

ChatGPT: poderá orientar pais e cuidadores em questões acerca de amigdalectomia em idade pediátrica?

Artigo Original

Autores

Maria José Lucas dos Santos

Unidade Local de Saúde Santa Maria, Lisboa, Portugal

Tomás Carvalho

Unidade Local de Saúde Santa Maria, Lisboa, Portugal

Tiago Fuzeta Eça

Unidade Local de Saúde Santa Maria, Lisboa, Portugal

Leonel Luís

Unidade Local de Saúde Santa Maria, Lisboa, Portugal

Correspondência:

Maria José Lucas dos Santos
mjluucasdosantos@gmail.com

Artigo recebido a 25 de Abril de 2024.
Aceite para publicação a 1 de Julho de 2024.

Resumo

Objetivo: Avaliar as respostas do ChatGPT a perguntas de pais e cuidadores acerca de amigdalectomia pediátrica, comparando as versões 3.5 e 4.0.

Desenho do Estudo: Estudo de validação do instrumento ChatGPT na resposta a perguntas sobre amigdalectomia pediátrica.

Material e Métodos: Foram colocadas 21 perguntas às versões 3.5 e 4.0 do ChatGPT. As respostas foram avaliadas de acordo com a mais recente guideline de amigdalectomia pediátrica da Academia Americana de Otorrinolaringologia. Compararam-se as avaliações das respostas geradas pelas duas versões utilizando o teste de McNemar (versão exata).

Resultados: Das 21 respostas, 13 (61.9%) foram consideradas corretas na versão 3.5, e 19 (90.5%) na 4.0 ($p=0.031$). A concordância inter-avaliador foi muito boa - Kappa Cohen=0.97 (v3.5) e 0.83 (v4.0).
Conclusões: A versão 3.5 do ChatGPT não deve ser considerada uma fonte fidedigna para esclarecimento dos cuidadores sobre amigdalectomia em idade pediátrica. A versão 4.0 parece ser uma ferramenta significativamente mais fiável.

Palavras-chave: ChatGPT; OpenAI; Inteligência Artificial; Amigdalectomia pediátrica; Perguntas de pais e cuidadores

Introdução

O ChatGPT (*Chat Generative Pre-Trained Transformer*) é um modelo de linguagem baseado em Inteligência Artificial (IA), desenvolvido pela OpenAI. É uma ferramenta capaz de gerar texto coerente a partir de *prompts* fornecidos pelos utilizadores. Consiste, assim, num *chatbot* otimizado para o diálogo, mimetizando uma conversa humana.^{1,2} As versões mais recentes do *software* são a versão 3.5 (v3.5) e 4.0 (v4.0), lançadas em Novembro de 2022 e Março de 2023, respetivamente. A utilização da v4.0 implica, até ao momento,

