Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Effectiveness and Students' Perception of Tutor-Guided AI Navigation in Undergraduate Medical Education in Sri Lanka: A Quasi-Experimental Study

Randombage PJS

prabathrandombagera@gmail.com

University of Kelaniya

Ahamed MYS

University of Kelaniya

Kulathunga TD

University of Kelaniya

H Wickramasekara

University of Kelaniya

Research Article

Keywords: "Artificial Intelligence", "Generative Artificial Intelligence", "Education, Medical", "Students, Medical"

Posted Date: April 3rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9116697/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Medical Education on May 18th, 2026. See the published version at https://doi.org/10.1186/s12909-026-09467-2.

Abstract

Background:

The integration of Artificial Intelligence (AI), particularly Large Language Models (LLMs) such as ChatGPT, into undergraduate medical education offers new opportunities for personalized learning. However, concerns remain regarding content accuracy, possible over-reliance and limited critical engagement when AI tools are used without guidance. Tutor-guided AI navigation has been proposed as a structured approach combining AI accessibility with educator oversight. This study aimed to evaluate the effectiveness and perception of tutor-guided AI navigation among medical undergraduates.

Methods:

The study was conducted among 87 final-year medical students at the Faculty of Medicine, University of Kelaniya, Sri Lanka. A quasi-experimental pre-test/post-test design compared knowledge gains following tutor-guided versus self-directed ChatGPT-assisted learning. A cross-sectional survey assessing three domains-Perception of Learning Experience, AI Usability and Confidence, Satisfaction and Overall Impression- was administered immediately after the session. Responses were measured using a five-point Likert scale. Internal consistency was evaluated with Cronbach's alpha and descriptive statistics were analysed using SPSS.

Results:

Baseline pre-test scores were comparable between groups ($W = 670.5$, $p = 0.153$). Both groups improved significantly from pre- to post-test ($p < 0.001$), with median scores increasing from 4 to 7. Knowledge gain between groups was not statistically significant (Welch's $t = 1.07$, $p = 0.144$), although the effect size (Hedges' $g = 0.36$) suggested a small-to-moderate advantage for tutor guidance. In the perception survey, reliability was high (Cronbach's alpha > 0.86). Students rated ease of use (4.19 ± 1.07), tutor-guided AI experience (4.03 ± 1.13), satisfaction (3.91 ± 0.94) and willingness to recommend (4.01 ± 0.97) favourably, while trust in AI accuracy was moderate (3.27 ± 0.91).

Conclusion:

Tutor-guided AI-assisted learning was well accepted by students and provided a positive, structured learning experience. Although it did not produce a statistically significant advantage in immediate knowledge gain over self-directed AI use, it was associated with favourable perceptions of usability, satisfaction and recommendation. These findings support the structured integration of tutor-guided AI into undergraduate medical education while maintaining educator oversight and critical appraisal of AI-generated content.

Background

Artificial Intelligence (AI) technologies are reshaping the educational landscape, particularly in the field of medical education. Large Language Models (LLMs) like ChatGPT have been popularized for their potential to make learning experiences more personalised, interactive and available on demand. These AI systems can integrate vast amounts of information, generate comprehensive explanations and simulate clinical reasoning scenarios, making them smart learning partners (1, 2, 12).

Recent research indicates that AI technologies can assist medical students in gaining and applying of information in multiple areas, such as clinical reasoning, anatomy and diagnostic interpretation (3, 4). LLMs like ChatGPT are particularly interesting due to their conversational interface, which allows for repeated questioning and explanation, mirroring the Socratic way of learning (5). Also, their 24/7 availability can make it easier for self-paced and distant learners to use them (6).

But even though they seem promising, there are still concerns regarding how safe and reliable unguided AI use for learning. LLMs are likely to make information that is wrong or misleading and they often do it with a high degree of confidence. This is called "AI hallucination" (7). This risk is especially concerning in medicine, where accuracy and clinical judgement are of the utmost importance. Also, if students don't get the right help, they can rely too much on AI-generated outputs, which could lead to shallow learning, less critical thinking and less growth in their clinical reasoning skills (8).

To overcome these issues, the idea of AI navigation with the help of a tutor has been put up. This teaching method combines the benefits of AI with the guidance and support that teachers give. Tutors can help students create good prompts, understand AI responses, check the accuracy of content and have thoughtful discussions. Structured facilitation may improve cognitive engagement, encourage metacognition and reduce the chances of learning passively or getting wrong information (9, 10).

