



# The computer will see you now: ChatGPT and artificial intelligence large language models for health information in urology—an invited perspective

Joseph Gabriel<sup>1</sup>^, Lidia Shafik<sup>2</sup>, Elizabeth Vincent<sup>1</sup>, Jirayr Ajzajian<sup>1</sup>, Ammar Alanbuki<sup>1</sup>, Timothy R. G. Lerner<sup>1</sup>

<sup>1</sup>Department of Urology, University Hospitals Sussex NHS Foundation Trust, Princess Royal Hospital, Haywards Heath, UK; <sup>2</sup>Department of Medicine, Frimley Health NHS Foundation Trust, Frimley, UK

*Correspondence to:* Joseph Gabriel, BMBS, MSc, FRCS(Urol). Speciality Trainee (ST6) in Urology, Department of Urology, University Hospitals Sussex NHS Foundation Trust, Princess Royal Hospital, Lewes Road, Haywards Heath, East Sussex, RH16 4EX, UK. Email: joseph.gabriel@nhs.net.

*Comment on:* Davis R, Eppler M, Ayo-Ajibola O, *et al.* Evaluating the Effectiveness of Artificial Intelligence-powered Large Language Models Application in Disseminating Appropriate and Readable Health Information in Urology. *J Urol* 2023;210:688-94.

**Keywords:** ChatGPT; artificial intelligence (AI); large language model (LLM); patient information; health information

Submitted Sep 27, 2023. Accepted for publication Nov 16, 2023. Published online Dec 07, 2023.

doi: 10.21037/tau-23-491

**View this article at:** <https://dx.doi.org/10.21037/tau-23-491>

ChatGPT has taken the world by storm since its launch in November 2022, breaking records as it reached 100 million active users in January 2023, making it the fastest-growing consumer application in history (1). The astronomical rise in the use of such artificial intelligence (AI) large language model (LLM) platforms in healthcare, particularly as potential patient education tools has made this an exciting and emerging area of research.

Davis *et al.* have written a robust and highly topical paper on the utility of ChatGPT 3.5 as a patient information tool in urology (2). Though not the first paper examining its use in urology (3-7) (contrary to the authors attestations), this is a well-executed study in which 18 potential patient questions, encompassing benign, oncological and emergency conditions (e.g., “My doctor told me I have prostate cancer, and I have to undergo prostatectomy. What does this mean?” and “Recently, my urine is cloudy and smells bad. What does this mean?”) were asked of ChatGPT. Three independent native English-speaking board-certified urologists then scored the answers it provided in three domains: accuracy, comprehensiveness and clarity using a 5-point Likert scale. For an answer to be

deemed “appropriate”, it required a score of  $\geq 4$  in each of the three domains.

Quantitative analysis showed that 14/18 (77.8%) of the platform’s responses were deemed “appropriate,” with the responses significantly ( $P < 0.05$ ) scoring higher on clarity relative to comprehensiveness or accuracy.

In addition to the appropriateness of the response, the authors assessed their readability using the Flesch Reading Ease and Flesh-Kincaid Reading Grade Level scores (8,9), two validated and widely-used readability criteria, used to measure how easy or difficult a text is to read. The readability was at a more difficult level than the American Medical Association recommendation for US patient information (10). As the authors highlight, it is imperative that any patient information is understandable and well-written in our highly specialised and technical field. It seems as though ChatGPT still has some ways to go before it produces medical information of a suitable level of readability. Work has also been done recently comparing ChatGPT to validated tools and human raters in assessing the readability of online medical information and have found it to be limited in its capabilities (11).

^ ORCID: 0000-0002-1238-9215.

Limitations of this study include the limited selection of questions asked of the chatbot, the fact that only the first reply to each question was taken as being representative of ChatGPT's response (the AI platform generates a unique response each time); perhaps the same question should be asked in multiple iterations in order to obtain a more representative sample of the platform's response in future studies. As the authors have also mentioned, the development of a new set of validated tools to be used for the specific purpose of such platforms is yet to be developed, and perhaps should be sought as a standardised method to compare this study's results against other studies. The authors must be lauded for their ongoing investigations that they allude to in pursuing further research in this sphere, particularly in the comparison of ChatGPT to other competitor platforms, such as Bard (Google) and Bing (Microsoft). We note that another group has very recently examined other AI chatbots including ChatGPT, Perplexity, Chat Sonic and Microsoft Bing (12).