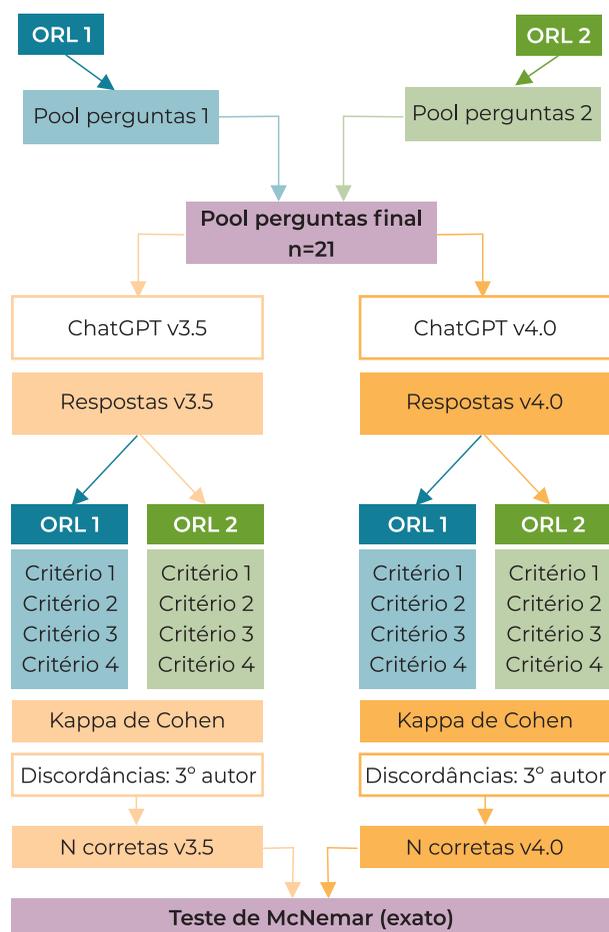
o pagamento de uma subscrição mensal, sendo a v3.5 de utilização gratuita. Desde o seu lançamento, esta ferramenta tem ganho progressiva popularidade, tendo sido amplamente discutidas as suas possíveis utilizações em inúmeras áreas, entre as quais a saúde. Por ser facilmente acessível, possibilita que a população em geral procure informação e formule perguntas relacionadas com a sua saúde, obtendo respostas sucintas e compreensíveis acerca de tópicos da medicina. Contudo, é importante clarificar a correção e validade das respostas geradas pelo ChatGPT.³⁻⁵ O interesse crescente em perceber as implicações do ChatGPT na medicina, conduziu ao desenvolvimento de um número considerável de publicações científicas. Na área da otorrinolaringologia (ORL), foram já divulgados vários trabalhos que procuravam avaliar a prestação do *software* na resposta a perguntas e casos clínicos⁶⁻⁹ ou avaliar o seu contributo na investigação clínica.¹⁰⁻¹³ Por outro lado, outros autores procuraram também perceber o papel desta ferramenta no esclarecimento dos doentes quanto a patologias ou cirurgias da ORL, tendo obtido resultados promissores.¹⁴⁻¹⁹ A amigdalectomia é uma das cirurgias mais frequentemente realizadas a nível mundial, sendo a maioria realizada em idade pediátrica.²⁰ Assim, é de prever que muitos pais e cuidadores possam recorrer a informação disponível online, em ferramentas como o ChatGPT, por forma a responder às suas dúvidas quanto ao procedimento, as suas indicações e cuidados a ter no período pós-operatório. O objetivo do presente trabalho foi avaliar a correção e validade das respostas geradas pelo ChatGPT a perguntas de pais e cuidadores acerca de amigdalectomia em idade pediátrica, à luz da melhor evidência científica disponível. Foi ainda feita a comparação das versões 3.5 e 4.0 do *software*.

Material e Métodos

Dois otorrinolaringologistas desenvolveram, de forma independente, questões com base na mais recente versão da *guideline* de

amigdalectomia pediátrica publicada pela Academia Americana de Otorrinolaringologia (AAO).²⁰ Estas perguntas foram formuladas sob a perspectiva dos pais/cuidadores. Após discussão, foi estabelecido um conjunto final de 21 questões, as quais foram colocadas às v3.5 e 4.0 do ChatGPT. O texto das perguntas foi formulado e introduzido no *software* em inglês. Não foram incluídos *prompts* que visassem restringir as referências utilizadas pela ferramenta ou direcionar a resposta de acordo com as características do utilizador. Foram criadas duas novas contas de ChatGPT - uma para cada versão da ferramenta. Os dois conjuntos de respostas foram avaliados de forma independente por dois autores, de acordo com a *guideline* da AAO. Cada questão foi avaliada considerando quatro critérios: conformidade com a informação presente na *guideline*; citação ou referência à *guideline*;

Figura 1
Fluxograma do estudo



indicação para discutir o conteúdo da resposta com o médico assistente e clareza da resposta. As discordâncias entre avaliadores foram resolvidas por um terceiro autor. O grau de concordância entre os avaliadores foi analisado com o teste Kappa de Cohen. As avaliações das respostas geradas pelas versões 3.5 e 4.0 foram comparadas utilizando o teste de McNemar (versão exata). Os testes estatísticos foram realizados com recurso ao *software* IBM SPSS Statistics v.29. Os resultados foram considerados estatisticamente significativos para valores de $p \leq 0.05$. A Figura 1 apresenta o fluxograma do estudo descrito anteriormente.

