

BRIEF COMMUNICATION

OPEN



Clinical

ChatGPT can help guide and empower patients after prostate cancer diagnosis

Harry Collin ^{1,2}✉, Kandice Keogh ^{1,2}, Marnique Basto ¹, Stacy Loeb ³ and Matthew J. Roberts ^{1,4}

© Crown 2024

BACKGROUND/OBJECTIVES: Patients often face uncertainty about what they should know after prostate cancer diagnosis. Web-based information is common but is at risk of being of poor quality or readability.

SUBJECTS/METHODS: We used ChatGPT, a freely available Artificial intelligence (AI) platform, to generate enquiries about prostate cancer that a newly diagnosed patient might ask and compared to Google search trends. Then, we evaluated ChatGPT responses to these questions for clinical appropriateness and quality using standardised tools.

RESULTS: ChatGPT generates broad and representative questions, and provides understandable, clinically sound advice.

CONCLUSIONS: AI can guide and empower patients after prostate cancer diagnosis through education. However, the limitations of the ChatGPT language-model must not be ignored and require further evaluation and optimisation in the healthcare field.

Prostate Cancer and Prostatic Diseases (2025) 28:513–515; <https://doi.org/10.1038/s41391-024-00864-6>

Patients often self-educate on their health problems, particularly if anxiety impacts information comprehension [1]. Web-based information (e.g. YouTube, TikTok, Instagram) can be of poor quality [2, 3] and poorly actionable (difficult for patients to identify what they can do based on the information presented).

ChatGPT 3.5 is an artificial intelligence (AI), large-language modelling chatbot that generates coherent responses to questions using language patterns from data published prior to September 2021 [4]. ChatGPT responses have been shown to be weakly distinguishable from those of healthcare providers, generating trust in AI chatbots by patients for health information [5]. Response generation does not use direct knowledge so is at risk of misinformation [6]. Thus, this study aimed to evaluate ChatGPT responses and health information quality for prostate cancer patient questions.

ChatGPT 3.5 (<https://chat.openai.com/>) was asked to generate the 10 most asked questions about prostate cancer. These questions were then compared against the top 25 worldwide Google searches during the 2022 calendar year using Google Trends (<https://trends.google.com/trends/>; Supplementary Tables 1 and 2). Each question was input into ChatGPT and a response generated (Supplementary Table 3).

A board-certified urologist graded each response as ‘appropriate’ or ‘inappropriate’ advice. The Patient Education Materials Assessment Tool (PEMAT) [7] adjudged content understandability (ability of patients to process and explain key messages) and actionability (ability of patients to identify what they can do based

on the information presented) measured as a percentage score of 0–100% (high quality $\geq 75\%$). Authorship, currentness and tone were appraised by the Quality Evaluation Scoring Tool (QUEST) [8] scored out of 28. Readability was reported as estimated reading grade level using the Flesch-Kincaid test.

Most (8/10) questions generated by ChatGPT were in the top 25 Google searches related to prostate cancer (Supplementary Table 2). ChatGPT answered these questions with clinically appropriate responses of variable length [mean 206 words, standard deviation (SD) 50 words]. Readability was at a college level.

Seven ChatGPT responses were deemed high quality patient information (Table 1), with a mean PEMAT score of 91.7% for understandability (SD 6.7%). All responses had clear purpose and did not contain distracting information. Most used everyday language with medical terms and numbers presented in an understandable manner, while one response used medical terms without explanation (referring to *BRCA* gene mutations). Six responses employed bullet points, while three responses lacked organisational features. Three responses would have benefited from visual aids. Responses scored a mean PEMAT score of 76% for actionability (SD 12.7%). All responses provided at least one action the user could take. Three responses did not explain the action in explicit steps and only one response addressed the user directly.

