process, and ethical considerations of the studies conducted using these AI systems. A more thorough examination of the robustness and generalizability of the findings presented in these publications would

have enhanced the article's credibility and provided readers with a better understanding of the current landscape of AI-driven research in the medical field.[7]

Moreover, the article implies that medical publishing is lagging behind the current development of AI, without providing sufficient evidence or analysis to support this claim. While AI technologies, including language models, have the potential to revolutionize healthcare, it is essential to maintain a cautious and critical approach when integrating them into medical research and practice.[8,9] The assertion that medical publishing is falling behind should be supported by concrete examples and a thorough examination of the current state of AI adoption in the medical field. Additionally, discussing potential challenges and ethical considerations associated with the use of AI language models in medical research would provide readers with a more comprehensive perspective.

I believe the article "ChatGPT surges ahead: GPT-4 has arrived in the arena of medical research" would benefit from a more balanced and in-depth discussion that addresses the limitations of ChatGPT, provides a comprehensive evaluation of GPT-4, examines the quality and implications of the increasing number of publications in this area, and substantiates claims regarding the state of medical publishing in relation to AI development.[10] By offering a more nuanced analysis, the article would contribute to a better understanding of the potential benefits and challenges associated with AI language models in medical research and practice.

REFERENCES

1. Wang YM, Chen TJ. ChatGPT surges ahead: GPT-4 has arrived in the arena of medical research. *J Chin Med Assoc* 2023;**86**:784.
2. Cascella M, Montomoli J, Bellini V, Bignami E. Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. *J Med Syst* 2023;**47**:33.
3. Vaishya R, Misra A, Vaish A. ChatGPT: is this version good for healthcare and research? *Diabetes Metab Syndr* 2023;**17**:102744.
4. Ruksakulpiwa S, Kumar A, Ajibade A. Using ChatGPT in medical research: current status and future directions. *J Multidiscip Healthc* 2023;**16**:1513–20.
5. Khan RA, Jawaid M, Khan AR, Sajjad M. ChatGPT-Reshaping medical education and clinical management. *Pak J Med Sci* 2023;**39**:605–7.
6. Arif TB, Munaf U, Ul-Haque I. The future of medical education and research: is ChatGPT a blessing or blight in disguise? *Med Educ Online* 2023;**28**:2181052.
7. Dahmen J, Kayaalp ME, Ollivier M, Pareek A, Hirschmann MT, Karlsson J, et al. Artificial intelligence bot ChatGPT in medical research: the potential game changer as a double-edged sword. *Knee Surg Sports Traumatol Arthrosc* 2023;**31**:1187–9.
8. Wang DQ, Feng LY, Ye JG, Zou JG, Zheng YF. Accelerating the integration of ChatGPT and other large-scale AI models into biomedical research and healthcare. *MedComm–Future Med* 2023;**2**:e43.
9. Xue VW, Lei P, Cho WC. The potential impact of ChatGPT in clinical and translational medicine. *Clin Transl Med* 2023;**13**:e1216.
10. Liebrenz 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.