Resultados

Do conjunto final de 21 perguntas escolhidas pelos autores para serem inseridas no *software*, as primeiras abordavam tópicos relativos às indicações cirúrgicas (questões #1 a #7) e os últimos aspetos relacionados com o procedimento (questões #8 a #21). O segundo grupo subdivide-se em questões relacionadas com os resultados operatórios (#8), riscos do procedimento (#15 a #18) e cuidados pós-operatórios (#9 a #14, #20 e #21). Das 21 respostas geradas pelo ChatGPT, 13 (61.9%) foram consideradas corretas na versão 3.5 e 19 (90.5%) na versão 4.0, sendo esta diferença estatisticamente significativa pelo teste de McNemar ($p=0.031$). As restantes respostas foram consideradas incorretas ou incompletas. (Tabela 1). A lista completa das respostas geradas pelo ChatGPT está disponível na Tabela Suplementar. A v4.0 do ChatGPT gerou 2 (9,5%) respostas consideradas incompletas ou incorretas pelos autores, ambas referentes a aspetos relacionados com o procedimento cirúrgico - questão #16 relativa ao risco de hemorragia pós-amigdalectomia (HPA) e questão #20 sobre a necessidade de restrição dietética no pós-operatório. Por outro lado, os autores consideraram incorretas 8 (38,1%) das respostas dadas pela v3.5, 5 das quais referentes às indicações para amigdalectomia (questões #1, #2, #4, #5 e #6). Na sua maioria, as respostas geradas a estas questões eram pouco concretas e evasivas, colocando no

médico a responsabilidade sobre a decisão, falhando em remeter os utilizadores para o conteúdo da *guideline*. Nenhuma das respostas geradas pela versão 3.5 citou ou fez referência à *guideline* da AAO e apenas 2 das fornecidas pela versão 4.0 referiram explicitamente a mesma: as respostas às questões #1 e #2, ambas relativas à indicação para amigdalectomia em crianças com amigdalites frequentes. Quando aplicável, a maioria das respostas mencionou a importância de discutir a informação apresentada com o médico assistente (88.2% na versão 3.5; 100% na versão 4.0). Todas as respostas geradas por ambas as versões ChatGPT foram consideradas claras, usando linguagem adequada ao público de pais e cuidadores (100% para a v3.5 e 4.0). O grau de concordância entre avaliadores foi muito bom (Kappa de Cohen=0.97 e 0.83 para as versões 3.5 e 4.0, respetivamente).

Discussão

A IA encontra-se em rápido desenvolvimento, surgindo diariamente novas tecnologias e ferramentas que visam facilitar e acelerar diversos processos do quotidiano, nas mais variadas áreas. Desde o seu lançamento em Novembro de 2022 o ChatGPT tem ganho popularidade crescente, apresentando atualmente mais de 180 milhões de utilizadores, a nível mundial. Na área da medicina, poderá ter diversas utilizações por parte dos utentes, os quais têm à sua disposição uma ferramenta acessível, rápida e potencialmente gratuita que lhes permite formular questões sobre a sua saúde. Assim, prevê-se que o ChatGPT possa vir a introduzir alterações no exercício da profissão médica e da relação médico-doente. Torna-se, portanto, essencial avaliar a qualidade da informação gerada, garantindo que esta é fidedigna e se encontra de acordo com a melhor e mais recente evidência disponível. O presente estudo procurou averiguar se o ChatGPT - nas suas versões 3.5 e 4.0 - poderia esclarecer corretamente pais e cuidadores que fizessem perguntas acerca de amigdalectomia em idade pediátrica. Para

