


EDITORIAL



How could ChatGPT impact my practice as an intensivist? An overview of potential applications, risks and limitations

Matthieu Komorowski^{1,5*} , Maria del Pilar Arias López^{2,3}  and Anthony C. Chang⁴

© 2023 Springer-Verlag GmbH Germany, part of Springer Nature

What is ChatGPT and how does it work?

Chat Generative Pre-trained Transformer (ChatGPT) is a very large natural language processing model that uses deep learning algorithms trained on vast amounts of internet data to generate human-like responses to user prompts. It was developed by a company called OpenAI and released in November 2022 [1].

Transformers are an architecture used primarily in the field of natural language processing, that aims to solve sequence-to-sequence tasks (e.g. language translation, which amounts to predicting a sequence of words in a target language from a sequence of words in a source language) while handling long-range dependencies with ease. They provide superior answers while requiring less training time than previous approaches such as recurrent neural networks [2]. In short, ChatGPT examines the words provided as input by the users, and predicts what words will follow while generating a coherent response in a human-like style. It has been fine-tuned for a myriad of language tasks including text generation and completion, translation, sentiment analysis, document summarisation, question and answering, and even generation and explanation of programming code [1].

So far, ChatGPT passed the Wharton MBA examination, law school examinations, United States Medical Licensing Examination (USMLEs), and more [3, 4]. Its role has been explored in different specialties including intensive care medicine, with variable results [5,

6]. How much of ChatGPT's "textbook" knowledge remains applicable to "real world" patients remains to be demonstrated.

Potential applications of ChatGPT in intensive care medicine

ChatGPT and large language models (LLMs) could potentially impact the clinical practice of intensive care medicine in significant ways (Table 1).

Supporting medical decisions

In the intensive care unit (ICU) setting, where there is a high level of uncertainty and where fast and accurate decisions are necessary, ChatGPT can be used as a glorified search engine, quickly providing information on medical conditions, treatments, and drug interactions, offering evidence-based recommendations, but also analysing (actual) patient data and predicting outcomes, all of which could potentially improve patient outcomes [7].

Handling of medical notes

As an artificial intelligence (AI) specialised in language problems, ChatGPT could assist clinicians in a number of tasks related to handling medical data and clinical notes. For example, it can be used to organise clinical notes and write accurate discharge summaries from unstructured input data [8].

Medical education

LLMs can also be used as educational tools, whereby medical students or clinical trainees interact with the AI to ask and/or answer clinical questions. Gunther Eysenbach provided a clear demonstration of such application in diabetes [9].

*Correspondence: m.komorowski14@imperial.ac.uk

¹ Division of Anaesthetics, Pain Medicine, and Intensive Care, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London SW7 2AZ, UK

Full author information is available at the end of the article

Table 1 Potential applications, limitations and risks of using large language models such as ChatGPT in medicine and the intensive care unit. For each application, we provide examples of how ChatGPT could be queried

Potential applications	Example of queries in ChatGPT	Limitations and risks
Providing fast and accurate information, suggesting differential diagnoses, diagnostic testing, management options, hence supporting clinical decision-making	[Providing a given history, physical exam, laboratory and radiology results:] "What are the most likely differential diagnoses, ranked by likelihood, along with explanations?" [After confirming the working diagnosis:] "Provide a suggested management plan"	Can "hallucinate", creating confident and completely wrong answers
Handling of medical data	[Providing unstructured medical notes for an entire ICU stay:] "Generate a structured discharge letter for this patient"	Data fabrication, plagiarism, scientific integrity
Enhancing communication between doctors and patients/relatives	[Providing a standard discharge letter:] "Explain this ICU discharge letter in lay terms"	Can provide biased answers
Medical education	[Providing a scenario for medical students:] "Discuss the differential diagnoses and the investigation plan to rule each of them in and out, with explanations"	Data privacy and security
Scientific writing	[Providing a journal manuscript:] "Discuss the strengths and limitations of this research article" [Providing the abstract of a scientific article in English:] "Translate this abstract into Spanish"	Automation of healthcare professions

There is much more work that will need to be done for a biomedical LLM to provide accurate and up-to-date answers to very specific and technical questions, but it has the potential to be a robust educational tool that is equal to or better than learning from textbooks or scientific literature [10].

