

EDITORIAL

Will ChatGPT replace human medical writers?

Raluca Enache¹, MD, PhD

ENT Sarafoleanu Medical Clinic, Bucharest, Romania



ChatGPT, which stands for Chat Generative Pre-Trained Transformer, developed by OpenAI and released in November 2022, is an artificial intelligence (AI) software which uses the natural language processing (NLP) technology to provide human-like conversational dialogue in various written content. The software works with algorithms programmed to understand and give appropriate responses to specific tasks or questions (e.g., write a text on a given subject, information about a certain topic, compose a message/ an e-mail/ a note with a specific content/to a specific person).

From its availability, AI was used by students for their homework or projects¹. To investigate the accuracy and integrity of ChatGPT for this purpose, Gilson et al.² evaluated the performance of this chatbot on questions included in the United States Medical Licensing Examination Step and Step 2 exams (clinical, biomedical, epidemiological, social-behavioural science). Analysing the data, the authors concluded that ChatGPT has the capacity to provide logical answers even to the incorrect answered questions, its performance being at the level of a third-year medical student. In other fields, AI impressed by generating poems or short stories written in the style of famous writers.

In healthcare, artificial intelligence was tried for the diagnosis and treatment of the diseases. In the '70s, at Stanford was developed MYCIN for the diagnosis of blood-borne bacterial infections^{3,4}, a programme which was poorly integrated with the work of clinicians and medical recording systems. Later, IBM's Watson was developed as a combination of a machine learning and NLP capabilities for cancer diagnosis and treatment. Unfortunately, the developers found difficult to teach Watson how to differentiate and how to address to particular types of cancer^{4,5}.

Artificial intelligence is already used by the medical staff

and the patients in administrative processes, diagnosis or certain information about their pathology/diseases. At the same time, there are firms and startups focusing on the development of prediction algorithms or learning models through which the healthcare system can benefit from AI.

The use of artificial intelligence, ChatGPT respectively, in scientific medical writing is still under debate, many ethical concerns being raised. From the very beginning, when registering on <https://chat.openai.com>, we are informed that "the system may occasionally generate incorrect or misleading information and produce offensive or biased content" (Figure 1) and which are its capabilities and limitations.

Cureus Journal of Medical Science published, in May 2023, an Open Access Technical Report presenting the use of ChatGPT for a case report writing. Buholayka et al.⁶ tasked ChatGPT to write in five different prompts a case report based on a draft written by oral and maxillofacial radiologists. The authors concluded that the reports generated by the software included many inconsistencies in the medical information, omissions in technical data, incorrect diagnosis and fabricated references.

In January 2023, Else reported that only 68% of the abstracts generated by ChatGPT were identified by investigators among all the scientific papers written by human researchers and ChatGPT⁷. In the same study, the AI-output detector spotted only 66% of the AI abstracts, while the plagiarism software missed them all. Ruksakulpiwat et al.⁸ showed in their review (2013) that "ChatGPT has the potential to revolutionize medical research" by identifying the use of this chatbot in drug development (33.33%), literature review writing (33.33%), medical report improvement, improving research conduct, data analysis, personalized medicine and provision of medical information (16.67% each). There are also other authors who sustain the idea that this chatbot can generate formal research articles⁹.

Corresponding author: Raluca Enache, MD, PhD, ENT Sarafoleanu Medical Clinic, 1 Lt. Av. Iuliu Tetrat Street, District 1, 011911, Bucharest, Romania