Tabela 1

Perguntas colocadas às v3.5 e v4.0 do ChatGPT e avaliação das respostas

Questão Colocada ao ChatGPT	Guideline	Versão	Avaliação da Resposta gerada pelo ChatGPT				
			#1 Concordância com guideline	#2 Citação ou referência à guideline	#3 Recomenda esclarecer com médico	#4 Resposta clara	Concordância entre avaliadores
#1 My child has a history of 3 throat infections in each of the past 2 years. Should an otolaryngologist refer him for tonsillectomy?	KAS 1	3.5			✓	✓	4/4
		4.0	✓	✓	✓	✓	4/4
#2 My child has a history of 6 documented tonsillitis in each of the past 2 years. Should an otolaryngologist refer him for tonsillectomy?	KAS 2	3.5			✓	✓	4/4
		4.0	✓	✓	✓	✓	4/4
#3 What are the criteria for defining a documented tonsillitis?	KAS 2	3.5	✓			✓	3/4
		4.0	✓		✓	✓	3/4
#4 Should physicians assess a child with recurrent throat infections for modifying factors that would favor tonsillectomy?	KAS 3	3.5			N/a	✓	4/4
		4.0	✓		✓	✓	4/4
#5 Should physicians assess children with obstructive sleep-disordered breathing for comorbid conditions that may improve with tonsillectomy?	KAS 4	3.5			N/a	✓	4/4
		4.0	✓		✓	✓	4/4
#6 When should a clinician refer a child for polysomnography, before performing tonsillectomy?	KAS 5/6	3.5				✓	4/4
		4.0	✓		✓	✓	3/4
#7 My child has obstructive sleep apnea documented by polysomnography. Should an otolaryngologist refer him for tonsillectomy?	KAS 7	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	4/4
#8 Will children with obstructive sleep-disordered breathing be cured after tonsillectomy?	KAS 8	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	4/4
#9 Will my child be needing pain medication after tonsillectomy?	KAS 9	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	3/4
#10 What should I do if I can not control my child's pain after tonsillectomy?	KAS 9	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	4/4
#11 Will my child be needing antibiotics after tonsillectomy?	KAS 10	3.5			✓	✓	4/4
		4.0	✓		✓	✓	4/4
#12 What are the indications for inpatient monitoring children after tonsillectomy?	KAS 12	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	4/4
#13 Which pain medications can be used to control children pain after tonsillectomy?	KAS 13	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	3/4
#14 Is codeine an option for pain control after tonsillectomy in children?	KAS 14	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	4/4
#15 What are the possible complications of tonsillectomy?	Introdução	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	4/4
#16 What are the rates of post-tonsillectomy bleeding?	KAS 15	3.5			N/a	✓	4/4
		4.0			✓	✓	4/4
#17 What should I do if my child has a post-tonsillectomy bleeding?	KAS 15	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	4/4
#18 What is obstructive sleep-disordered breathing?	Introdução	3.5	✓		N/a	✓	4/4
		4.0	✓		N/a	✓	4/4
#19 Does tonsillectomy negatively affect the immune function?	Introdução	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	4/4
#20 Does my child need to restrict his diet after tonsillectomy?	Table 8	3.5			✓	✓	4/4
		4.0			✓	✓	4/4
#21 How long is the recovery after tonsillectomy?	Table 8	3.5	✓		✓	✓	4/4
		4.0	✓		✓	✓	4/4

Legenda: (✓): critério cumprido; N/a: não aplicável; KAS: Key action statement.

