ORIGINAL RESEARCH

# ChatGPT: Is This Patient Education Tool for Urological Malignancies Readable for the General Population?

Ivan Thia [ID], Manmeet Saluja

Department of Urology, Royal Perth Hospital, Perth, WA, Australia

Correspondence: Ivan Thia, Email ivan.thia@health.wa.gov.au

**Background:** With widespread adoption of technological advancements in everyday life, patients are now increasingly able and willing to obtain information about their health conditions, treatment options, and indeed expected outcomes via the convenience of any device than can access the worldwide web. This introduces another aspect of patient care in the provision of healthcare for the modern doctor. ChatGPT is the first of an increasing number of self learning programs that have been released recently which may revolutionize and impact healthcare delivery.

**Methods:** The aim of this study is to obtain an objective measure of the readability of information provided on ChatGPT when compared with current validated patient information sheets provided by government health institutions in Western Australia. The same structured questions were input into the program for three major urological malignancies (urothelial, renal, and prostate), with the response generated evaluated with a validated readability scoring system – Flesch-Kincaid reading ease score. The same scoring system was then applied to current patient information sheets in circulation from Cancer Council Australia and UpToDate.

**Results:** Findings in this study looking at ease of readability of information provided on ChatGPT as compared to other government bodies and healthcare institutions confirm that they are non-inferior and may be a useful tool or adjunct to the traditional clinic based consultations. Ease of use of the information generated from ChatGPT was increased further when the question was modified to target an audience of 16 years of age, the average level of education attained by an Australian.

**Discussion:** Future research can be done to look into incorporating the use of similar technologies to increase efficiency in the healthcare system and reduce healthcare costs.

**Keywords:** ChatGPT, innovative education, patient education, urological malignancies

## Background

Cancer incidence and prevalence is rising steadily over time, with only 47,414 cases diagnosed in 1982 in Australia, rising to 162,163 in 2021.[1] Correspondingly, diagnosis of new cases of urological malignancies have increased as well, from 6,545 cases in 1982 to 31,988 cases in 2022.[1] This sharp increase in the diagnosis of cancer can be attributed to improved health awareness through public health campaigns, increased use of screening services, technological advancements and proliferation, as well as an aging population and modifiable lifestyle factors.[2] However, rural and remote communities have consistently suffered from poorer health outcomes, with high disease burden, low population screening, delayed treatment, and low overall 5-year cancer survival rates compared with metropolitan populations.[3] Part of the reason for this disparity in health outcomes is attributed to lower education levels, poor health literacy, and lack of access to timely specialist healthcare services. Western Australia is also unique amongst the countries and states of the world, being the largest and most rural state, with 28% of the population living in rural and remote areas.[4] Therefore, access to healthcare is vital to improve the overall general health of these populations. The advent of the worldwide web and increasing adoption of virtual technology has "closed the gap" dramatically in the last decade, as can be witnessed during the COVID-19 pandemic.[5] Utilizing Telehealth conferencing, online resources, multi-way conversations over platforms

**31**

such as Microsoft Teams and Zoom to name but a few equips the modern day health professional with a multitude of methods to communicate with patients effectively.[6] Similarly, patients can now have access to a large volume of information about their health conditions at the click of a mouse, without needing to seek medical advice or explanations from a medical professional.

ChatGPT is an online, open access platform that utilizes semi-supervised or self supervised learning tools called large language models (LLMs) to generate human-like responses to routine conversational inputs.[7] In essence, it is a condensed search engine used to answer a specific targeted question posed by the user and relay the answers in an eloquent, contextually appropriate, conversational manner. Since the release of ChatGPT 3.5 in November 2022, there have been numerous papers on ChatGPT of late, detailing the benefits and pitfalls of its use in medicine, especially in the area of patient information provision.[8–11] The current ChatGPT-4 has even been shown to be able to generate literature papers that would have otherwise been successfully published in reputable journals, being able to bypass traditional plagiarism software.[12] The number of recorded users have increased dramatically as well, with more than 100 million registered users within 2 months of its launch. The potential of tapping on ChatGPT to aid in providing useful, up to date information across a variety of medical conditions to all patients and improve access to healthcare information remains very enticing despite emerging issues with its ethical consideration, credibility, and limitations.