ORCID: <https://orcid.org/0000-0002-2841-6265>

e-mail: r.enache@rinologie.ro

Received for publication: June 11, 2023 / **Accepted:** June 21, 2023

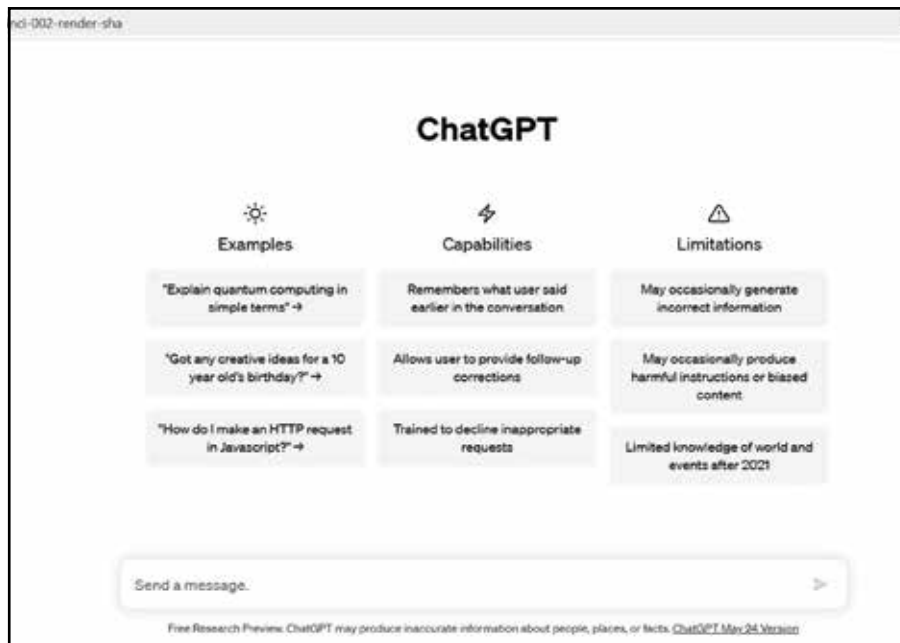


Figure 1. ChatGPT notification note from <https://chat.openai.com/>.

I tested the ChatGPT in a case report writing. I used a draft of a real case report (Figure 2) and the task was generated in three different chats (Figure 3, Figure 4, Figure 5). I identified the following: 1). In all three generated case reports, important data was omitted (details from the MRI report for example); 2). In Chat 2, the generated case report missed information also

about the surgical procedure and included false details, such as “the patient was subsequently referred to a foreign ENT clinic”; 3). The case report in Chat 3 was the shortest, combining less information when compared to the previous two.

ChatGPT can generate data about a certain general subject, but it fails to present a full, detailed, accurate report of a given



Figure 2. The original draft of the case report.

Case Report:

A 35-year-old female presented to the ENT specialist with complaints of right hypoacusis, fullness sensation, and pulsatile tinnitus in the right ear, which had been present for approximately one year. Upon clinical examination, a reddish-blue, pulsatile mass was observed behind the intact right tympanic membrane, without any clinical manifestations of associated facial nerve palsy. Pure-tone audiometry revealed mild conductive hearing loss in the right ear. Cervical and cerebral magnetic resonance imaging (MRI), both native and contrast-enhanced, demonstrated an extra-axial mass in the right temporal bone, centered on the right jugular foramen, with extension into the middle ear cavity and invasion of the sigmoid sinus.

Angiography performed through a right femoral artery approach confirmed the presence of a voluminous jugulotympanic tumoral mass with intradural component on the posterior part of the right petrous pyramid. The tumor displayed significant vascularity, receiving arterial afferents from branches of the external carotid artery, including the occipital, posterior auricular, and internal carotid arteries. Preoperative embolization was performed using polyvinyl alcohol particles (PVA) to occlude the vascular supply, with the exception of a small portion vascularized by terminal branches of the right antero-inferior cerebral artery, due to the risk of ischemic lesions.

The patient was subsequently treated surgically in another ENT clinic abroad. The surgical intervention involved the removal of the extradural part of the tumor and the portion inside the internal auditory canal. During the same procedure, reconstruction of the right facial nerve was performed through tarsorrhaphy. The surgical cavity was obliterated with abdominal fat.

Postoperatively, the patient developed right facial nerve palsy, right ear hearing loss, swallowing disorders (particularly for liquids), and dysphonia. To address the facial nerve palsy, the patient underwent therapy at a recovery clinic specifically targeting facial nerve rehabilitation.

Figure 3. Chat 1 – first case report generated by ChatGPT.

Case Presentation:

A 35-year-old female presented to our ENT clinic with complaints of right hypoacusis, fullness sensation, and pulsatile tinnitus in the right ear. These symptoms had been present for approximately one year. Upon clinical examination, a reddish-blue, pulsatile mass was observed behind the intact right tympanic membrane, without any signs of associated facial nerve palsy. Pure-tone audiometry revealed mild conductive hearing loss in the right ear. Cervical and cerebral magnetic resonance imaging (MRI) scans with contrast demonstrated an extra-axial mass centered on the right jugular foramen, extending into the temporal bone and partially occupying the internal auditory canal. The tumor displayed intense vascularity and received arterial supply from branches of the external and internal carotid arteries.

Treatment and Management:

The patient underwent angiography, which identified the vascular supply of the tumor and facilitated preoperative embolization using polyvinyl alcohol particles. However, due to the risk of ischemic complications, a small portion of the tumor remained vascularized by terminal branches of the right antero-inferior cerebral artery. The patient was subsequently referred to a foreign ENT clinic, where surgical intervention was performed. The procedure involved removal of the extradural component of the tumor and resection of the intracanalicular portion. Additionally, reconstruction of the right facial nerve using tarsorrhaphy was performed, and the surgical cavity was filled with abdominal fat.

Outcome and Follow-up:

Postoperatively, the patient developed right facial nerve palsy, right ear cophosis, swallowing disorders (particularly for liquids), and dysphonia. The patient received rehabilitative therapy for the right facial nerve in a recovery clinic. Long-term follow-up is ongoing to assess the functional outcomes and to monitor for tumor recurrence.