New research backs up the idea that scaffolded AI learning is useful. Research conducted by Noroozi et al. and Jin et al. indicates that guided AI interactions enhance information retention, increase learner satisfaction and improve problem-solving skills in comparison to self-directed AI usage (9, 11). Moreover, ethical standards are progressively promoting human supervision in the implementation of AI within educational contexts to prevent bias and excessive dependence (13).

Even with these changes, there isn't much real-world data on how well tutor-guided AI models work in undergraduate medical education, especially in low- and middle-income countries (LMICs). To responsibly and effectively add AI technologies to medical curriculum, it's important to understand how students feel about them, how easy they are to use and how they affect learning.

This study evaluated the effectiveness of tutor-guided AI navigation in undergraduate learning across three domains: perception of the learning experience, usability and confidence with AI and satisfaction and overall impression.

Methods

Study Design

This research was conducted in two parts:

Quasi-experimental study: A pre-test/post-test design compared knowledge acquisition in students who underwent the tutor-guided AI session versus those who used ChatGPT independently (intervention: n = 59; control: n = 28).

Cross-sectional survey: Assessed students' perception after the tutor guided learning session.

Study Setting and Participants

The study was conducted at the Faculty of Medicine, University of Kelaniya, Sri Lanka among final-year undergraduate medical students. Participation was voluntary and informed consent was obtained from all participants. The study was approved by the institutional Ethics Review Committee (Ref No: P/116/06/2025).

Intervention

The intervention group participated in a 90-minute session that integrated ChatGPT-based learning facilitated by a tutor. The tutor provided structured prompts aligned with key learning objectives in Obstetrics and Gynaecology and guided the students in interpreting and critiquing AI-generated content. The control group used ChatGPT independently to explore the same content.

Assessments

Knowledge Tests: Students in both groups completed identical pre- and post-session MCQs.

A validated Single Best Answer (SBA) knowledge test was developed based on authoritative clinical guidelines (e.g., SLCOG, NICE NG3 and local obstetric diabetes protocols).

Perception Questionnaire: A pre-defined, self-administered questionnaire comprising 5-point Likert scale items was used to evaluate the learning experience, usability, confidence using AI, trust in AI-generated information, overall satisfaction and likelihood of recommending the method.

Statistical Analysis

All statistical analyses were performed using SPPS version 26. The Wilcoxon rank-sum test assessed within-group changes, while Welch's t-test compared the magnitude of knowledge gains between groups.

Hedges' g was used to estimate effect size. Reliability of perception items was evaluated using Cronbach's alpha.

Results

Knowledge Acquisition

The pre-test scores were statistically comparable between the two groups ($W = 670.5$, $p = 0.153$). Both groups demonstrated significant improvement from pre-test to post-test ($p < 0.001$), with median scores increasing from 4 to 7. However, the difference in the degree of improvement between the groups was not statistically significant (Welch's $t = 1.07$, $p = 0.144$), though the effect size (Hedges' $g = 0.36$) indicated a small-to-moderate benefit for the tutor-guided group (Fig. 1).

Students' Perception

The perception survey, completed by 80 students, showed high reliability across all domains ($\alpha > 0.86$). Students reported high levels of satisfaction and ease of use, with mean scores as depicted in Table 1 and Fig. 2.

Table 1
Students' Perception of Tutor Guided AI Assisted Learning

Response	Mean Score ± SD
Ease of use	4.19 ± 1.07
Experience of tutor-guided AI navigation	4.03 ± 1.13
Trust in AI accuracy	3.27 ± 0.91
Overall satisfaction	3.91 ± 0.94
Willingness to recommend	4.01 ± 0.97

Students particularly appreciated the structured nature of the sessions and the opportunity to discuss AI output critically with a tutor.

Discussion

Interpretation of Results

The results demonstrate that both tutor-guided and self-directed AI use can significantly improve short-term knowledge acquisition. While the difference between groups was not statistically significant, the structured guidance offered in the intervention appeared to enhance learning engagement and perceived value.

