The Laryngoscope

Letter to the Editor

In Reference to *Utilization of Artificial Intelligence in the Creation of Patient Information on Laryngology Topics*

Dear Editor,

We would like to commend Tran et al.[1] for their insightful study evaluating the readability and quality of patient education materials on laryngology topics generated by ChatGPT-3.5 compared to the American Academy of Otolaryngology-Head and Neck Surgery. The study's findings that ChatGPT materials require a higher reading level and are less readable, yet provide comparable information quality, underscore the transformative potential and challenges of artificial intelligence (AI) in patient education. This work is particularly timely, as AI's role in healthcare communication continues to expand. The quality of ChatGPT's information in head and neck surgery has been the focus of numerous publications in recent months,[2–5] reflecting the growing interest and importance of this topic.

The study opens important avenues for discussion, especially considering the rapid pace of AI evolution. First, the authors utilized ChatGPT-3.5, whereas newer iterations, like GPT-4o, demonstrate significantly improved reasoning capabilities and consistently higher quality outputs.[6–8] These advances could address some of the shortcomings identified in the study, further narrowing the gap between AI-generated content and professional guidelines.

Second, existing evaluation tools for online information are not tailored to AI's unique attributes. Traditional tools often assess the authority of the source, a metric irrelevant to AI. Consequently, questions addressing authority were marked as "Not Applicable," limiting the scope of the analysis. Validated AI-specific assessment tools, such as the Quality Assessment of Medical AI[9] and the AI Patient Information Index,[10] are already proposed and could provide more nuanced and accurate evaluations of AI-generated content.

Lastly, the study's use of decontextualized prompts highlights the importance of contextualized inputs in harnessing AI's full potential. When prompts lack clarity about the intended audience, such as their educational level or specific needs, AI may generate outputs that are less readable or accessible to the average user.[11,12] This decontextualized approach could be a contributing factor to the poorer readability scores observed in the study. By specifying the audience, such as tailoring content for patients with average reading skills, AI can produce information that is more appropriately aligned with the user's capabilities and needs.[13] Addressing this limitation would not only improve readability but also enhance the overall utility of AI-generated patient education materials.

Future research incorporating advanced AI systems, validated assessment tools, and contextualized queries will be instrumental in optimizing AI's role in healthcare communication. This study is an essential step in that direction.

ACKNOWLEDGMENTS

The authors have nothing to report.

Luigi Angelo Vaira, MD, PhD
Giacomo De Riu, MD
Maxillofacial Surgery Operative Unit, Department of Medical, Surgical and Experimental Sciences, University of Sassari, Sassari, Italy

Antonino Maniaci, MD, PhD
Department of Medicine and Surgery, University of Enna Kore, Enna, Italy

Miguel Mayo-Yáñez, MD, PhD
Otorhinolaryngology, Head and Neck Surgery Department, Complexo Hospitalario Universitario A Coruña (CHUAC), A Coruña, Galicia, Spain

Alberto Maria Saibene, MD, MA
Otolaryngology Unit, Santi Paolo e Carlo Hospital, Department of Health Sciences, University of Milan, Milan, Italy

Carlos Miguel Chiesa-Estomba, MD, PhD, MS
Department of Otorhinolaryngology-Head & Neck Surgery, Hospital Universitario Donostia, San Sebastian, Spain

Jerome R. Lechien, MD, PhD, FACS
Department of Surgery, Mons School of Medicine, UMONS Research Institute for Health Sciences and Technology, University of Mons (UMons), Mons, Belgium
Department of Otolaryngology-Head Neck Surgery, Elsan Hospital, Paris, France

Send correspondence to Luigi Angelo Vaira, Viale San Pietro 43/B, Sassari, Italy. Email: lavaira@uniss.it

DOI: 10.1002/lary.32032

Editor's Note: This Manuscript was accepted for publication on December 02, 2024.
The authors have no other funding, financial relationships, or conflicts of interest to disclose.