## Aim

The aim of this study is to determine if information provided by ChatGPT for urological malignancies is at an appropriate readability level for the target audience. The average Australian according to a census from the Australian Bureau of Statistics in 2021 would have received mandatory education to year 10 equivalent in secondary school, with an average Australian having a reading ability equivalent to year 8, or between age 13–14.[13] Therefore, a reasonable level at which medical information should be pitched for the majority of the population should be at year 8. The Flesch-Kincaid reading ease score developed in 1975 for the United States Navy remains one of the most widely used validated readability tests for English texts.[14] Although very simplistic in nature, this tool allows for a validated evaluation of the texts in question without use of complex linguistic metrics and hence has been adopted in this study.[15] This assessment tool takes into account the length of sentences and the number of syllables per word to determine the level of reading difficulty of a text. It is scored from 0–100, with a higher score reflecting a text that is easier to read. The reading ability for a Year 8 individual corresponds to a Flesch-Kincaid score of 60.

## Study Design

A number of questions were pre-selected by two Urology consultants from Royal Perth Hospital whilst taking into consideration existing patient information sheets to ensure comparability for the three most common primary urological cancers, including prostate, bladder, and renal cancers (Table 1). The questions were written based on a questionnaire databank collected from patients seen in clinics to determine the most frequent questions that patients have regarding their own illnesses. Questions were posed at a generic population level as well as at the level of year 8 repeatedly with answers recorded and assessed for ease of readability with the Flesch-Kincaid reading ease score. Three other sources of cancer-related information were collected as a comparison; information from the Cancer Council Australia as well as the predominantly American databases UpToDate and the Mayo Clinic. Cancer Council Australia (CCA) is the pre-eminent national organization in undertaking cancer research, patient support and education, and development of a cancer control policy.[16–18] UpToDate and the Mayo Clinic, on the other hand, are clinician authored resources that provides current, evidence based statements about a variety of medical conditions to aid clinicians in making point-of-care decisions.[19–29] Content was taken from the patient information section of the databases, with "basic" and "beyond basic" discussions from UpToDate being evaluated independently. The information was collected from existing online resources made available to the general public.

Content of information provided was assessed independently by two separate Urology specialists to determine if it was adequate and accurate. Each question was posed five times to ChatGPT with Flesch-Kincaid score counted as an average and used for comparison between all three sources. A further inclusion criteria on questions posed on ChatGPT whereby the target audience was selected at year 8 education level was also evaluated as a comparison. Statistical