QUEST scores for information quality were poor (mean score 11.3 [40.4%], SD 2 [7.2%]), attributable to lack of information

¹Department of Urology, Royal Brisbane & Women's Hospital, Brisbane, QLD, Australia. ²Faculty of Medicine, The University of Queensland, Brisbane, QLD, Australia. ³Departments of Urology and Population Health, New York University Langone Health, New York, NY, USA. ⁴University of Queensland Centre for Clinical Research, Brisbane, QLD, Australia. ✉email: [hjcollin@gmail.com](mailto:hjacollin@gmail.com)

Table 1. Content quality scores (PEMAT and QUEST) for ChatGPT-generated responses.

Question	PEMAT		QUEST (out of 28)	QUEST (% of total)
	Understandability	Actionability		
1	90%	100%	13	46.4%
2	90%	80%	9	32.1%
3	100%	60%	10	35.7%
4	100%	80%	10	35.7%
5	80%	60%	15	53.6%
6	85%	60%	13	46.4%
7	92%	80%	13	46.4%
8	90%	80%	10	35.7%
9	100%	80%	10	35.7%
10	90%	80%	10	35.7%
Mean	91.7%	76.0%	11.30	40.4%
STD	6.7%	12.6%	2.00	7.2%

authorship, reference to expert sources and indication of currentness. All responses employed cautious vocabulary and eight responses referenced the patient-physician relationship.

Overall, ChatGPT-generated questions matched common patient enquiries and suggests that AI can be a companion for patients who are generally unsure what to ask about prostate cancer.

Responses to these questions were consistently clinically appropriate, signifying that ChatGPT has the potential to offer reliable information to patients. Our study reports higher accuracy (100%) when compared to ChatGPT responses to prostate cancer questions adapted from online patient materials (94.2%) [9], suggesting that patient questions generated by ChatGPT itself can be answered accurately. ChatGPT may be a superior online platform for healthcare advice accuracy compared to others including YouTube [2], TikTok [2] and Instagram [3], known for reporting significant misinformation. However, evaluation with QUEST highlighted issues of ChatGPT information transparency concordant with other studies including a lack of authorship, referencing [10] and currentness [2, 3]. Such transparency limitations were offset by cautious vocabulary to indicate inherent medical uncertainty for patients that are less obvious among online resources.

ChatGPT responses were understandable, reflecting the clear purpose and everyday language in responses. Additionally, each response identified one action the reader could take. While ChatGPT has been shown to confer variable understandability (66.7%) and actionability (40%) across uro-oncology domains [10], our study showed better understandability (91.7%) and actionability (76%), likely due to the relatively basic information need from patients. However, college-level readability may not be applicable for all patients. The absence of explicit steps in some responses may affect the patient's ability to make informed decisions about their health. Lack of visual aids hindered understandability, which may direct patients to alternative platforms such as TikTok [2] and Instagram [3] where misinformation and poor actionability is more likely to be encountered.

While this study shows that AI is promising for education and patient information, there are clear limitations worth noting. For instance, only 26% of ChatGPT responses to clinical questions on prostate cancer, based on EAU Guidelines, were accurate [11]. ChatGPT-generated questions may not fully capture the breadth or depth for patients from various backgrounds and states of mind. Moreover, we did not account for progressive information gathering, where patients often ask additional questions through dynamic dialogue. Therefore, professional oversight is recommended for clinical applications to ensure clinical safety, effectiveness and reliability. Second, patient preference for education source and

format is an important consideration for wider implementation. Patients over 60 years of age have found chatbots usable with low cognitive load [12], but accessibility and effective use is unclear. Additionally, version control of AI chatbots warrant consideration, as ChatGPT 4.0 (from ChatGPT 3.5) is more current and customisable but less accessible due to paywall. Currently, we expect ChatGPT 3.5 to be accessed more by patients, however future research should continually compare various AI models for effective, accurate prostate cancer information. Customisation of AI chatbots for specific tasks presents an opportunity for development of AI models tailored specifically for patient needs.