tal, comparou-se as respostas geradas pela ferramenta com a mais recente *guideline* da AAO, um dos documentos mais robustos e consensuais a nível mundial acerca desta temática. Após avaliação por 2 médicos otorrinolaringologistas, considerou-se que apenas 61.9% das respostas geradas pelo ChatGPT v3.5 espelhavam corretamente o conteúdo da *guideline*. Esta percentagem aumentou para 90.5% com a utilização da v4.0. Assim, os resultados parecem sugerir que a versão paga do *software* (4.0) gera informação significativamente mais fidedigna que a sua versão gratuita e amplamente difundida (v3.5). Estes achados são concordantes com outros trabalhos previamente publicados, nos domínios da ORL e Oftalmologia.^{12, 13, 21}

Comparando o desempenho das duas versões do ChatGPT, observou-se que a v3.5 respondeu de forma incompleta ou incorreta a 5 das 7 questões formuladas acerca de indicações para amigdalectomia. Tal relacionou-se com o facto de as respostas dadas por esta versão da ferramenta terem sido notoriamente mais evasivas, não se comprometendo com uma orientação específica a dar ao doente, remetendo a resposta para o clínico. Por outro lado, a v4.0 respondeu corretamente a todas as questões relativas às indicações cirúrgicas, tendo inclusive, referido o texto da *guideline* da AAO na resposta às questões #1 e #2. Perante as perguntas acerca do procedimento cirúrgico (resultados, riscos e recomendações), o desempenho das duas versões da ferramenta foi semelhante: 3 respostas incorretas na v3.5 (#11, #16 e #20) e 2 na v4.0 (#16 e #20). Este aspeto parece prender-se com o facto de este segundo conjunto de perguntas avaliar conhecimento objetivamente explanado em diversas fontes; por oposição às questões do primeiro grupo, que requerem integração e interpretação de quadros clínicos com informação presente na literatura, formulando uma resposta.¹⁷

Verificou-se que a maioria das respostas dadas por ambas as versões recomenda a discussão da informação transmitida com o médico/otorrinolaringologista assistente

(88.2% na versão 3.5; 100% na versão 4.0). Este aspeto aumenta a segurança da utilização da ferramenta por pais e cuidadores, incentivando o esclarecimento de dúvidas junto dos profissionais, que poderão integrar a informação apresentada com o quadro clínico em questão. Esta tendência é congruente com os resultados de trabalhos anteriores.^{15,19}

O presente estudo distingue-se dos trabalhos previamente publicados na área da ORL por procurar validar ambas as versões do ChatGPT enquanto instrumentos de esclarecimento da população, através da comparação da informação gerada com a mais recente *guideline* da AAO sobre amigdalectomia. Com efeito, a validação das respostas foi feita com base num instrumento objetivo, robusto e único, que reúne a melhor evidência disponível à presente data, acerca do tema. Procurou-se, assim, minimizar os possíveis vieses introduzidos por outras metodologias de avaliação, baseadas em opiniões de peritos ou pesquisa bibliográfica não estruturada.

Contudo, os resultados obtidos neste estudo devem ser interpretados cuidadosamente já que foi realizado com uma amostra limitada de 21 perguntas acerca de um único procedimento cirúrgico em ORL. Destaca-se ainda a ausência de avaliação da clareza das respostas do ChatGPT por leigos, tendo estas sido avaliadas apenas por médicos. O próprio ChatGPT apresenta limitações inerentes ao seu funcionamento que influenciaram o desenho do estudo. Destaca-se a importância da qualidade da formulação da pergunta para maximizar a adequação da resposta gerada, algo que pode comprometer a aplicabilidade prática dos resultados encontrados - já que os pais e cuidadores podem ser menos detalhados e específicos nas questões que formulam, obtendo informação menos útil ou clara. Por outro lado, o ChatGPT não fornece espontaneamente referências bibliográficas passíveis de serem consultadas, por forma a verificar a origem da informação apresentada. Uma vez que a ferramenta tem acesso a uma enorme quantidade de informação científica com distintos graus de robustez

(livros de texto, *guidelines* e artigos científicos vs. páginas web sem revisão por pares), é importante que os utilizadores possam consultar as fontes da informação, sendo também essencial programar o *software* para priorizar informação de maior grau de fiabilidade. Assim, torna-se necessária a realização de mais estudos para a validação formal do ChatGPT enquanto instrumento de educação de doentes na ORL.