**Table 1** Readability comparison between various information sources using Flesch-Kincaid readability score

| Questions Posed | Flesch-Kincaid Score (*p* value) |
|---|---|
| **General information about prostate cancer** | |
| ChatGPT | 43.0 ± 2.9 |
| ChatGPT (Year 8) | 57.7 ± 2.9 |
| Cancer Council Australia | 51.8 (*p* = 0.09, *p* = 0.11) |
| UpToDate (Basics) | 65.4 (*p* <0.05, *p* = 0.18) |
| Mayo Clinic | 53.5 (*p* <0.05, *p* = 0.14) |
| **General information about prostate cancer screening** | |
| ChatGPT | 34.9 ± 2.8 |
| ChatGPT (Year 8) | 45.5 ± 10.9 |
| UpToDate (Basics) | 71.6 (*p* <0.05, *p* = 0.06) |
| UptoDate (Beyond Basics) | 45.5 (*p* = 0.14, *p* = 1) |
| **Low risk prostate cancer treatment** | |
| ChatGPT | 18.6 ± 1.8 |
| ChatGPT (Year 8) | 36.7 ± 2.2 |
| UpToDate (Basics) | 52.3 (*p* <0.05, *p* <0.05) |
| **Advanced prostate cancer treatment** | |
| ChatGPT | 28.4 ± 3.3 |
| ChatGPT (Year 8) | 46.6 ± 1.1 |
| UpToDate (Beyond Basics) | 42.2 (*p* = 0.05, *p* = 0.06) |
| **General information about renal cell cancer** | |
| ChatGPT | 26.8 ± 4.0 |
| ChatGPT (Year 8) | 55.1 ± 1.5 |
| Cancer Council Australia | 48.2 (*p* <0.05, *p* <0.05) |
| UptoDate (Basics) | 68.5 (*p* <0.05, *p* <0.05) |
| UpToDate (Beyond Basics) | 43.8 (*p* <0.05, *p* <0.05) |
| Mayo Clinic | 57.2 (*p* <0.05, *p* = 0.15) |
| **General information about bladder cancer** | |
| ChatGPT | 36.2 ± 2.9 |
| ChatGPT (Year 8) | 51.6 ± 2.6 |
| Cancer Council Australia | 42.4 (*p* = 0.17, *p* = 0.07) |
| UpToDate (Basics) | 66.8 (*p* <0.05, *p* <0.05) |
| Mayo Clinic | 50.3 (*p* <0.05, *p* = 0.33) |
| **Bladder cancer diagnosis and staging** | |
| ChatGPT | 35.6 ± 0.7 |
| ChatGPT (Year 8) | 49.6 ± 1.6 |
| UpToDate (Beyond Basics) | 42.6 (*p* <0.05, *p* <0.05) |
| **Non-muscle invasive bladder cancer treatment** | |
| ChatGPT | 25.9 ± 1.1 |
| ChatGPT (Year 8) | 36.6 ± 1.3 |
| UpToDate (Beyond Basics) | 37.8 (*p* <0.05, *p* = 0.06) |
| **Muscle invasive bladder cancer treatment** | |
| ChatGPT | 31.2 ± 3.6 |
| ChatGPT (Year 8) | 61.1 ± 2.5 |
| UpToDate (Beyond Basics) | 35.8 (*p* = 0.17, *p* <0.05) |

significance was calculated using a two sample *T*-test to determine if there was a difference in readability level between the different sources of information.

# Results

Figure 1. Readability comparison between different sources using Flesch-Kincaid readability score.

The average readability score for information provided by ChatGPT was 31.2 ± 12.6, compared with the 41.3 ± 5.5 (*p* = 0.26) for UpToDate Beyond Basics, 47.5 ± 5.1 (*p* = 0.17) for CCA, 53.7 ± 3.5 for the Mayo Clinic (*p* = 0.11), and 64.9 ± 12.6 (*p* = 0.09) for UpToDate Basics (Figure 1). This meant that, overall, the difference in readability of information provided by chatGPT was not statistically significant comparable to those provided by CCA, the Mayo Clinic, and UpToDate.