Educational Value of Tutor-Guided AI Navigation

Tutor-guided AI learning may bridge the gap between traditional pedagogy and emerging educational technologies. The presence of a tutor allowed for real-time clarification of misconceptions, critical appraisal of AI-generated information and reinforcement of learning through dialogue. This hybrid approach is especially relevant in settings where students may lack the digital literacy or confidence to engage meaningfully with AI tools. Guided use of AI fosters greater confidence among students in critically evaluating and applying AI-generated responses within academic and clinical contexts (14, 16, 17).

Moreover, the positive perception of this model among students underscores its potential scalability. As AI tools become increasingly integrated into educational platforms, faculty training and structured implementation will be key to maximizing their impact (15, 19).

Implications for Undergraduate Medical Education

The findings suggest that AI tools such as ChatGPT can be effectively integrated into undergraduate medical education without compromising learning outcomes. However, the added value of tutor guidance appears to lie not in immediate knowledge gains but in enhancing learning quality, engagement and professional formation. This supports the view that AI functions most effectively as a complementary educational tool, preserving the central role of human instruction while enriching the learning experience (18, 19).

This has important implications for curriculum design. Rather than viewing AI as a replacement for teaching, educational institutions should consider AI as a co-learning tool. Tutor-guided AI activities can promote deeper learning by encouraging questioning, contextualisation and reflection-skills essential for clinical practice. When structured appropriately, AI-supported learning has been shown to achieve outcomes comparable to expert-led instruction, reinforcing its potential as a scalable adjunct in undergraduate training (20).

The findings align with previous studies by Noroozi et al. and Jin et al., which suggest that structured AI use can enhance learning engagement (9, 11). Wadden and Seaborn have highlighted the ethical and pedagogical risks of unregulated AI use in education (7, 8). This study contributes novel data from a South Asian LMIC context, providing valuable insights for resource-limited settings.

This study had several limitations. The sample size of the control group was smaller, which may limit the generalisability of the findings and the statistical power to detect between-group differences. The study only measured immediate knowledge gain, the long-term retention and application in clinical contexts were not assessed.

The intervention consisted of a single session. Repeated exposure to tutor guided AI navigation may yield different outcomes. The perception survey demonstrated high reliability, the inclusion of qualitative interviews may provide deeper insight into student experience and learning processes.

Conclusion

Tutor-guided AI navigation using ChatGPT is a viable, effective and well-received approach for medical education. Although short-term knowledge gains were similar to self-directed AI use, the tutor-guided model improved satisfaction, critical engagement and perceived value. It represents a safe, structured way to integrate AI into undergraduate curricula while maintaining pedagogical and ethical oversight.

Study demonstrates that AI-assisted learning using ChatGPT is an effective educational tool for undergraduate medical students, resulting in significant short-term knowledge gains in a core Obstetrics and Gynaecology topic. Tutor-guided AI navigation did not confer a statistically significant advantage in immediate knowledge acquisition compared with self-directed AI use; however, it was associated with a superior learning experience, higher satisfaction and greater perceived educational value.

These findings suggest that the primary benefit of tutor-guided AI lies in enhancing engagement, confidence and reflective learning rather than solely improving test performance. Importantly, students exhibited a balanced level of trust in AI, reinforcing the need for structured guidance to support critical appraisal and professional judgment.

Recommendations

This study recommends integrating tutor-guided AI into undergraduate medical curricula as a structured, blended learning approach rather than an unstructured, self-directed tool. Incorporating AI into sessions like case-based discussions, problem-based learning and formative assessments can align with curricular goals while fostering critical thinking, contextual understanding and reflective learning. This approach positions AI as a complementary educational resource that enhances learning depth and consistency without undermining established pedagogical frameworks.

Effective implementation of tutor-guided AI learning requires systematic faculty development. Educators must be equipped with competencies in AI prompt formulation, verification of AI-generated content and facilitation techniques that encourage analytical thinking and reflective discussion.

In parallel, institutions should establish clear ethical guidelines to govern the responsible use of AI in medical education. Such frameworks should address issues of academic integrity, appropriate reliance on AI-generated information and the necessity of human oversight in learning and assessment. To inform evidence-based curriculum reform, future research should adopt longitudinal designs to evaluate the impact of tutor-guided AI learning on knowledge retention, clinical reasoning and decision-making in authentic clinical environments. These investigations will be critical in determining the sustained educational value of guided AI integration and in shaping best practices for its scalable implementation.