Figure 4. Chat 2 – second case report generated by ChatGPT.

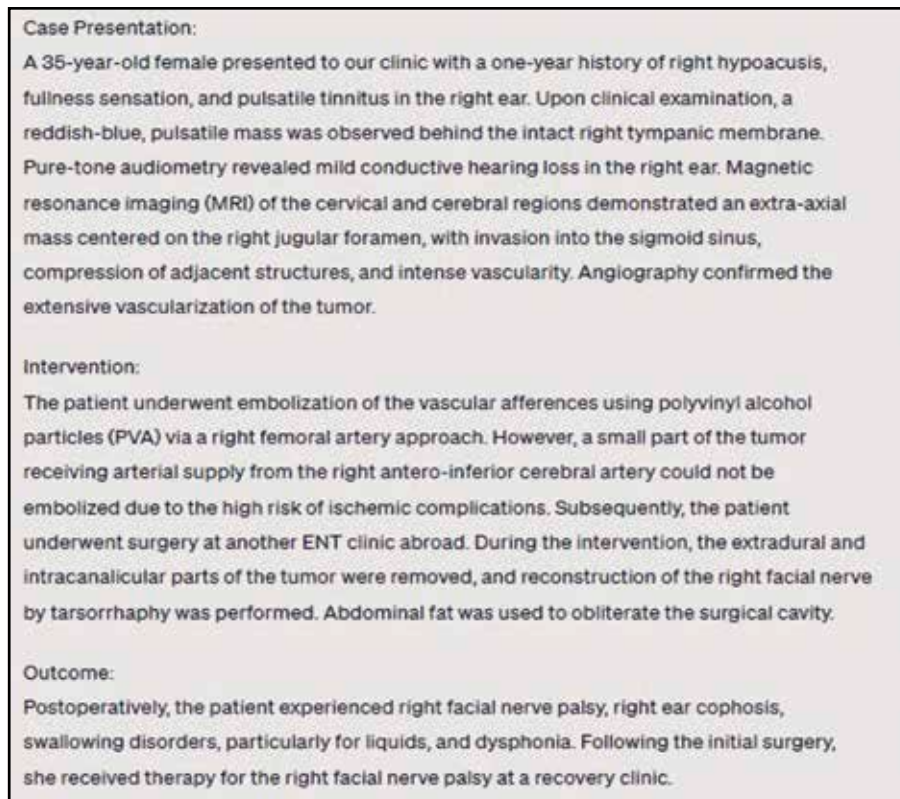


Figure 5. Chat 3 – third case report generated by ChatGPT.

medical case. Another limitation is the absence of references, and the ability to cheat plagiarism checking. All authors who evaluated ChatGPT in medical writing raised the same problems – ethical, originality, scientific integrity issues, biases – and concluded, that in any scientific writing in which ChatGPT is used, the human factor must carefully check, filter and review the correctness and integrity of the medical information provided.

Personally, I do not think that ChatGPT will be able to replace the human factor in writing scientific articles in medicine, in the near future. The chatbot will really help through the speed of information filtering, by exemplifying an editorial model, but in the end the researcher will be the one who will give the final form (scientific content, correctness of data, authenticity), personal experience having an essential role in the way of presentation of the cases, in the interpretation of the results obtained from the research carried out.

Financial disclosure: None.

Conflict of interest: The author has no conflict of interest to declare.

REFERENCES

1. Graham F. Daily briefing: will chatGPT kill the essay assignment? *Nature* 12, 2022.
2. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. Medical Licensing examination? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ.* 2023;e45312. DOI: 10.2196/45312.
3. Bush J. How AI is taking the scut work out of health care . *Harvard Business Review* 2018. <https://hbr.org/2018/03/how-ai-is-taking-the-scut-work-out-of-health-care>.
4. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthcare Journal.* 2019;6(2):94-8.
5. Buchanan BG , Shortliffe EH. Rule-based expert systems: The MYCIN e xperiments of the Stanford heuristic programming project. Reading: Addison Wesley , 1984 .
6. Buholayka M, Zouabi R, Tadinada A. The readiness of ChatGPT to write scientific case reports independently: A comparative evaluation between human and artificial intelligence. *Cureus.* 2023;15(5):e39386. DOI: 10.7759/cureus.39386.
7. Else, H. (2023). Abstracts written by ChatGPT fool scientists. *Nature* 613, 423. DOI: 10.1038/d41586-023-00056-7.
8. Ruksakulpiwat S, Kumar A, Ajibade A. Using ChatGPT in medical research: Current status and future directions. *Journal of Multidisciplinary Health Care.* 2023;16:1513-20.
9. Gordijn B, Have HT. ChatGPT: evolution or revolution? *Med Health Care and Philos.* 2023;26:1-2. DOI: 10.1007/s11019-023-10136-0.