However, when individual cancer information was looked into in greater detail, apart from questions regarding general information about prostate cancer, all other questions posed with the general public as the target group were at
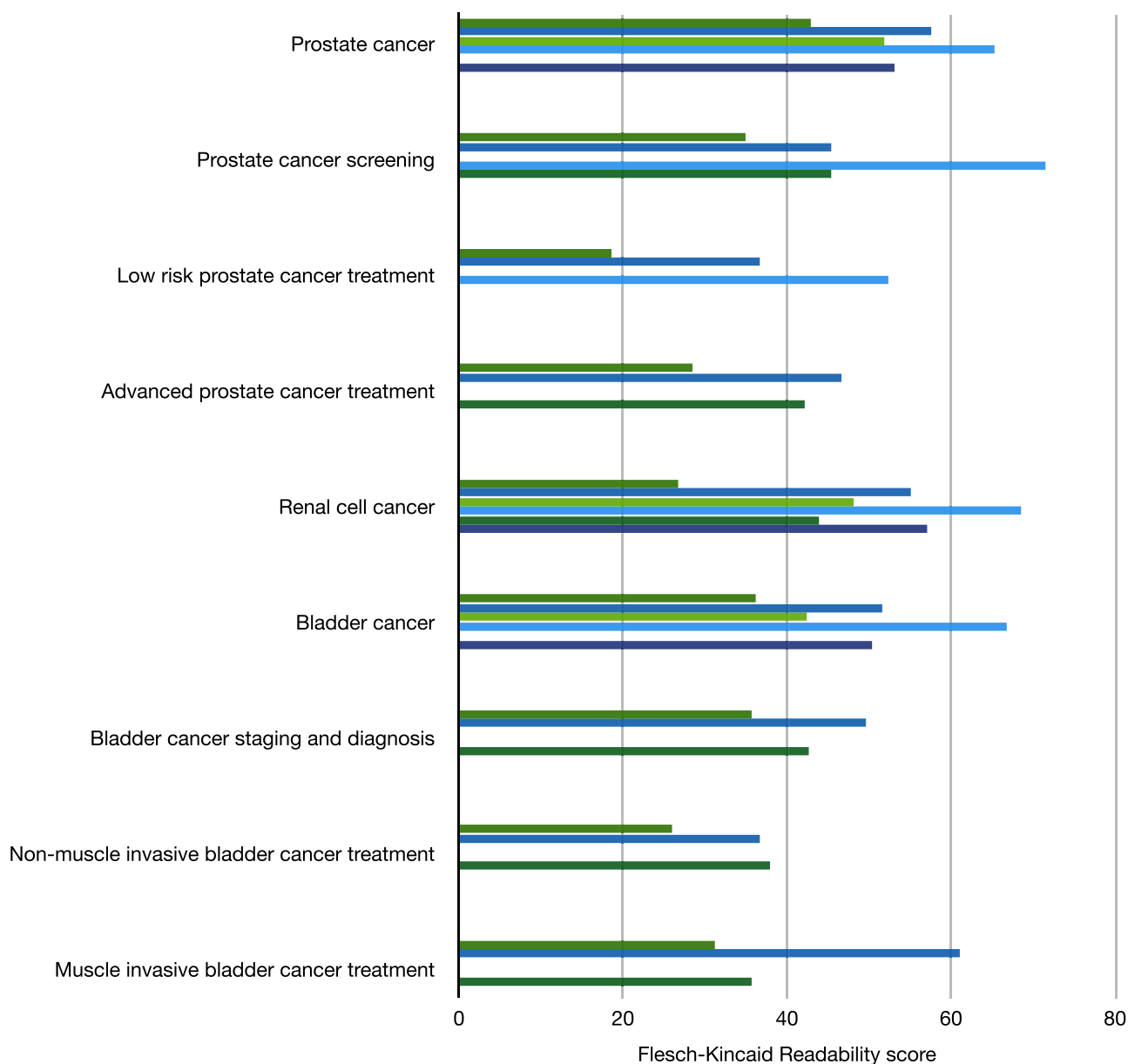


**Figure 1** Readability comparison between different sources using Flesch-Kincaid readability score.

a more difficult readability level when compared to basic patient information sheets provided on UpToDate. When comparison was made with the more in-depth (Beyond Basics) information sheet on UpToDate, readability between the two were comparable. Ease of use between information provided on ChatGPT and CCA was comparable for information regarding prostate and bladder cancer, with the readability of information on renal cancer being significantly easier for information provided on CCA.

When the additional criteria was included for information to be pitched to a 16-year-old audience, we found that the information provided was made significantly more readable, with an average improvement of $16.25 \pm 12.05$ on the readability scale. In most cases, this made the information provided easier to understand when compared to information on CCA or the Beyond Basics section on UpToDate.

Information provided for each category was examined to determine if it was accurate and sufficient to allow an average person without medical knowledge to learn more about the cancer in question. Through rigorous reviews independently by two specialist urologists, it was deemed to contain the basic information required for an understanding of each condition without being overwhelming or misleading. However, the consensus was that information provided by CCA and the Mayo Clinic were more patient focused and covered more aspects of the urological malignancies, including peri-operative care, surveillance, and patient support information, which is a vital aspect and very useful in aiding patients in understanding their own condition.

## Discussion

The instant acceptance and proliferate use of ChatGPT makes it difficult to regulate and monitor or validate information that would otherwise be detrimental to patients receiving care or impact on patient–doctor relationships. In addition, information provided by government bodies, health organizations, and medical institutions have been specially designed for use as educational tools by patients, and have been screened and amended to suit a variety of target audiences, taking into consideration health literacy of the population, whereas information provided by artificial intelligence (AI) softwares may not exhibit the same motivation. To date, this is the first analysis of the readability of AI software in providing urological medical information to patients in the literature.

