Author's Response

Advancing AI Horizons: Scientific Conversations on Tympanoplasty Postoperative Management

Aynur Aliyeva[1,2]

[1]Division of Otorhinolaryngology-Head and Neck Surgery, The Catholic University St.Mary Hospital Medical Center, Seoul, South Korea
[2]The Neuroscience Doctoral Program, Yeditepe University, İstanbul, Türkiye

ORCID IDs of the author: A.A. 0000-0001-9398-4261.

Dear Editor,

Before addressing the scientific points raised by the reviewer, I would like to express my sincere gratitude for their thoughtful and detailed feedback on our manuscript.[1] Receiving such insightful comments enhances the impact of our publication, and I firmly believe that this type of academic dialogue significantly contributes to the advancement of science. This exchange improves the quality of our work and fosters collective growth in the otolaryngology community.

I would now like to address the reviewer's concerns point by point:

1. SAMPLE SIZE AND GENERALIZABILITY
The reviewer noted that our sample size of 10 postoperative questions was limited. While we agree that evaluating more questions would provide broader insights, we have acknowledged this limitation in our manuscript. It is also important to mention that many published studies have employed similar or even smaller sample sizes when evaluating ChatGPT in medical contexts. For example, Armbruster J. et al. used 10 questions evaluated by one medical professional; Hirosawa T. et al. used five differential diagnosis questions; Aliyeva A. et al. assessed five questions related to cochlear implantation; Grünebaum et al. evaluated 14 questions in obstetrics and gynecology; and Zhang et al. assessed 14 medical questions and Johnson SB et al. reviewed 13 questions related to cancer using ChatGPT.[2-7] While increasing the number of questions is undoubtedly beneficial, our study aligns with existing literature regarding sample size.

2. DIVERSITY OF MEDICAL PROFESSIONALS
The feedback in our study was obtained from ten otolaryngology specialists with varying years of experience. Although we acknowledge that including professionals from diverse medical backgrounds would enhance the evaluation process, our methodology is supported by previous studies. For instance, Liu et al. and Johnson SB et al. included only five clinicians to assess AI-generated medical responses.[7,8]

3. AWARENESS OF AI RESPONSES AND STATISTICAL ANALYSIS
The reviewer expressed concerns regarding potential bias, as our study did not clarify whether medical professionals were aware of the AI source during evaluation. We confirm that all evaluators were blinded to the AI source to minimize bias. While our study reported accuracy ratings ranging from 80% to 100%, consistent with existing literature, we acknowledge the importance of employing more comprehensive statistical methods. Additionally, we are currently working on a manuscript comparing ChatGPT-4 and Claude 3.5 Sonet in tinnitus management, which has been submitted and is under review. In that study, we

Corresponding author: Aynur Aliyeva, e-mail: aynur_aliyeva86@mail.ru

incorporate confidence intervals, p-values, and a broader range of statistical parameters. Once published, we would be delighted to engage in further discussions

Finally, I would like to mention that our publication stands as one of the first studies exploring the use of ChatGPT-4.0 in postoperative management for tympanoplasty patients. One key aspect that makes our study unique is that ChatGPT-4 was not pre-trained with specific guidelines or medical information related to tympanoplasty, underscoring its baseline capabilities. We hope this research is an initial stepping stone for future otolaryngology and AI-assisted patient management studies.

FUTURE DIRECTIONS
In the era of AI and machine learning, our primary focus, along with that of many medical researchers, is to develop strategies that effectively integrate AI into otolaryngology and medical practice while addressing ethical considerations, mitigating biases, and fostering trust in AI applications.[9-11] We also encourage researchers to utilize larger datasets, perform more comprehensive statistical analyses, and investigate a wider range of AI models tailored to postoperative management and other related medical fields.

Again, Thank you for the opportunity to engage in this meaningful scientific dialogue. It has been a great pleasure.

Sincerely,

Dr. Aynur Aliyeva

REFENECSES
1. Aliyeva A, Alaskarov E, Sari E. Postoperative management of tympanoplasty with ChatGPT-4.0. *J Int Adv Otol*. 2025;21(1):1-6. [CrossRef]
2. Armbruster J, Bussmann F, Rothhaas C, Titze N, Grützner PA, Freischmidt H. "Doctor ChatGPT, Can You Help Me?" the Patient's perspective: cross-sectional study. *J Med Internet Res*. 2024;26:e58831. [CrossRef]
3. Hirosawa T, Harada Y, Yokose M, Sakamoto T, Kawamura R, Shimizu T. Diagnostic accuracy of differential-diagnosis lists generated by Generative Pretrained Transformer 3 Chatbot for Clinical Vignettes with Common Chief Complaints: A Pilot Study. *Int J Environ Res Public Health*. 2023;20(4):3378. [CrossRef]
4. Aliyeva A, Sari E, Alaskarov E, Nasirov R. Enhancing postoperative cochlear implant care with ChatGPT-4: a study on artificial intelligence (ai)-assisted patient education and support. *Cureus*. 2024;16(2):e53897. [CrossRef]
5. Grünebaum A, Chervenak J, Pollet SL, Katz A, Chervenak FA. The exciting potential for ChatGPT in obstetrics and gynecology. *Am J Obstet Gynecol*. 2023;228(6):696-705. [CrossRef]
6. Zhang JS, Yoon C, Williams DKA, Pinkas A. Exploring the usage of ChatGPT among medical students in the United States. *J Med Educ Curricular Dev*. 2024;11:23821205241264695. [CrossRef]
7. Johnson SB, King AJ, Warner EL, Aneja S, Kann BH, Bylund CL. Using ChatGPT to evaluate cancer myths and misconceptions: artificial intelligence and cancer information. *JNCI Cancer Spectr*. 2023;7(2):pkad015. [CrossRef]
8. Liu S, Wright AP, Patterson BL, et al. Assessing the value of ChatGPT for clinical decision support optimization. *medRxiv*. 2023: 2023.02.21.23286254. [CrossRef]. Preprint posted online on February 23, 2023.2023.02.21.23286254.
9. Aliyeva A, Sari E. Be or not to be with ChatGPT? *Cureus*. 2023;15(11):e48366. [CrossRef]
10. Han JS, Lim JH, Kim Y, et al. Hearing rehabilitation with a chat-based mobile auditory training program in experienced hearing aid users: prospective randomized controlled study. *JMIR mHealth uHealth*. 2024;12:e50292. [CrossRef]
11. Aliyeva A. "Bot or Not": Turing problem in otolaryngology. *Cureus*. 2023;15(11):e48170. [CrossRef]