This work has been developed within the framework of the project e.INS—Ecosystem of Innovation for Next Generation Sardinia (cod. ECS 00000038) funded by the Italian Ministry for Research and Education (MUR) under the National Recovery and Resilience Plan (NRRP)—MISSION 4 COMPONENT 2, "From research to business" INVESTMENT 1.5, "Creation and strengthening of Ecosystems of innovation," and construction of "Territorial R&D Leaders," CUP J83C21000320007.

BIBLIOGRAPHY

1. Tram Q-L, Huynh PP, Le B, Jiang N. Utilization of artificial intelligence in the creation of patient information on laryngology topics. *Laryngoscope*. 2025;135(4):1295-1300. https://doi.org/10.1002/lary.31891.
2. Vaira LA, Lechien JR, Abbate V, et al. Accuracy of ChatGPT-generated information on head and neck and Oromaxillofacial surgery: a multicenter collaborative analysis. *Otolaryngol Head Neck Surg*. 2024;170:1492-1503.
3. Lechien JR, Naunheim MR, Maniaci A, et al. Performance and consistency of ChatGPT-4 versus otolaryngologists: a clinical case series. *Otolaryngol Head Neck Surg*. 2024;170:1519-1526.
4. Lorenzi A, Pugliese G, Maniaci A, et al. Reliability of large language models for advanced head and neck malignancies management: a comparison between ChatGPT 4 and Gemini advanced. *Eur Arch Otorhinolaryngol*. 2024;281:5001-5006. https://doi.org/10.1007/s00405-024-08746-2.
5. Teixeira-Marques F, Medeiros N, Nazaré F, et al. Exploring the role of ChatGPT in clinical decision-making in otorhinolaryngology: a ChatGPT designed study. *Eur Arch Otorhinolaryngol*. 2024;281(4):2023-2030.
6. Lechien JR, Briganti G, Vaira LA. Accuracy of ChatGPT-3.5 and −4 in providing scientific references in otolaryngology-head and neck surgery. *Eur Arch Otorhinolaryngol*. 2024;281:2159-2165.
7. Mayo-Yáñez M, Lechien JR, Maria-Saibene A, Vaira LA, Maniaci A, Chiesa-Estomba CM. Examining the performance of ChatGPT 3.5 and Microsoft copilot in otolaryngology: a comparative study with otolaryngologists' evaluation. *Indian J Otolaryngol Head Neck Surg*. 2024;76(4):3465-3469.
8. Massey PA, Montgomery C, Zhang AS. Comparison of ChatGPT-3.5, ChatGPT-4, and Orthopaedic resident performance on Orthopaedic assessment examinations. *J Am Acad Orthop Surg*. 2023;31(23):1173-1179.
9. Vaira LA, Lechien JR, Abbate V, et al. Validation of the quality analysis of medical artificial intelligence (QAMAI) tool: a new tool to assess the quality of health information provided by AI platforms. *Eur Arch Otorhinolaryngol*. 2024;281(11):6123-6131.
10. Lechien JR, Maniaci A, Gengler I, Hans S, Chiesa-Estomba CM, Vaira LA. Validity and reliability of an instrument evaluating the performance of intelligent chatbot: the artificial intelligence performance instrument (AIPI). *Eur Arch Otorhinolaryngol*. 2024;281:2063-2079.
11. Campbell DJ, Estephan LE, Sina EM, et al. Evaluating ChatGPT responses on thyroid nodules for patient education. *Thyroid*. 2024;34(3):371-377.
12. Lee TJ, Campbell DJ, Rao AK, et al. Evaluating ChatGPT responses on atrial fibrillation for patient education. *Cureus*. 2024;16(6):e61680.
13. Vaira LA, Lechien JR, Abbate V, et al. Enhancing AI chatbot responses in healthcare: the SMART prompt structure in head and neck surgery. *OTO Open*. 2025;9(1):e70075. https://doi.org/10.21203/rs.3.rs-4953716/v1.