Enhancing communication

ChatGPT can be used as a communication tool between healthcare providers, patients, and their families, for example to explain or translate technical information. The ability of ChatGPT to summarise lay and specialised information is often remarkable, so it represents an opportunity for technical communication among clinicians, as well as plain language for communication with patients and their families [11].

Scientific writing

ChatGPT can assist in the writing process of scientific papers, identify research questions, provide an overview of the current state of the field, and assist with tasks such as formatting and language and content review [12].

Using ChatGPT, free translations can reduce costs and facilitate access to publishing for the authors from non-English speaking countries, hence helping with democratisation and diffusion of knowledge [13].

Limitations and risks of large language models

However, it is important to note that ChatGPT is a machine learning model, and as such, it has limitations,

and can be used in "wrong" ways, knowingly or unknowingly (Table 1).

Hallucinations

Bluntly put, ChatGPT (at the time of writing) cannot be (fully) trusted, since its output can be flat-out incorrect. The so-called "hallucination" phenomenon refers to the ability of ChatGPT to confidently produce answers that look believable but may be incorrect or nonsensical. For example, it may invent publications, references or Pubmed IDs, or write case reports about non-existing diseases [14]. It cannot replace the expertise of trained healthcare professionals, and its recommendations should always be considered in the context of the patient's unique circumstances.

Misuses

Given their ability to generate realistic looking data, there are a number of possible misuses of LLMs, including fabricating research data or results to meet funding or publication requirements, generating fake news or misinformation, and plagiarising [7, 10]. This has serious implications for the integrity of the scientific record, given the risk of introducing not only errors but also plagiarised content into publications. This could result in future research or health policy decisions being made on the basis of false information [10].

Perpetuating biases

These models train on existing medical research that could carry biases, causing the AI program to risk

perpetuating these same biases [10]. Since most scientific publications are from high-income countries, ChatGPT's output will be based on evidence from these countries, and its applicability to low- and middle-income settings must be questioned.

Data privacy and security

A significant concern related to privacy pertains to the data shared with ChatGPT through user prompts. As we request ChatGPT to answer queries, we may unintentionally disclose sensitive information and expose it to the public. Since every user prompt is incorporated into ChatGPT's database, they may be utilised to enhance the tool's training and conceivably be incorporated in replies to other users' prompts [10].

Automation of jobs and wider societal risks

Finally, advanced LLMs are a source of concerns and wider societal risks, which have been expressed by machine learning experts and key opinion leaders, calling for a halt in training these models [15].

While it is unlikely that LLMs will ever replace physicians, if they can pass medical exams, suggest differential diagnoses, investigation and treatment plans, summarise information and communicate effectively with patients and their relatives, what is left for human doctors to do? As stated in the Future of Life Institute letter, "Should we automate away all the jobs, including the fulfilling ones?" [15].

Take-home messages

In the fast-paced and high stress environment of the ICU, LLMs like ChatGPT should be considered as potential valuable tools to support medical and non-medical staff in daily activities, education and research. Importantly, users must be aware of the limitations and potential misuses of these tools, including their capacity to generate confident and completely wrong answers or perpetuate existing biases and inequalities. However, this is a fast evolving field, and (rapid) progress on LLMs performance, safety and uncertainty awareness is to be expected.

Author details

¹ Division of Anaesthetics, Pain Medicine, and Intensive Care, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London SW7 2AZ, UK. ² Hospital de Niños Ricardo Gutierrez, Intermediate Care Unit, Gallo 1330, C1425EFD Buenos Aires, Argentina. ³ Argentina Society of Intensive Care, SATI-Q Paediatric Program. Av. Cnel. Niceto Vega 4617, C1414BEA Buenos Aires, Argentina. ⁴ Children's Hospital of Orange County Sharon Disney Lund Medical Intelligence and Innovation Institute, 1201 W. La Veta Ave, Orange, CA 92868-3874, USA. ⁵ Intensive Care Unit, Charing Cross Hospital, Fulham Palace Road, London W6 8RF, UK.