In conclusion, ChatGPT can offer understandable, clinically appropriate initial advice for newly diagnosed prostate cancer patients. ChatGPT-generated questions reflect what patients typically ask, which can be helpful for patients who otherwise don't know what to ask. AI chatbots can augment patient education through access to healthcare advice. As AI continues to evolve, mindful integration, ongoing refinement, and collaboration with healthcare professionals and medical databases will be pivotal to harnessing its full potential to empower patients through effective information delivery.

DATA AVAILABILITY

The data analysed in this study is available from the corresponding author upon request.

REFERENCES

1. Bell RA, Hu X, Orrange SE, Kravitz RL. Lingering questions and doubts: Online information-seeking of support forum members following their medical visits. *Patient Educ Couns*. 2011;85:525–8.
2. Abramson M, Feiertag N, Javidi D, Babar M, Loeb S, Watts K. Accuracy of prostate cancer screening recommendations for high-risk populations on YouTube and TikTok. *BJUJ Compass*. 2023;4:206–13.
3. Xu AJ, Myrie A, Taylor JI, Matulewicz R, Gao T, Pérez-Rosas V, et al. Instagram and prostate cancer: using validated instruments to assess the quality of information on social media. *Prostate Cancer Prostatic Dis*. 2022;25:791–3.
4. OpenAI. GPT-4 Technical Report. arXiv; 2023. Available from: <http://arxiv.org/abs/2303.08774>.
5. Nov O, Singh N, Mann D. Putting ChatGPT's Medical Advice to the (Turing) Test: Survey Study. *JMIR Med Educ*. 2023;9:e46939.
6. Whiles BB, Bird VG, Canales BK, DiBianco JM, Terry RS. Caution! AI Bot Has Entered the Patient Chat: ChatGPT Has Limitations in Providing Accurate Urologic Healthcare Advice. *Urology*. 2023;180:278–84.
7. Shoemaker SJ, Wolf MS, Brach C. Development of the Patient Education Materials Assessment Tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. *Patient Educ Couns*. 2014;96:395–403.

8. Robillard JM, Jun JH, Lai JA, Feng TL. The QUEST for quality online health information: validation of a short quantitative tool. *BMC Med Inform Decis Mak*. 2018;18:87.
9. Caglar U, Yildiz O, Meric A, Ayranci A, Yusuf R, Sarilar O, et al. Evaluating the performance of ChatGPT in answering questions related to benign prostate hyperplasia and prostate cancer. *Minerva Urol Nephrol*. 2023;75:729–33.
10. Musheyev D, Pan A, Loeb S, Kabarriti AE. How Well Do Artificial Intelligence Chatbots Respond to the Top Search Queries About Urological Malignancies? *Eur Urol*. 2024;85:13–6.
11. Lombardo R, Gallo G, Stira J, Turchi B, Santoro G, Riolo S, et al. Quality of information and appropriateness of Open AI outputs for prostate cancer. *Prostate Cancer Prostatic Dis*. 2024;1–3.
12. Wilczewski H, Soni H, Ivanova J, Ong T, Barrera JF, Bunnell BE, et al. Older adults' experience with virtual conversational agents for health data collection. *Front Digit Health*. 2023;5:1125926.

ACKNOWLEDGEMENTS

MJR is supported by a Clinician Research Fellowship from the Metro North Office of Research, Queensland Health and Research Support Package from The University of Queensland.

AUTHOR CONTRIBUTIONS

HC and MJR helped to oversee and direct the project. HC, KK, MB, SL and MJR were responsible for design of the study. HC and KK were responsible for response & data collection and drafting the manuscript. MJR was responsible for reviewing ChatGPT responses as a board-certified urologist. MB, SL and MJR were responsible for drafting and critical revision of the manuscript.

FUNDING

Open Access funding enabled and organized by CAUL and its Member Institutions.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41391-024-00864-6>.

Correspondence and requests for materials should be addressed to Harry Collin.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© Crown 2024