https://doi.org/10.1007/s00345-023-04563-0

ChatGPT and most frequent urological diseases: analysing the quality of information and potential risks for patients

Juliusz Jan Szczesniewski[1,2] · Carlos Tellez Fouz[1] · Alejandra Ramos Alba[3,4] · Francisco Javier Diaz Goizueta[1] · Ana García Tello[1] · Luis Llanes González[1,5]

Abstract

Purpose Artificial intelligence (AI) is a set of systems or combinations of algorithms, which mimic human intelligence. ChatGPT is software with artificial intelligence which was recently developed by OpenAI. One of its potential uses could be to consult the information about pathologies and treatments. Our objective was to assess the quality of the information provided by AI like ChatGPT and establish if it is a secure source of information for patients.

Methods Questions about bladder cancer, prostate cancer, renal cancer, benign prostatic hypertrophy (BPH), and urinary stones were queried through ChatGPT 4.0. Two urologists analysed the responses provided by ChatGPT using DISCERN questionary and a brief instrument for evaluating the quality of informed consent documents.

Results The overall information provided in all pathologies was well-balanced. In each pathology was explained its anatomical location, affected population and a description of the symptoms. It concluded with the established risk factors and possible treatment. All treatment answers had a moderate quality score with DISCERN (3 of 5 points). The answers about surgical options contain the recovery time, type of anaesthesia, and potential complications. After analysing all the responses related to each disease, all pathologies except BPH achieved a DISCERN score of 4.

Conclusions ChatGPT information should be used with caution since the chatbot does not disclose the sources of information and may contain bias even with simple questions related to the basics of urologic diseases.

Keywords ChatGPT · Urology · Patient information · Artificial intelligence · Quality of information

Introduction

Artificial intelligence (AI) is a set of systems or combinations of algorithms, whose purpose is to create machines that mimic human intelligence. Most of those systems can improve as the information they gather increases. ChatGPT is a software Chatbot with an artificial intelligence language model (specifically generative language) released in 2022 by OpenAI, and it is specialized in dialogue. It can generate well-structured and coherent written texts, create consistent answers to users' questions, analyse problems, clarify doubts, generate solutions, reports, summaries, or e-mails [1].

Unfortunately, the information found on internet is not always right. Furthermore, even if the information is correct, it could be misunderstood by the user. When related to health, this risk becomes greater especially if the patient uses it for self-diagnosis [2]. Multiple studies have been conducted to assess information about urological illness on the internet. For most pathologies such as bladder cancer, benign prostatic hypertrophy, or prostate cancer, a risk of biased information has been observed [3–5].

Algorithm-based tools, such as ChatGPT or other similar AI chatbots, can provide multiple benefits in terms of health

✉ Juliusz Jan Szczesniewski
juliusz.szcz@gmail.com

1 Department of Urology, Getafe University Hospital, Carretera Madrid-Toledo km 12,500 Getafe, 28905 Madrid, Spain

2 Faculty of Medicine, University of Salamanca, C. Alfonso X el Sabio, s/n, 37007 Salamanca, Spain

3 DXC Technology, C. de José Echegaray, Las Rozas, 28232 Madrid, Spain

4 Rey Juan Carlos University, P. de los Artilleros 38, 28032 Madrid, Spain

5 Francisco de Vitoria University, Carretera Pozuelo a Majadahonda, Km 1.800, 28223 Madrid, Spain

promotion. They could help the patient make better informed decisions related to their health status [6, 7], and obtain more direct, fast, convenient, and specific information. This new approach could reduce unnecessary consultations and visits. However, the risks still exist, and patients may ask inappropriate treatments or not seek medical attention when needed [8].

Our objective was to assess the quality of the information provided by AI like ChatGPT and establish if it is a reliable source of information for our patients.

Materials and methods

On March 18th, 2023, we formulated several questions in English to the latest version of ChatGPT 4.0 about the following urologic diseases: bladder cancer, prostate cancer, renal cancer, benign prostatic hypertrophy (BPH), and urinary stones. We asked a general question about the pathology (What is *pathology*?), followed by a question about its general treatment (What options of treatment exist for *pathology*?). We also asked about the most frequent surgical treatment for each disease: transurethral resection of bladder tumour, transurethral resection of the prostate, Holmium enucleation of Prostate (HoLEP), radical prostatectomy, radical nephrectomy, ureteroscopy (URS), percutaneous nephrolithotomy (PCNL) and shock wave lithotripsy (ESWL). The responses provided are available for consultation in the open science framework (OSF) repository [9, 10].