Abbreviations

- AI – Artificial Intelligence

- LLMs – Large Language Models
- LMICs – Low- and Middle-Income Countries
- MCQs – Multiple Choice Questions

Declarations

Ethics approval and consent to participate

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Ethical approval for this study was obtained from the Ethics Review Committee of the Faculty of Medicine, University of Kelaniya, Sri Lanka (Ref No: P/116/06/2025). Participation was voluntary and informed consent was obtained from all participants prior to data collection.

Consent for publication

Not applicable.

Competing interests

All authors wish to confirm that there are no known conflicts of interest associated with this publication and there has been no financial support for this work that could have influenced its outcome.

Funding

None

Author Contribution

Randombage P.J.S. , the lead author, contributed to conceptualization, methodology, project administration, supervision, writing, editing and revising the manuscript.

M.Y.S. Ahamed, T.D. Kulathunga and H. Wickramasekara contributed to data collection, data analysis, manuscript drafting and manuscript revision. All authors read and approved the final manuscript.

Acknowledgements

Institutions: Department of Obstetrics & Gynaecology, Faculty of Medicine, University of Kelaniya, Department of Medical Education, Faculty of Medicine, University of Kelaniya, Sri Lanka

Participants: Final-year medical students of the Faculty of Medicine, University of Kelaniya.

Data Availability

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Ellaway RH, Masters K. AMEE Guide 32: e-learning in medical education. Part 1: learning, teaching and assessment. Med Teach. 2008;30(5):455–73.
2. Tegos S, Demetriadis S. Conversational agents as learning companions: a review. Educ Technol Res Dev. 2022;70(1):1–28.
3. Chan KS, Zary N. Applications and challenges of artificial intelligence in medical education: an integrative review. JMIR Med Educ. 2019;5(1):e13811.
4. Park SH, Han K. Methodologic guide for evaluating artificial intelligence in medical education. Korean J Med Educ. 2023;35(1):1–10.
5. Wadden TA. Trust but verify: using artificial intelligence in education responsibly. Acad Med. 2023;98(6):707–10.
6. Alsoufi A, Alsuyihili A, Msherghi A, et al. Impact of the COVID-19 pandemic on medical education: medical students' knowledge, attitudes and practices regarding electronic learning. PLoS ONE. 2020;15(11):e0242905.
7. Shen Y, Heacock L, Elias J, et al. ChatGPT and other large language models are double-edged swords. Radiology. 2023;307(2):e230163.
8. Seaborn K, Fels DI. Ethical concerns in AI-enabled learning environments. AI Ethics. 2023;3(2):141–50.
9. Noroozi O, Biemans HJ, Mulder M, Chizari M. The effectiveness of structured AI-based feedback in learning complex tasks: a review. BMC Med Educ. 2023;23(1):112.
10. Holmboe ES, Sherbino J, Englander R, Snell L, Frank JR. A call to action: the controversy of and rationale for competency-based medical education. Med Teach. 2017;39(6):574–81.
11. Jin R, Yuan H, Zhang Y, Wang L. Artificial intelligence tutors and clinical reasoning development: a systematic review. Med Sci Educ. 2023;33(3):321–9.
12. Weng SF, Reps J, Kai J, Garibaldi JM, Qureshi N. Can machine learning improve cardiovascular risk prediction using routine clinical data? PLoS ONE. 2017;12(4):e0174944.
13. United Nations Educational, Scientific and Cultural Organization. Guidance for generative AI in education and research. Paris: UNESCO; 2023.
14. Sami A, Tanveer F, Sajwani K, et al. Medical students' attitudes toward artificial intelligence in education: perception, effectiveness and credibility. BMC Med Educ. 2025;25:82.
15. Naseer MA, Saeed S, Afzal A, et al. Integrating artificial intelligence into medical curricula: a mixed-methods study exploring student and faculty perspectives. BMC Med Educ. 2025;25:273.