In conclusion, this is a welcome article in the burgeoning area that is LLMs, which are almost certainly in use by the general public to enquire about their urological health, and whose role can only be anticipated to grow. It is plain that the pitfalls of undiscerning use of LLMs can include information of a variable quality and utility to patients. More recent publications in the field of AI LLMs and urology include their potential utility in administrative tasks from as diverse areas as facilitating note taking (13) to appraising letters of recommendations for residencies (14). Various ethical, legal and social implications have certainly arisen from the spread of such technologies that need further exploration (15). As we mentioned in our own paper, it is imperative that urologists remain vigilant to these fallible tools when consulting patients in clinic, and that we take key roles as active stakeholders in the development and regulation of such tools when it comes to our patients' wellbeing.

## Acknowledgments

*Funding:* None.

## Footnote

*Provenance and Peer Review:* This article was commissioned by the editorial office, *Translational Andrology and Urology*. The article has undergone external peer review.

*Peer Review File:* Available at <https://tau.amegroups.com/article/view/10.21037/tau-23-491/prf>

*Conflicts of Interest:* All authors have completed the ICMJE uniform disclosure form (available at <https://tau.amegroups.com/article/view/10.21037/tau-23-491/coif>). The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

*Open Access Statement:* This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

## References

1. Hu K. ChatGPT sets record for fastest-growing user base - analyst note. Reuters [Internet]. 2023 Feb 2 [cited 2023 Apr 4]; Available online: <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/>
2. Davis R, Eppler M, Ayo-Ajibola O, et al. Evaluating the Effectiveness of Artificial Intelligence-powered Large Language Models Application in Disseminating Appropriate and Readable Health Information in Urology. *J Urol* 2023;210:688-94.
3. Gabriel J, Shafik L, Alanbuki A, et al. The utility of the ChatGPT artificial intelligence tool for patient education and enquiry in robotic radical prostatectomy. *Int Urol Nephrol* 2023;55:2717-32.
4. Coskun B, Ocakoglu G, Yetemen M, et al. Can ChatGPT, an Artificial Intelligence Language Model, Provide Accurate and High-quality Patient Information on Prostate Cancer? *Urology* 2023;180:35-58.
5. Cocci A, Pezzoli M, Lo Re M, et al. Quality of information and appropriateness of ChatGPT outputs for urology patients. *Prostate Cancer Prostatic Dis* 2023. [Epub ahead of print]. doi: 10.1038/s41391-023-00705-y.

6. Zhou Z, Wang X, Li X, et al. Is ChatGPT an Evidence-based Doctor? *Eur Urol* 2023;84:355-6.
7. Zhu L, Mou W, Chen R. Can the ChatGPT and other large language models with internet-connected database solve the questions and concerns of patient with prostate cancer and help democratize medical knowledge? *J Transl Med* 2023;21:269.
8. Kincaid J, Fishburne R, Rogers R, Chissom B. Derivation Of New Readability Formulas (Automated Readability Index, Fog Count And Flesch Reading Ease Formula) For Navy Enlisted Personnel. Institute for Simulation and Training [Internet]. 1975 Jan 1. Available online: <https://stars.library.ucf.edu/istlibrary/56>
9. FLESCH R. A new readability yardstick. *J Appl Psychol* 1948;32:221-33.
10. Weiss BD. Health Literacy: A Manual for Clinicians. American Medical Association. 2003. Available online: <http://lib.ncfh.org/pdfs/6617.pdf>
11. Golan R, Ripps SJ, Reddy R, et al. ChatGPT's Ability to Assess Quality and Readability of Online Medical Information: Evidence From a Cross-Sectional Study. *Cureus* 2023;15:e42214.
12. Musheyev D, Pan A, Loeb S, et al. How Well Do Artificial Intelligence Chatbots Respond to the Top Search Queries About Urological Malignancies? *Eur Urol* 2023. [Epub ahead of print]. doi:10.1016/j.eururo.2023.07.004
13. Talyshinskii A, Naik N, Hameed BMZ, et al. Expanding horizons and navigating challenges for enhanced clinical workflows: ChatGPT in urology. *Front Surg* 2023;10:1257191.
14. Barrett A, Hekman L, Ellis JL, et al. Utilization of ChatGPT for Appraising Letters of Recommendation in Urology Residency Applications: Ready for Prime Time? *J Urol* 2023;210:833-4.
15. Adhikari K, Naik N, Hameed BZ, et al. Exploring the Ethical, Legal, and Social Implications of ChatGPT in Urology. *Curr Urol Rep* 2023. [Epub ahead of print]. doi: 10.1007/s11934-023-01185-2.

**Cite this article as:** Gabriel J, Shafik L, Vincent E, Ajzajian J, Alanbuki A, Larner TRG. The computer will see you now: ChatGPT and artificial intelligence large language models for health information in urology—an invited perspective. *Transl Androl Urol* 2023;12(12):1772-1774. doi: 10.21037/tau-23-491