We collected the answers for each question, analysed the number of words per answer and checked how much incorrect information was present in every answer according to the European Urology Association (EAU) clinical guidelines. We then proceeded to evaluate the answers concerning the treatment and global information of each disease using the validated DISCERN quality questionnaire [9]. This validated test is made up of by 15 questions and an overall quality rating (16th questions), and it has been designed to judge and standardise the quality of the information on the different treatment options. Each question is scored from 1 (minimum) to 5 (maximum). To assess the question about surgical treatment, we used the instrument for evaluating the quality of informed consent documents developed by Spatz et al. [11]. The questionnaire is formed by 8 items. Seven items related to the content of the procedure (how it is performed, rationale for the procedure, quantitative and qualitative probability of risk, benefits, and possible alternatives) [11]. There is a last regarding timing, which was no analysed due to lack of patients in the study. The maximum possible score was 15 points. The employed tool is disponible as Supplementary Information.

The above-mentioned data was analysed by two urological general attending with 10 years' experience (FJDG and AMGT). Both were familiarised with the tests used, consulting previously appropriate manuals. Errors found in the answers were pooled and the results obtained from the DISCERN and the informed consent instrument were obtained by performing a mean of both scores.

No patients were enrolled in our study, evaluation by the study ethics committee was not required.

Results

The overall information provided in all pathologies is well-balanced and follows an established structure. Firstly, each pathology is explained with its anatomical location, affected population and a description of the most frequent symptoms. It concludes with the established risk factors and possible treatment.

In the case of bladder cancer, ChatGPT recommends that the patients receive regular check-ups from their doctor, which was not in other pathologies. Regarding prostate cancer, it is emphasized that the tumour is often not life-threatening and identifies PSA and digital rectal exam as a screening method. Prostate cancer is the only area with an error (Table 1). Concerning benign prostatic hypertrophy, the emphasis is placed on its benign condition and its impact on quality of life. Patients are advised that their quality of life should guide their treatment, which may be managed through changes in their habits. The formation of lithiasis is briefly described and it is noted that it can occur in different sections of the urinary tract. The description of renal cancer is the shortest mentioning the cause is not fully understood. Nevertheless, it does mention some established risks factor.

All answers correctly identified the different treatment modalities, with a moderate quality score of 3 out of 5 on the DISCERN-16 scale (as shown in Table 1). However, the lowest scores were related to sources of information, which were missing. An additional question was required to obtain

Table 1 Word count, number of errors and score in the 16th question in DISCERN instrument regarding the answers about overall information and treatment

	Overall information		Treatment	
	Word count	Number of errors	Word count	DISCERN-16th item
Bladder cancer	249	0	286	3
HBP	207	0	466	3
Urinary stones	204	0	191	3
Prostate cancer	222	1	279	3
Kidney cancer	200	0	247	3

sources such as Medline or American Cancer Society websites. The information provided on the different treatments was balanced, outlining their possible benefits. However, details on the risks, complications, and potential impact on the patient's quality of life were lacking. Additionally, no information was provided on what would happen if the patient chooses not to undergo treatment. It was highlighted that the decision on the therapeutic option should be made in consultation with a specialist.

Regarding bladder cancer, the treatment approach depends on the stage of the disease, as explained by the source. For superficial cases, transurethral resection is the most used method, with other treatments being reserved for more advanced stages. ChatGPT also correctly mentions Bacillus Calmette-Guerin instillations as a possible treatment option for early-stage bladder cancer.

In the case of BPH, the treatment options are more diverse, including medication, as well as minimally invasive procedures like laser surgery. However, ChatGPT only mentions transurethral resection as a surgical option and does not differentiate between different surgical techniques such as holmium enucleation of the prostate (HoLEP) or conventional open surgery. Therefore, although ChatGPT provides comprehensive information, it may not be as complete as it omits some of the available treatments.

