COMMENT

Comment to: AI's deep dive into complex pediatric inguinal hernia issues—a challenge to traditional guidelines?

H. Daungsupawong[1] · V. Wiwanitkit[2]

Dear Editor,

We read with a great interest on "AI's deep dive into complex pediatric inguinal hernia issues: a challenge to traditional guidelines? [1]." The purpose of this study was to employ ChatGPT, an artificial intelligence software, to analyze disputed topics in pediatric inguinal hernia surgery and compare its responses to the European Association of Pediatric Surgeons (EUPSA) standards. Six contentious problems were examined, with two distinct responses generated for each. These comments were then compared to systematic studies and recommended practices. Validation was accomplished through content analysis and expert judgments. The study evaluated the consistency or disagreement between ChatGPT's responses and the guidelines, as well as the responses' quality, reliability, and applicability.

The results showed that ChatGPT generated responses that were mainly aligned with the standards, but there were occasional variations and conflicts. The average quality score was 3.33, while the average reliability score was 2.75. One weakness of this study is that it only compared ChatGPT's responses to EUPSA standards. To enable a full evaluation of the AI program's performance, a greater range of criteria and expert perspectives would be beneficial. More study should be conducted to improve the dependability and usability of AI algorithms such as ChatGPT in the medical industry. This could entail improving training methods and using more diverse and extensive datasets.

Furthermore, studies should look into the potential benefits and drawbacks of utilizing AI algorithms as decision support tools in clinical practice, taking into account aspects like patient preferences, personalised care, and ethical concerns. The shortcomings of ChatGPT in correctly reading data and answering scientific inquiries indicate the necessity for additional study and improvement of AI language models. When assessing ChatGPT-4's performance in a clinical setting, biases in the training data must also be taken into account. A chatbot could provide a false reference if there is insufficient human control or verification, which would cause new problems [2, 3]. Future studies should focus on enhancing the models' precision, openness, and ethical issues.

Declarations

Conflict of interest Authors declare no conflict of interest.

References

1. Wang G, Liu Q, Chen G, Xia B, Zeng D, Chen G, Guo C (2023) AI's deep dive into complex pediatric inguinal hernia issues: a challenge to traditional guidelines? Hernia. https://doi.org/10.1007/s10029-023-02900-1
2. Kleebayoon A, Wiwanitkit V (2023) Artificial intelligence, chatbots, plagiarism and basic honesty: comment. Cell Mol Bioeng 16(2):173174
3. Kleebayoon A, Wiwanitkit V (2023) Comment on "How may ChatGPT impact medical teaching?" Rev Assoc Med Bras (1992) 69(8):e20230593

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

✉ H. Daungsupawong
hinpetchdaung@gmail.com

✉ V. Wiwanitkit
wviroj@yahoo.com

1 Phonhong, Lao People's Democratic Republic

2 Chandigarh University, Mohali, Punjab, India