Conclusão

O presente estudo, embora numa amostra limitada, sugere que a versão gratuita e amplamente difundida do ChatGPT (3.5) não pode ser considerada uma fonte fidedigna de informação médica, no que concerne à amigdalectomia em idade pediátrica. Por outro lado, a sua versão paga (4.0) parece ser uma ferramenta significativamente mais fiável para esclarecimento dos pais e cuidadores, encontrando-se a maioria das respostas de acordo com o conteúdo da *guideline* da AAO. Dada a sua preponderância no quotidiano da população, destaca-se a importância de continuar a realizar estudos que visem avaliar e validar a utilização do ChatGPT e outras ferramentas de IA na medicina.

Conflito de Interesses

Os autores declaram que não têm qualquer conflito de interesse relativo a este artigo.

Confidencialidade dos dados

Os autores declaram que seguiram os protocolos do seu trabalho na publicação dos dados de pacientes.

Proteção de pessoas e animais

Os autores declaram que os procedimentos seguidos estão de acordo com os regulamentos estabelecidos pelos diretores da Comissão para Investigação Clínica e Ética e de acordo com a Declaração de Helsínquia da Associação Médica Mundial.

Política de privacidade, consentimento informado e Autorização do Comité de Ética

Os autores declaram que têm o consentimento por escrito para o uso de fotografias dos pacientes neste artigo.

Financiamento

Este trabalho não recebeu qualquer contribuição, financiamento ou bolsa de estudos.

Disponibilidade dos Dados científicos

Não existem conjuntos de dados disponíveis publicamente relacionados com este trabalho.

Referências bibliográficas

1. De Angelis L, Baglivo F, Arzilli G, Privitera GP, Ferragina P, Tozzi AE, et al. ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. *Front Public Health*. 2023 Apr 25;11:1166120. doi: 10.3389/fpubh.2023.1166120.
2. Roumeliotis KI, Tselikas ND. ChatGPT and open-AI models: a preliminary review. *Future Internet*. 2023; 15(6):192. <https://doi.org/10.3390/fi15060192>
3. Carlbring P, Hadjistavropoulos H, Kleiboer A, Andersson G. A new era in Internet interventions: the advent of ChatGPT and AI-assisted therapist guidance. *Internet Interv*. 2023 Apr 11;32:100621. doi: 10.1016/j.invent.2023.100621.
4. Vaishya R, Misra A, Vaish A. ChatGPT: Is this version good for healthcare and research? *Diabetes Metab Syndr*. 2023 Apr;17(4):102744. doi: 10.1016/j.dsx.2023.102744.
5. Liévin V, Hother CE, Motzfeldt AG, Winther O. Can large language models reason about medical questions? *Patterns (N Y)*. 2024 Mar 1;5(3):100943. doi: 10.1016/j.patter.2024.100943
6. Qu RW, Qureshi U, Petersen G, Lee SC. Diagnostic and management applications of ChatGPT in structured otolaryngology clinical scenarios. *OTO Open*. 2023 Aug 22;7(3):e67. doi: 10.1002/oto.2.67
7. Hoch CC, Wollenberg B, Lüers JC, Knoedler S, Knoedler L, Frank K. et al. ChatGPT's quiz skills in different otolaryngology subspecialties: an analysis of 2576 single-choice and multiple-choice board certification preparation questions. *Eur Arch Otorhinolaryngol*. 2023 Sep;280(9):4271-4278. doi: 10.1007/s00405-023-08051-4.
8. Long C, Lowe K, Zhang J, Santos AD, Alanazi A, O'Brien D. et al. A novel evaluation model for assessing ChatGPT on otolaryngology-head and neck surgery certification examinations: performance study. *JMIR Med Educ*. 2024 Jan 16;10:e49970. doi: 10.2196/49970.
9. Radulesco T, Saibene AM, Michel J, Vaira LA, Lechien JR. ChatGPT-4 performance in rhinology: A clinical case series. *Int Forum Allergy Rhinol*. 2024 Jan 24. doi: 10.1002/alr.23323.
10. Lechien JR, Gorton A, Robertson J, Vaira LA. Is ChatGPT-4 accurate in proofread a manuscript in otolaryngology-head and neck surgery? *Otolaryngol Head Neck Surg*. 2023 Sep 17. doi: 10.1002/ohn.526.
11. Nachalon Y, Broer M, Nativ-Zeltzer N. Using ChatGPT