This study demonstrates the remarkable versatility and potential of ChatGPT when used appropriately. The instantaneous difference in language use, medical jargon, and resultant ease of readability with a single modification of age of the target audience is presented above. Gender, occupation, and ethnicity, among other non-modifiable personal information included in the question posed, would elicit even more targeted and improved individualized responses. The information included in the responses hase been demonstrated to be non-inferior to that presented by established healthcare governing bodies in terms of readability and lends weight to the usefulness of similar AI programs as adjuncts in dissemination of health literature and increasing health literacy.

The difficulty in the employment of intuitive AI technologies in medical practice stems from the inability to control and regulate what is being presented when questions are posed. Indeed, even when the same question is posed a different answer is generated with varying content and readability with consistency. Furthermore, the constant inundation of information on a daily basis means that information need to be constantly updated. The current ChatGPT-4 model is still based on formation from September 2021, meaning that current medical advancements are not included in the generated responses.

There are several limitations to this study. Firstly, a direct comparison of readability was made based on information acquired and does not look in depth into separate content of the various sources. It is difficult to qualitatively assess information provided to patients based on a retrospective cross-sectional study, though further studies in the future can be performed in this regard, looking at patient satisfaction and understanding when presented with the various sources.

Secondly, the Flesch-Kincaid reading ease score does not incorporate the impact that diagrams, images, audiovisual content, and indeed webpage design and layout have on patient understanding. The CCA and Mayo Clinic information sheets employ these tools extensively and this helps to further improve ease of understanding. To our knowledge, there is no tool that reliably assesses non-text content of digital or print information, thereby introducing bias into our study.

This is the first study to date looking at readability of information generated by ChatGPT when prompted about urological malignancies. Our study confirms that the majority of information provided by medical institutions around the

world is above the average reading ability of an Australian layperson, ChatGPT included. The complex nature of urological malignancies necessitate a higher education level to be able to understand and assimilate the concepts presented. Therefore, a clinician is vital in helping patients to understand about their disease, available treatment options, and how it impacts on their activities of daily living so that patients can make an informed choice regarding shared treatment decision. Open access information systems like ChatGPT should only be used as an adjunct to clinical consults until further improvements can be made in the future.

A potential development in the future may include certain aspects of ChatGPT in the creation of a natural language chatbot with a medical journal database so that information extracted include only validated and peer reviewed data rather than indiscriminate crawling through the worldwide web. This database could be updated periodically to include the most up to date information so that the responses generated are current.

## Conclusion

ChatGPT, when used appropriately, can revolutionize healthcare by providing a useful platform or tool with which to improve health literacy within the community. In its current state, ChatGPT cannot be considered as a source of factual information as the generated responses require further evaluation and validation. Further studies can be performed to determine how technology can make healthcare even more accessible and affordable to rural and remote communities, which traditionally suffer from worse healthcare outcomes.

## Ethics Statement

Given that ChatGPT is a public platform with accessibility freely available to the general public, there are no permission requirements for the use of information generated in this study. All information is kept confidential and available for review if required to ensure reproducibility. There were also no human subjects involved in this study, therefore an ethics committee was not formed for its review.

## Disclosure

The authors report no conflicts of interest in this work.

