

Comparative Impact of ChatGPT and Conventional Search Tools on Clinical Reasoning Performance: A Randomized Crossover Study in Preclinical Medical Students [Letter]

Riya Kalra

Maharishi Markandeshwar Institute of Physiotherapy and Rehabilitation, MM(DU), Mullana-Ambala, India

Correspondence: Riya Kalra, Maharishi Markandeshwar Institute of Physiotherapy and Rehabilitation, MM(DU), Mullana-Ambala, India,
Email riyakalra14oct@gmail.com

Dear editor

We read with great interest the article by Nartthanarung et al titled “Comparative Impact of ChatGPT and Conventional Search Tools on Clinical Reasoning Performance: A Randomized Crossover Study in Preclinical Medical Students”.¹ The study provides timely insights into the integration of Artificial Intelligence (AI) tools within medical education and highlights the complementary role of large language models and traditional search strategies in enhancing clinical reasoning. While the findings are promising, several methodological aspects warrant further clarification and critical appraisal.

First, although the randomized crossover design strengthens internal validity by allowing participants to serve as their own controls, the potential for carryover and period effects remains a concern. The authors implemented a 60-minute washout period; however, given the cognitive nature of clinical reasoning and learning retention, this duration may be insufficient to eliminate residual learning effects between interventions. The progressive improvement in scores across phases raises the possibility that learning or practice effects, rather than the intervention itself, contributed substantially to the observed outcomes.²

Second, the sampling framework is limited to a single institution with a relatively small cohort ($n = 46$), which restricts the generalizability of the findings. Although the authors justified the sample size statistically, the inclusion of a homogeneous group of second-year medical students may not adequately represent variability in baseline clinical reasoning abilities across different educational settings.³

Third, the assessment of clinical reasoning using an eight-point rubric, although structured, raises questions regarding its psychometric robustness. While content validity was reviewed by faculty members, the absence of reported inter-rater reliability, construct validity, and sensitivity to change limits confidence in the measurement of such a complex construct. Clinical reasoning is inherently multidimensional, and reliance on a single rubric-based tool may oversimplify its evaluation.⁴

Furthermore, the study utilized paired t-tests for statistical analysis, which are appropriate for within-subject comparisons; however, the absence of adjustments for multiple comparisons may increase the risk of Type I error. Additionally, while effect sizes were reported, the interpretation of cumulative improvement across phases should be approached cautiously due to the potential confounding influence of repeated testing.

Another important consideration is the lack of standardization in ChatGPT prompting strategies. Participants were allowed self-directed prompting, which introduces variability and limits reproducibility. Given that prompt engineering significantly influences AI-generated outputs, this factor represents a critical source of bias that should be controlled or at least systematically documented.

Moreover, while the study acknowledges the risk of AI-generated hallucinations, there is limited exploration of how accuracy and reliability of responses were ensured during the intervention. Without verification mechanisms, it remains unclear whether improved performance reflects enhanced reasoning or reliance on potentially inaccurate synthesized information.

Lastly, the study focuses on short-term outcomes, with no assessment of long-term knowledge retention, transfer to clinical settings, or impact on higher-order reasoning skills. Future research should incorporate longitudinal designs and include objective performance measures in real or simulated clinical environments to better understand the educational implications of AI integration.

In conclusion, while this study contributes valuable preliminary evidence supporting the integration of AI tools in medical education, addressing the aforementioned methodological concerns would strengthen the validity and applicability of the findings. Rigorous, multi-center studies with standardized protocols and robust assessment frameworks are essential to establish the true educational value of AI-assisted learning.

Funding

This communication received no funding.

Disclosure

The author declares no conflicts of interest in this communication.

References

1. Nartthanarung A, Plangsiri K, Kongmalai P. Comparative Impact of ChatGPT and Conventional Search Tools on Clinical Reasoning Performance: a Randomized Crossover Study in Preclinical Medical Students. *Adv Med Educ Practice*. 2026;17:603679. doi:10.2147/AMEP.S603679
2. Kitago T, Ryan SL, Mazzoni P, Krakauer JW, Haith AM. Unlearning versus savings in visuomotor adaptation: comparing effects of washout, passage of time, and removal of errors on motor memory. *Front Human Neurosci*. 2013;7:43462. doi:10.3389/fnhum.2013.00307
3. Bellomo R, Warrillow SJ, Reade MC. Why we should be wary of single-center trials. *Crit Care Med*. 2009;37(12):3114–3119. doi:10.1097/CCM.0b013e3181bc7bd5
4. Fitzner K. Reliability and validity a quick review. *Diabetes Educator*. 2007;33(5):775–780. doi:10.1177/0145721707308172

Dove Medical Press encourages responsible, free and frank academic debate. The content of the Advances in Medical Education and Practice 'letters to the editor' section does not necessarily represent the views of Dove Medical Press, its officers, agents, employees, related entities or the Advances in Medical Education and Practice editors. While all reasonable steps have been taken to confirm the content of each letter, Dove Medical Press accepts no liability in respect of any letter, nor is it responsible for the content and accuracy of any letter to the editor.

Advances in Medical Education and Practice

Publish your work in this journal

Advances in Medical Education and Practice is an international, peer-reviewed, open access journal that aims to present and publish research on Medical Education covering medical, dental, nursing and allied health care professional education. The journal covers undergraduate education, postgraduate training and continuing medical education including emerging trends and innovative models linking education, research, and health care services. The manuscript management system is completely online and includes a very quick and fair peer-review system. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <http://www.dovepress.com/advances-in-medical-education-and-practice-journal>

Dovepress
Taylor & Francis Group