As for prostate cancer, ChatGPT identifies various treatment modalities, including active surveillance, surgery, radiotherapy, hormone therapy, and chemotherapy. However, it does not provide indications for each option and once again highlights the importance of consulting with a doctor to determine the best treatment approach for each patient's specific case.

The primary treatment option identified for renal cancer is nephrectomy, with partial nephrectomy being mentioned as an alternative option. Targeted therapy or immunotherapy are cited as additional treatment options for advanced stages. The importance of making a joint decision regarding the treatment plan is emphasized.

Regarding lithiasis treatment, the approach depends on size, composition, location, and symptoms. The initial treatment options include analgesia, hydration, and medical expulsive therapy. ESWL, URS, and PCNL are the surgical procedures mentioned. The possibility of catheterization in cases of urolithiasis is not addressed, either as an emergent treatment of urosepsis or after definitive treatment.

The information regarding surgical treatment complements the general treatment responses provided. For instance, it explains that in the case of bladder cancer, transurethral resection of the bladder is usually the initial step, with other treatments being reserved for more advanced stages. It provides clear information on the potential risks involved, the duration of the procedure, and its specifics.

In all cases, the recovery time, type of anaesthesia, and potential complications are listed. However, there is no quantitative data on complications or any indication of how the procedures may impact the patient's quality of life. Moreover, unless specifically asked, no other treatment options are mentioned, apart from bladder cancer.

The description of HoLEP is brief, only mentioning that the prostate is removed via suction without providing a detailed explanation of the procedure. It notes that the risk and complication rates are lower than with traditional treatments. Radical prostatectomy is discussed in more detail, including different surgical approaches and types. Nephrectomy is differentiated between partial and radical procedures. When discussing lithiasis treatment, the focus is on renal procedures and flexible scopes used in URS, with no mention of urethral lithiasis. Laser and basket options are identified, but there is no discussion of post-procedure drain or catheter requirements. ESWL is noted to be effective for all lithiasis and renal lithiasis smaller than 2 cm. PCNL is specifically for larger lithiasis that cannot be treated with ESWL or other methods, and its score of 11 is the highest among the treatments listed in Table 2.

After analysing all the responses, which include disease definition, potential treatments, and detailed surgical options, all pathologies achieved a DISCERN score of 4,

Table 2 Analysis of answers regarding each surgical treatment: word count and score in the instrument for evaluating the quality of informed consent documents developed by Spatz et al. [11]

Surgery	Word count	Score in instrument for evaluating the quality of informed consent documents [11] Maximum score = 15
Transurethral resection of bladder	235	9
Transurethral resection of prostate	252	9
Holmium enucleation of prostate	162	4
Ureteroscopy	232	8
Percutaneous nephrolithotomy	208	11
Shock wave lithotripsy	208	6
Radical nephrectomy	247	7
Radical prostatectomy	200	6

except BPH which obtained a score of 3. This improvement is due to the identification of potential associated risks and the consideration of various treatment options. None of the pathologies provides additional literature on the disease unless a new question is asked, and there is no mention of sources of information. BPH received a lower score due to potential biases in its treatment options and descriptions.

Discussion

The rapid advancement of technology has significantly impacted our daily lives. Similarly, the fields of medicine and urology have also seen significant changes with a wealth of information now available through social media and audiovisual platforms [12]. With the emergence of artificial intelligence, patients now can search for information regarding their condition [13]. This study evaluates the potential responses that urological patients can receive through this new technology.

The quality of information on social networks regarding various urological pathologies has been examined. Most of the content related to conditions such as prostate and bladder cancer was found to be inadequate [3, 4]. Additionally, in some cases, the information had a commercial bias [5].

ChatGPT has been trained using a dialogue model that encompasses all freely available information on the internet, including human conversations. Its knowledge base covers various areas, not limited to medicine alone. The model may have assimilated data from clinical practice guidelines, professional texts, or scientific articles found on the internet, but it does not have access to restricted private health system data [1, 14]. Our study suggests that the quality of the information provided by ChatGPT may vary depending on the pathology being discussed. Well-researched conditions such as prostate or renal cancer may have higher-quality information, while newer procedures may have lower-quality information.

