Introduction

In November 2022, OpenAI released the public ChatGPT, a chatbot with the ability to communicate in plain English and called by some as a “tipping point for AI.” Far from being a comprehensive review, the intention of this study is a cursory review of ChatGPT’s diagnostic capability, ability to convey pathophysiology in simple terms, accuracy in management recommendations, and appropriateness in follow up and post operative recommendations in common otolaryngologic conditions.

Methods and Materials

Assessment of ChatGPT’s current diagnostic capability, ability to convey pathophysiology in simple terms, accuracy in management recommendations, and appropriateness in follow-up and postoperative recommendations in common otolaryngologic conditions was explored in December 2022. A qualitative analysis of adenotonsillectomy (T&A), tympanoplasty (TP), endoscopic sinus surgery (ESS), parotidectomy (PT), and total laryngectomy (TL) was performed.

We asked ChatGPT:
• “How do I know if I need (procedure)?”
• “What are treatment alternatives to (procedure)?”
• “What are the risks of (procedure)?”
• “How is (procedure) performed?”
• “What is the recovery process for (procedure)?”

Results

<table>
<thead>
<tr>
<th>Procedure</th>
<th>How to know if I need...?</th>
<th>What are the treatment alternatives to...?</th>
<th>What are the risks of...?</th>
<th>How is it performed?</th>
<th>What is the recovery process for...?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tonsillectomy</td>
<td>Positive: accurate generalized procedure description; safely recommends need for physician evaluation</td>
<td>Positive: suggested watching and waiting and antibiotics; mentioned alternative treatments alleviate symptoms but do not treat the underlying problem</td>
<td>Positive: covered appropriate risks; discussed issues of swallowing and dehydration</td>
<td>Positive: mentions techniques including cold knife, electrocautery, and ablation</td>
<td>Positive: describes accurate post-op course, pain, recovery, discomfort, rest, and saltwater rinses</td>
</tr>
<tr>
<td>Tymanoplasty</td>
<td>Positive: accurate procedure description; discussed details of possible causes of cholesteatoma</td>
<td>Negative: no aberrant suggestions</td>
<td>Negative: no aberrant risks mentioned</td>
<td>Positive: describes postauricular and transcanal approaches; covers accurate steps</td>
<td>Negative: no mention of avoiding insufflation or Valsalva</td>
</tr>
<tr>
<td>Endoscopic Sinus Surgery</td>
<td>Positive: highlights need to fail medical therapy then proceed with surgical workup</td>
<td>Negative: isolates indications to chronic sinuses</td>
<td>Positive: highlights major risks including graft failure and hearing loss</td>
<td>Positive: focuses on facial nerve injury and results of surgery</td>
<td>Positive: states need for general anesthesia, neck incision, and dissection of the facial nerve</td>
</tr>
<tr>
<td>Parotidectomy</td>
<td>Positive: provides basic anatomy and function of parotid; mentions surgical evaluation by ENT</td>
<td>Negative: focuses on masses; does not mention chronic parotid diseases as an indication</td>
<td>Positive: states options based on specific pathology; mentions radiation and chemotherapy alternative; emphasizes surgery as best option for cancer</td>
<td>Positive: focuses on facial nerve injury and results of surgery</td>
<td>Positive: accurately describes incision care and pain management</td>
</tr>
<tr>
<td>Total Laryngectomy</td>
<td>Positive: Mentions only for cancer of larynx; resection size depending on tumor size and characteristics</td>
<td>Negative: no aberrant surgical indications</td>
<td>Positive: highlights bleeding, infection, and damage to surrounding structures</td>
<td>Positive: highlights basic steps of procedure including incisions and dissection</td>
<td>Positive: highlights recovery basics and rehabilitation of speech and swallowing</td>
</tr>
</tbody>
</table>

Discussion

• In terms of management recommendations, ChatGPT was able to give generalized statements of evaluation, need for intervention, and the basics of the procedure without major aberrant errors or risks of safety.
• ChatGPT was successful in providing appropriate treatment alternatives in all procedures tested.
• When queried for methodology, risks, and procedural steps, ChatGPT lacked precision in the order of procedural steps, missed key surgical details, and did not accurately provide all major risks of each procedure.
• In terms of the recovery process, ChatGPT showed promise in T&A, TP, ESS, and PT but struggled in the complexity of TL, stating the patient could speak immediately after surgery without speech therapy.

Conclusions

• ChatGPT accurately demonstrated the need for intervention, management recommendations, and treatment alternatives in common ENT procedures.
• However, ChatGPT was not able to replace an otolaryngologist’s clinical reasoning necessary to discuss procedural methodology, risks, and the recovery process in complex procedures.
• As AI becomes further integrated into healthcare, we should continue to explore its indications, evaluate its limits, and refine its use to the otolaryngologist’s advantage.

Contact
Michael E. Hoffer, MD, FACS
University of Miami Health System
Professor of Otolaryngology and Neurological Surgery
Division of Neurootology
1120 NW 14th St, Miami, FL 33136
meh630@med.miami.edu
(305) 243-1484

References

Disclosures

The authors nor the University of Miami Health System are affiliated with OpenAI or the ChatGPT software utilized in this study. The queries and answers provided by OpenAI’s ChatGPT Version 3.5 are not representative of the views of the University of Miami Health System, the Department of Otolaryngology at the University of Miami Health System, or the authors of the study.