16. Seneviratne T, Kodikara K, Abeykoon I, Palpola W. Medical undergraduates' perceptions of artificial intelligence in medical education: a qualitative study. JMIR Med Educ. 2025;11:e73798.
17. Jackson P, Ponath Sukumaran G, Babu C, et al. Artificial intelligence in medical education: perceptions of medical students. BMC Med Educ. 2024;24:804.
18. Li Q, Qin Y. Artificial intelligence in medical education: medical student perceptions and curriculum design recommendations. BMC Med Educ. 2023;23:852.
19. McLaughlin JE, Dal Ponte C, Lyons K. Student perceptions of generative AI as a virtual tutor to support collaborative research training for health professionals. BMC Med Educ. 2025;25:895.
20. Fazlollahi AM, Bakhaidar M, Alsayegh S, et al. Effect of artificial intelligence tutoring versus expert instruction on learning simulated surgical skills among medical students: a randomized clinical trial. JAMA Netw Open. 2022;5(2):e2149008.

Figures

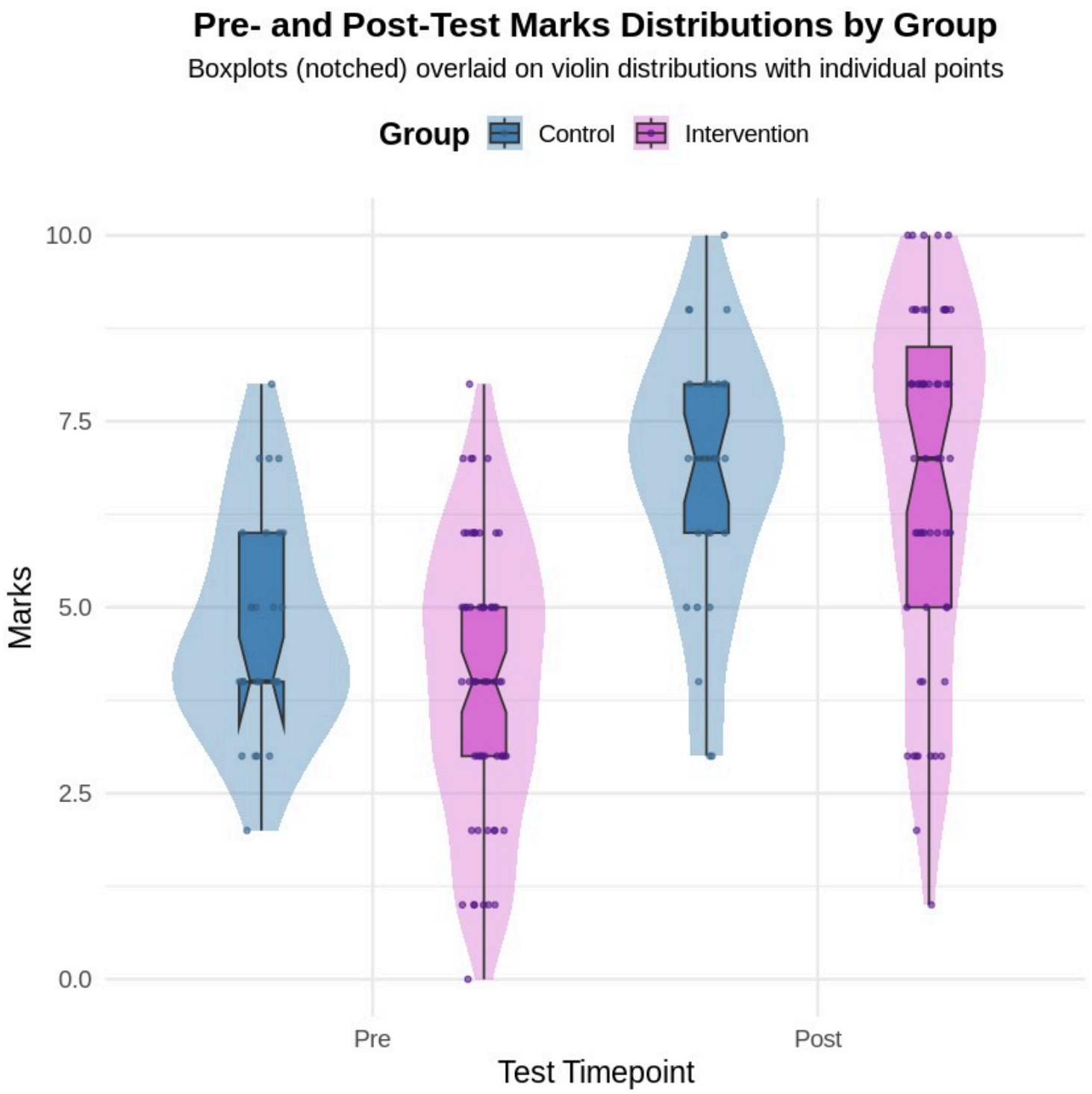


Figure 1

Pre and Post Test Marks Distribution by Groups

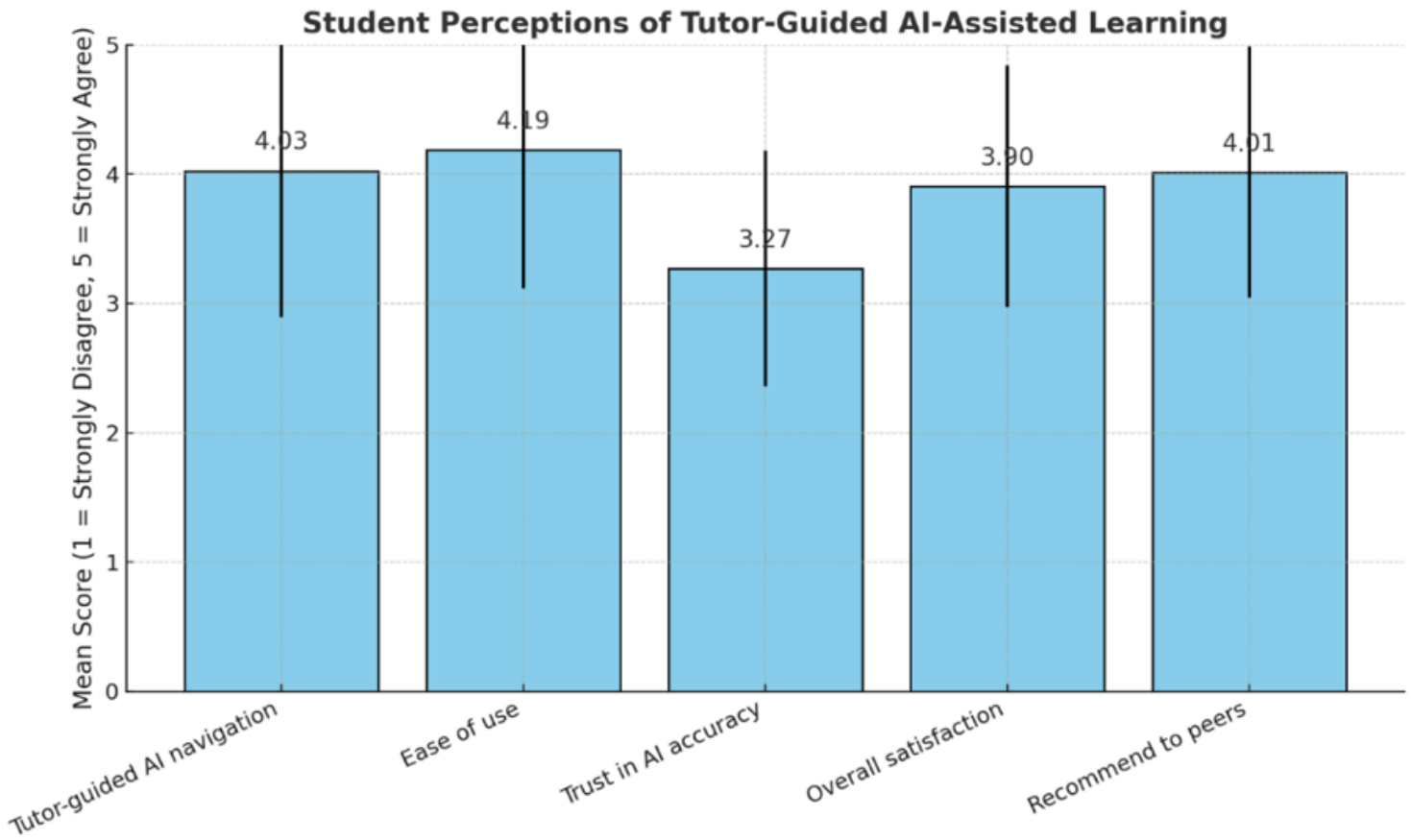


Figure 2

Students' Perception of Tutor Guided AI Assisted Learning

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- QuestionnaireforBMC.docx