Acknowledgements

We are grateful to Alexandre Joosten, Louis Kreitmann, Dominic Marshall, and Justin Salciccioli for their critical input in this manuscript. ChatGPT was used at the early drafting phase of this article to suggest ideas for the content.

Data availability

Not applicable.

Declarations

Conflicts of interest

None declared.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 12 April 2023 Accepted: 5 May 2023

Published: 31 May 2023

References

1. OpenAI. Introducing ChatGPT. (2022) <https://openai.com/blog/chatgpt#OpenAI>. Accessed 7 April 2023
2. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A, Kaiser L, Polosukhin I (2017). Attention is all you need. *Advances in neural information processing systems*. 30. <https://doi.org/10.48550/arXiv.1706.03762>. <https://arxiv.org/abs/1706.03762>. Accessed 8 April 2023
3. Gilson A, Safranek C, Huang T, Socrates V, Chi L, Taylor RA, Chartash D (2023) How does ChatGPT perform on the United States medical licensing examination? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ* 9:e45312. <https://doi.org/10.2196/45312>
4. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, Madriaga M, Aggabao R, Diaz-Candido G, Maningo J, Tseng V (2023) Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health* 2(2):e0000198. <https://doi.org/10.1371/journal.pdig.0000198>
5. Sabry Abdel-Messih M, Kamel Boulos MN (2023) ChatGPT in clinical toxicology. *JMIR Med Educ* 9:e46876. <https://doi.org/10.2196/46876>
6. Alberts IL, Mercolli L, Pyka T, Prenosil G, Shi K, Rominger A, Afshar-Oromieh A (2023) Large language models (LLM) and ChatGPT: what will the impact on nuclear medicine be? *Eur J Nucl Med Mol Imag* 50(6):1–4. <https://doi.org/10.1007/s00259-023-06172-w>
7. Cascella M, Montomoli J, Bellini V, Bignami E (2023) Evaluating the feasibility of ChatGPT in healthcare: an analysis of multiple clinical and research scenarios. *J Med Syst* 47(1):33. <https://doi.org/10.1007/s10916-023-01925-4>
8. Patel SB, Lam K (2023) ChatGPT: the future of discharge summaries? *Lancet Digit Health* 5(3):e107–e108. [https://doi.org/10.1016/S2589-7500\(23\)00021-3](https://doi.org/10.1016/S2589-7500(23)00021-3)
9. -Eysenbach G (2023) The Role of ChatGPT, Generative Language Models, and Artificial Intelligence in Medical Education: A Conversation With ChatGPT and a Call for Papers *JMIR Med Educ* 2023;9:e46885-<https://mededu.jmir.org/2023/1/e46885>. DOI: <https://doi.org/10.2196/46885>. Accessed 7 April 2023
10. Sallam M (2023) ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare (Basel)* 11(6):887. <https://doi.org/10.3390/healthcare11060887>
11. Abbey A (2023) Artificially intelligent reflection? Smoke and mirrors and a tale of two perspectives. *Intensive Care Med*. <https://doi.org/10.1007/s00134-023-07008-9>
12. Salvagno M, Taccone FS, Gerli AG (2023) Can artificial intelligence help for scientific writing? *Crit Care* 27(1):99

13. Kim SG (2023) Using ChatGPT for language editing in scientific articles. *Maxillofac Plast Reconstr Surg* 45(1):13
14. Alkaissi H, McFarlane SI (2023) Artificial Hallucinations in ChatGPT: implications in scientific writing. *Cureus* 15(2):e35179. <https://doi.org/10.7759/cureus.35179>
15. The Future of Life Institute. Pause Giant AI Experiments: An Open Letter (2023). <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> Accessed 9 April 2023