## References

1. ABS (Australian Bureau of Statistics). *Population Projections, Australia, 2017 (Base) – 2066*. ABS cat. no. 3222.0. Canberra: ABS;2018.
2. van Hoogstraten L, Vrieling A, van der Heijden A, et al. Global trends in the epidemiology of bladder cancer: challenges for public health and clinical practice. *Nat Rev Clin Oncol*. 2023;20:287–304. doi:10.1038/s41571-023-00744-3
3. Bhatia S, Landier W, Paskett E, et al. Rural-Urban Disparities in Cancer Outcomes: opportunities for Future Research. *JNCI*. 2022;114(7):940–952. doi:10.1093/jnci/djac030
4. Australian Bureau of Statistics. *Regional Population*. Available from:. Canberra: ABS; 2019.
5. Gray K, Chapman W, Khan U, et al. The Rapid Development of Virtual Care Tools in Response to COVID-19 Case Studies in Three Australian Health Services. *JMIR Form Res*. 2022;6(4):e32619. doi:10.2196/32619
6. Hargreaves C, Clarke AP, Lester KR. Microsoft Teams and team performance in the COVID-19 pandemic within an NHS Trust Community Service in North-West England. *Team Performance Management*. 2022;28(1/2):79–94. doi:10.1108/TPM-11-2021-0082
7. Biswas SS. Role of Chat GPT in Public Health. *Ann Biomed Eng*. 2023;51(5):868–869. doi:10.1007/s10439-023-03172-7
8. Dave T, Athaluri S, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Front Artif Intell*. 2023;6. doi:10.3389/frai.2023.1169595
9. Cakir H, Caglar U, Yildiz O, et al. Evaluating the performance of ChatGPT in answering questions related to urolithiasis. *Int Urol Nephrol*. 2023. doi:10.1007/s11255-023-03773-0.
10. Caglar U, Yildiz O, Meric A, et al. Evaluating the performance of ChatGPT in answering questions related to paediatric urology. *J Paed Urol*. 2023. doi:10.1016/j.jpurol.2023.08.003
11. Whiles B, Bird V, Canales B, Di Bianco J, Terry R. Caution! AI Bot has entered the patient chat: chatGPT has limitations in providing accurate urologic healthcare advice. *Urology*. 2023;180:278–284. doi:10.1016/j.urology.2023.07.010
12. Ruksakulpiwat S, Kumar A, Ajibade A. Using ChatGPT in Medical Research: current Status and Future Directions. *J Multidisciplinary Healthcare*. 2023;16:1513–1520. doi:10.2147/JMDH.S413470
13. Australian Bureau of Statistics. *Education and Training: Census*. Canberra: ABS; 2021. Available from: https://www.abs.gov.au/statistics/people/education/education-and-training-census/2021. Accessed January 8, 2024.
14. Solnyshkina M, Zamaletdinov R, Gorodetskaya L, Gabitov A. Evaluating Text Complexity and Flesch-Kincaid Grade Level. *J Social Studies Educ Res*. 2017;8(3):238–248.
15. Lines NA. The Past, Problems, and Potential of Readability Analysis. *Chance*. 2022;35(2):16–24. doi:10.1080/09332480.2022.2066411

16. Cancer Council Australia. Understanding bladder cancer; 2022. Available from: https://www.cancer.org.au/cancer-information/types-of-cancer/bladder-cancer. Accessed January 8, 2024.

17. Cancer Council Australia. Understanding kidney cancer; 2020. Available from: https://www.cancer.org.au/cancer-information/types-of-cancer/kidney-cancer. Accessed January 8, 2024.

18. Cancer Council Australia. Understanding prostate cancer; 2022. Available from: https://www.cancer.org.au/cancer-information/tyoes-of-cancer/prostate-cancer. Accessed January 8, 2024.

19. Baskin LS. Patient education: prostate cancer (the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.

20. Baskin LS. Patient education: prostate cancer screening (the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.

21. Clemens JQ. Patient education: prostate cancer (beyond the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.

22. Clemens JQ. Patient education: choosing treatment for low-risk prostate cancer (beyond the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.

23. Foust-Wright CE. Patient education: treatment for advanced prostate cancer (beyond the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.

24. Kassouf W. Patient education: kidney cancer (the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.

25. Khera M. Patient education: renal cell carcinoma (kidney cancer) (beyond the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.

26. Lotan Y. Patient education: bladder cancer (the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.

27. Mattoo TK. Patient education: bladder cancer diagnosis and staging (beyond the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.

28. McVary KT. Patient education: bladder cancer treatment; non-muscle invasive (superficial) cancer (beyond the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.

29. O'Leary M. Patient education: bladder cancer treatment; muscle invasive cancer (beyond the basics). Post TW, ed. UpToDate. Waltham, MA: UpToDate Inc. Available from: http://www.uptodate.com. Accessed July 29, 2023.