ChatGPT as a Novel Consultant in Infectious Diseases and Clinical Microbiology

Hinpetch Daungsupawong[1] (iD), Viroj Wiwanitkit[2] (iD)

[1]Private Academic Consultant, Phonhong, Lao People's Democratic Republic
[2]Saveetha Medical College, Saveetha Institute of Medical and Technical Sciences, Chennai, India

To the Editor,

Tunçer et al. brought us interesting observations from the study "How Reliable is ChatGPT as a Novel Consultant in Infectious Diseases and Clinical Microbiology?" (1). In conclusion, 200 questions about infectious diseases from different platforms were used in the study, along with recommendations from reliable sources. The replies were predicated on predetermined standards, and the questions were meticulously selected and arranged for coherence and clarity. The study's scoring mechanism made it possible to evaluate ChatGPT's precision in giving accurate responses to every query. Overall, ChatGPT performed well, according to the results, with inquiries about urinary tract infections showing the lowest accuracy.

However, the study's findings also revealed potential areas for improvement. For instance, there was a notable discrepancy in accuracy rates between questions sourced from guidelines and those from social media sites. When it came to appropriately answering queries on social media platforms, ChatGPT performed better than guidelines in terms of response rate. This disparity could indicate a limitation in the model's capacity to comprehend and deliver precise data in line with accepted standards. Subsequent investigations could concentrate on resolving this disparity and enhancing ChatGPT's efficacy with regard to questions that follow guidelines.

Furthermore, method faults in the study may include biases in the selection of questions from social media platforms and standards. The study may have needed a broad and representative sample of questions from both sources, which could have biased the findings. Future research could benefit from a more methodical approach to question selection to ensure a balanced representation of topics and sources.

Looking ahead, future research could focus on fine-tuning the study's grading system to account for variations in model responses. This could contribute to a more nuanced assessment of ChatGPT's performance and potentially enhance its ability to provide accurate responses across a wider range of inquiries. Moreover, exploring strategies to improve ChatGPT's accuracy in responding to guideline-based questions could be a valuable area of investigation, with the aim of optimizing its utility in providing accurate and reliable information on infectious diseases.

REFERENCES

1 Tunçer G, Güçlü KG. How reliable is ChatGPT as a novel consultant in infectious diseases and clinical microbiology? Infect Dis Clin Microbiol. 2024;6(1):55-9. [CrossRef]

The Authors Reply

We appreciate Daungsupawong and Wiwanitkitet's interest in our study. They only point out the significant difference in ChatGPT's correct response rates between guidelines and social media questions. They argue that this could be caused by question selections. We agree that bias is possible when choosing the questions, especially from social media. However, we want to emphasize that we meticulously selected the questions derived from social media, taking care to reduce the risk of selection bias. We, as specialists, evaluated these questions for grammatical and linguistic clarity. Moreover, when we evaluate the methodology of the studies on the performance of artificial intelligence tools, it is seen that similar methods have been used for selecting questions (1-3).

Because ChatGPT mostly uses publicly available official websites such as the World Health Organization (WHO) or the Centers for Disease Control and Prevention (CDC), which focus on public health informatics, it has accurately answered the questions derived from social media. As a result, since the guidelines are directed to experts, the questions derived from the guidelines are more complex. Asking professional questions results in poor performance. Therefore, we found a natural and expected result that ChatGPT is more capable of answering social media questions than guideline questions.

Gülşah Tunçer[1] [ID], Kadir Görkem Güçlü[2] [ID]

[1] Bilecik Training and Research Hospital, Department of Infectious Diseases and Clinical Microbiology, Bilecik, Türkiye
[2] Haseki Training and Research Hospital, Department of Infectious Diseases and Clinical Microbiology, İstanbul, Türkiye

REFERENCES

1 Cinar C. Analyzing the performance of ChatGPT about osteoporosis. Cureus. 2023;15(9):e45890. [CrossRef]

2 Caglar U, Yildiz O, Meric A, Ayranci A, Yusuf R, Sarilar O, et al. Evaluating the performance of ChatGPT in answering questions related to benign prostate hyperplasia and prostate cancer. Minerva Urol Nephrol. 2023;75(6):729-33. [CrossRef]

3 Dyckhoff-Shen S, Koedel U, Brouwer MC, Bodilsen J, Klein M. ChatGPT fails challenging the recent ESCMID brain abscess guideline. J Neurol. 2024;271(4):2086-101. [CrossRef]