ChatGPT has been identified as a potential tool for public health in providing advice on aspects such as quality of life, vaccination, screening, and risk factor reduction. However, its limitations include the lack of context and direct interaction with healthcare professionals [6]. In the future, the development of AI and machine-learning methodologies could further enhance clinical practice by facilitating decision-making in diagnosis and treatment, as well as aiding training using simulation models [13].

The ChatGPT chat learning method has the potential to introduce biases by incorporating untruthful information from internet sources, which may contain a commercial component [1, 14]. In our study, we observed this bias in the information related to BPH pathology. Traditional surgeries, such as transurethral resection of the prostate or open adenectomy, were not mentioned as a standard surgical treatment. Instead, other procedures such as HoLEP and minimally invasive procedures were defined as effective with less impact on the quality of life and fewer side effects. In other studies, analysing responses to questions about diabetes, it was observed that ChatGPT provides adequate and understandable answers to general questions, but it starts to fail once more complex dialogue boxes are asked [15]. However, no commercial brands or preferences for any type of treatment were observed.

Although our study is limited to the patient's inquiries to ChatGPT, it could potentially be utilized to facilitate healthcare professionals' daily work and decision-making. For instance, ChatGPT could transcribe the information provided to the patient during the medical visit, giving them a summary of their visit and an overview of their pathology [13]. Also, ChatGPT was capable of summarizing the patient-physician conversation, providing a medical record note, and accurately answering questions about a clinical case of chronic obstructive pulmonary disease [14]. Our study's results support this potential use, as the response set analysed showed good quality and detailed information for each surgery. If the medical staff provide this information, and AI analyse it together with other evidence, it could be shared with patients in written form and inform the treatment plan and follow-up planning [14].

One limitation of the ChatGPT is its inability to verify the information. It only provides additional information if specifically requested and does not disclose the bibliography used to obtain the answer. This leads to a poor DISCERN score for questions 7–8 since the sources of information cannot be verified for accuracy.

It is important to acknowledge that we do not have direct control over the responses generated by AI. Professional associations should engage with developers to ensure that accurate and tailored answers about common urological conditions are provided by AI in a clear and detailed manner, following the lines of action established for social media [16]. It should be noted that ChatGPT is not the only AI model available, and others such as Google's LaMDA are currently under development. In the future, there may be the possibility of developing customized medical dialogue boxes [13, 14, 17]. However, these chatbots may assist in better practice but will not replace the clinical expertise (such as physical examination) [13].

Our study has a few limitations. Firstly, we did not conduct a validated test to compare the information provided by ChatGPT with clinical practice guidelines. Additionally, the informed consent questionnaire we used was not entirely suitable for a free-text format like that provided by ChatGPT. Furthermore, the responses to the posed inquiries may differ depending on the formulation of the questions, which need not be identical to the one we have asked [18]. This fact must

be considered, especially in patients whose questions may differ from those asked by a urologist and the interpretation of the answers may vary. However, the questions posed were simple and valid questionnaires have been employed to assess them.

Conclusions

ChatGPT has demonstrated its ability to provide information in a user-friendly manner, however the information should be used with caution since it doesn`t disclose the sources and may contain bias even with simple questions related to the basics of urologic diseases. To obtain a higher quality of information ChatGPT requires asking multiple questions. As technology continues to advance, artificial intelligence and other new technologies may become useful tools in the doctor-patient relationship.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00345-023-04563-0.

Author contributions JJS: Project development, data collection and analysis, manuscript writing. CTF: data collection. ARA: data analysis, manuscript writing. AGT: Manuscript writing. FJDG: Manuscript writing. LLG: Project development, manuscript editing, final approval.

Funding All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Data sharing The questions and answers of ChatGPT can be consulted in Open Science Framework repository (https://doi.org/10.17605/OSF.IO/8UNQV). The correspondence author will provide data from valid questionarries upon reasoned request.

Declarations

Conflict of interest The authors declare no conflict of interests.

Ethics approval and informed consent No human patients were involved in the study. No need Ethics Committee was required.

