

Reply: Comment on “The new frontier: utilizing ChatGPT to expand craniofacial research”

Andi Zhang¹, Ethan Dimock², Rohun Gupta¹, Kevin Chen¹

¹Division of Plastic Surgery, Saint Louis University School of Medicine, St. Louis, MO, USA

²Oakland University William Beaumont School of Medicine, Rochester, MI, USA

Reply:

We would like to respond to the comment by Daungsupawong and Wiwanitkit on our published article “The new frontier: utilizing ChatGPT to expand craniofacial research” [1]. Our study demonstrated ChatGPT’s ability to generate novel systematic review ideas within the field of craniofacial surgery. Our results showed an average 57.5% total accuracy across both general and specific topics, 39% accuracy for general topics, and 76% for specific topics.

The authors in the reply letter discussed two main points: (1) the need for further investigation of ChatGPT’s ability to generate specialized research ideas and (2) the usage of only four well-known medical resources for the cross-referencing and assessment of ChatGPT’s accuracy. We will address each point individually below: we agree with the commenters that ChatGPT’s accuracy and dependability in producing research ideas were called into question by this study. The inability of ChatGPT to produce novel general topics highlights the difficulty of acquiring a specialized knowledge base, such as in craniofacial surgery, and the need to examine whether later iterations of ChatGPT with larger and more up-to-date training databases will perform better.

Regarding the decision to utilize four major medical literature databases to cross-reference ChatGPT’s accuracy for generating

novel research ideas; we believe that the current blend of broad and specialized medical literature databases is optimal and provides an accurate reflection of the state of literature currently available. In any type of literature search or review, a balance needs to be achieved between minimizing the manual search burden for the investigators and not missing relevant references, thereby reducing the validity of the research [2]. However, the authors agree that increasing the number of databases searched would contribute to a more comprehensive search. It is possible that by increasing the number of databases searched, ChatGPT’s accuracy might decrease slightly when previously considered novel ideas are found in the newly added medical literature databases.

Overall, we agree with the commentors’ observation that artificial intelligence (AI) in craniofacial research needs to be studied further. With the release of more advanced large language mode in recent months, their abilities to generate general and specialized topics need to be examined.

NOTES

Conflict of interest

No potential conflict of interest relevant to this article was reported.

Funding

None.

Acknowledgments

AI declaration: the author used a language editing computational tool in preparation of the article.

Correspondence:

Andi Zhang
Division of Plastic and Reconstructive Surgery, Saint Louis University School of Medicine, SLUCare Academic Pavilion 1008 S. Spring Ave, Suite 1500 St. Louis, MO 63110, USA
E-mail: andyzhang214@gmail.com

How to cite this article:

Zhang A, Dimock E, Gupta R, Chen K. Reply: Comment on “The new frontier: utilizing ChatGPT to expand craniofacial research”. Arch Craniofac Surg 2024;25(4):207-208. <https://doi.org/10.7181/acfs.2024.00423>

Received July 25, 2024 / Revised July 25, 2024 / Accepted August 10, 2024

ORCID

Andi Zhang <https://orcid.org/0009-0005-2851-0736>
Ethan Dimock <https://orcid.org/0009-0001-8423-3965>
Rohun Gupta <https://orcid.org/0000-0003-1491-2441>
Kevin Chen <https://orcid.org/0000-0003-2545-6381>

Author contributions

Conceptualization: Andi Zhang, Ethan Dimock, Rohun Gupta, Kevin Chen. Writing - original draft: Andi Zhang. Writing - review & editing: Andi Zhang, Ethan Dimock, Rohun Gupta, Kevin Chen.

REFERENCES

1. Zhang A, Dimock E, Gupta R, Chen K. The new frontier: utilizing ChatGPT to expand craniofacial research. *Arch Craniofac Surg* 2024;25:116-22.
2. Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Syst Rev* 2017;6:245.