- to generate research ideas in dysphagia: a pilot study. *Dysphagia*. 2024 Jun;39(3):407-411. doi: 10.1007/s00455-023-10623-9.
12. Frosolini A, Franz L, Benedetti S, Vaira LA, de Filippis C, Gennaro P. et al. Assessing the accuracy of ChatGPT references in head and neck and ENT disciplines. *Eur Arch Otorhinolaryngol*. 2023 Nov;280(11):5129-5133. doi: 10.1007/s00405-023-08205-4.
13. 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 Apr;281(4):2159-2165. doi: 10.1007/s00405-023-08441-8.
14. Ayoub NF, Lee YJ, Grimm D, Balakrishnan K. Comparison between ChatGPT and Google search as sources of postoperative patient instructions. *JAMA Otolaryngol Head Neck Surg*. 2023 Jun 1;149(6):556-558. doi: 10.1001/jamaoto.2023.0704.
15. Moise A, Centomo-Bozzo A, Orishchak O, Alnoury MK, Daniel SJ. Can ChatGPT guide parents on tympanostomy tube insertion? *Children (Basel)*. 2023 Sep 30;10(10):1634. doi: 10.3390/children10101634.
16. Campbell DJ, Estephan LE, Mastrolonardo EV, Amin DR, Huntley CT, Boon MS. Evaluating ChatGPT responses on obstructive sleep apnea for patient education. *J Clin Sleep Med*. 2023 Dec 1;19(12):1989-1995. doi: 10.5664/jcsm.10728.
17. Zalzal HG, Abraham A, Cheng J, Shah RK. Can ChatGPT help patients answer their otolaryngology questions? *Laryngoscope Investig Otolaryngol*. 2023 Dec 9;9(1):e1193. doi: 10.1002/liv.1193.
18. Soto-Galindo GA, Capelleras M, Cruellas M, Apaydin F. Effectiveness of ChatGPT in identifying and accurately guiding patients in rhinoplasty complications. *Facial Plast Surg*. 2023 Dec 27. doi: 10.1055/a-2218-6984.
19. Langlie J, Kamrava B, Pasick LJ, Mei C, Hoffer ME. Artificial intelligence and ChatGPT: an otolaryngology patient's ally or foe? *Am J Otolaryngol*. 2024 May-Jun;45(3):104220. doi: 10.1016/j.amjoto.2024.104220.
20. Mitchell RB, Archer SM, Ishman SL, Rosenfeld RM, Coles S, Finestone SA. et al. Clinical practice guideline: tonsillectomy in children (update)- executive summary. *Otolaryngol Head Neck Surg*. 2019 Feb;160(2):187-205. doi: 10.1177/0194599818807917.
21. Taloni A, Borselli M, Scarsi V, Rossi C, Coco G, Scorcia V. et al. Comparative performance of humans versus GPT-4.0 and GPT-3.5 in the self-assessment program of American Academy of Ophthalmology. *Sci Rep*. 2023 Oct 29;13(1):18562. doi: 10.1038/s41598-023-45837-2.