References

1. OpenAI Help Center ChatGPT general FAQ. https://help.openai.com/en/articles/6783457-chatgpt-general-faq. Accessed 9 Mar 2023
2. Fraser H, Coiera E, Wong D (2018) Safety of patient-facing digital symptom checkers. Lancet 392(10161):2263–2264. https://doi.org/10.1016/S0140-6736(18)32819-8
3. Loeb S, Sengupta S, Butaney M et al (2019) Dissemination of misinformative and biased information about prostate cancer on YouTube. Eur Urol 75(4):564–567. https://doi.org/10.1016/j.eururo.2018.10.056
4. Loeb S, Reines K, Abu-Salha Y et al (2021) Quality of bladder cancer information on YouTube. Eur Urol 79(1):56–59. https://doi.org/10.1016/j.eururo.2020.09.014
5. Betschart P, Pratsinis M, Müllhaupt G et al (2020) Information on surgical treatment of benign prostatic hyperplasia on YouTube is highly biased and misleading. BJU Int 125(4):595–601. https://doi.org/10.1111/bju.14971
6. Biswas SS (2023) Role of ChatGPT in Public Health. Ann Biomed Eng. https://doi.org/10.1007/s10439-023-03172-7. (**published online ahead of print, 2023 Mar 15**)
7. Baclic O, Tunis M, Young K, Doan C, Swerdfeger H, Schonfeld J (2020) Challenges and opportunities for public health made possible by advances in natural language processing. Can Commun Dis Rep 46(6):161–168. https://doi.org/10.14745/ccdr.v46i06a02
8. Millenson ML, Baldwin JL, Zipperer L, Singh H (2018) Beyond Dr Google: the evidence on consumer-facing digital tools for diagnosis. Diagnosis (Berl) 5(3):95–105. https://doi.org/10.1515/dx-2018-0009
9. Szczesniewski JJ, Tellez Fouz C, Ramos Alba A, Garcia Tello A, Diaz Goizueta FJ, Llanes Gonzalez L (2023) Answers of ChatGPT to questions about urologic diseases. OSF (Database). https://doi.org/10.17605/OSF.IO/8UNQV
10. Charnock D, Shepperd S, Needham G, Gann R (1999) DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 53(2):105–111. https://doi.org/10.1136/jech.53.2.105
11. Spatz ES, Suter LG, George E et al (2020) An instrument for assessing the quality of informed consent documents for elective procedures: development and testing. BMJ Open 10(5):e033297. https://doi.org/10.1136/bmjopen-2019-033297
12. Rivas JG, Socarrás MR, Blanco LT (2016) Social media in urology: opportunities, applications, appropriate use and new horizons. Cent Eur J Urol 69(3):293–298. https://doi.org/10.5173/ceju.2016.848
13. Haug CJ, Drazen JM (2023) Artificial Intelligence and machine learning in clinical medicine. N Engl J Med 388(13):1201–1208. https://doi.org/10.1056/NEJMra2302038
14. Lee P, Bubeck S, Petro J (2023) Benefits, limits, and risks of GPT-4 as an AI Chatbot for medicine. N Engl J Med 388(13):1233–1239. https://doi.org/10.1056/NEJMsr2214184
15. Sng GGR, Tung JYM, Lim DYZ, Bee YM (2023) Potential and pitfalls of ChatGPT and natural-language artificial intelligence models for diabetes education. Diabetes Care. https://doi.org/10.2337/dc23-0197. (**published online ahead of print, 2023 Mar 15**)
16. Fode M, Jensen CFS, Østergren PB (2021) How should the medical community respond to the low quality of medical information on social media? Eur Urol 79(1):60–61. https://doi.org/10.1016/j.eururo.2020.09.050
17. Singhal K, Azizi S, Tu T et al (2022) Large language models encode clinical knowledge. http://arxiv.org/abs/2212.13138. Accessed 30 Mar 2023
18 Whiles BB, Bird VG, Canales BK, DiBianco JM, Terry RS (2023) Caution! AI bot has entered the patient chat: ChatGPT has limitations in providing accurate urologic healthcare advice. Urology. https://doi.org/10.1016/j.urology.2023.07.010

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).
Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.
These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.
We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.
While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.
These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.
To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.
Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.
If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com