

Original Article

Artificial Intelligence Chatbot as a Companion for Cancer Patients About Most Common Questions: Analysis of Readability and Quality

Efe Cem Erdat, Elif Berna Köksoy, Güngör Utkan

Ankara University Faculty of Medicine, Department of Medical Oncology, Ankara, Türkiye

ABSTRACT

Aim: Advances on artificial intelligence (AI) have led to development of AI chatbots and more people are using AI chatbots to seek answers to their questions every day. We conducted this study to investigate the readability and quality of answers generated by large language model AI chatbots as companions in answering questions for cancer patients.

Methods: After surveying 508 patients admitted to the outpatient clinic of Ankara University Faculty of Medicine, Department of Medical Oncology, we selected the most frequently asked questions about the four most common cancer types and general cancer knowledge. We asked these questions of ChatGPT (an AI chatbot from OpenAI) and calculated readability and quality scores, and the statistical difference between suggested and calculated reading scores. Means and the t-tests (one-way and/or paired) were used for statistical analysis.

Results: A total of 57 questions, including those about colorectal, breast, lung, prostate cancer, and general cancer questions, were selected for analysis. The mean Flesch Reading Ease Score for all questions was 48.18 [standard deviation (SD) ± 11.65], which was significantly lower than the suggested reading score of 60 points ($p < 0.01$). The mean score for graded readability scores was 13.21 (SD ± 2.49), which was consistent with college-level readability and significantly higher than a suggested value of 6th graders ($p < 0.01$). The mean DISCERN score of all questions was 51.98 (SD ± 7.27) and the Global Quality Score was 3.91 (SD ± 0.69). Breast cancer responses were easier to read on graded scales ($p = 0.02$) and had higher quality ($p = 0.05$).

Conclusion: ChatGPT may be a good companion for cancer patients despite its limitations, but it should be used carefully.

Keywords: Artificial intelligence, chatbots, cancer education, ChatGPT

Introduction

Approximately 40% of individuals are expected to develop cancer during their lifetime, and nearly 1.9 million new cases of cancer are expected to be diagnosed in the United States in 2022 [1]. Most cancer patients have questions they want to ask their oncologist about their cancer; however, most oncologists spend less than 25 minutes per visit with each patient [2]. And the time they do spend with their patients is not enough to answer all the patients' questions. Consequently, most questions remain unanswered, and patients use search engines for their questions, and most of the answers they find may be misleading. Patients are concerned about cancer

and strive to find reliable sources of online information such as blogs, videos, and news sites [3,4]. Most patients tend to use trustworthy websites; they need a reliable source of information. Given that many individuals have low health literacy, it is essential to provide adequate and reliable health information to meet patients' needs [5].

Advances in neural networks, deep learning, and artificial intelligence (AI) have led to the development of large language model (LLM) AI chatbots. While older chatbots could respond in simple sentences, newer chatbots such as ChatGPT (an AI chatbot from OpenAI) have advanced to the point of generating more sophisticated responses that can fool even experienced scientists [6,7]. Since its public opening in

Address for Correspondence: Efe Cem Erdat MD, Ankara University Faculty of Medicine, Department of Medical Oncology, Ankara, Türkiye

E-mail: cemerdat@gmail.com, ecerdat@ankara.edu.tr **ORCID ID:** orcid.org/0000-0002-1250-1297

Received: 12.08.2024 **Accepted:** 20.12.2024 **Epub:** 06.03.2025 **Publication Date:** xxxxxxxxxxxxxx

Cite this article as: Erdat EC, Köksoy EB, Utkan G. Artificial intelligence chatbot as a companion for cancer patients about most common questions: analysis of readability and quality. Acta Haematol Oncol Turc. [Epub Ahead of Print]



November 2022, ChatGPT has become a prominent source of information that is likely to outrank popular search engines. Nowadays, an increasing number of people are using ChatGPT to access online information, and more companies are starting to develop new AI chatbots.

As an increasing number of people use ChatGPT, it is likely that cancer patients and their relatives especially will utilize ChatGPT to access more medical information. Therefore, we should be concerned about the quality and readability of the information provided by AI chatbots such as ChatGPT.

This study aimed to evaluate the appropriateness and readability of the chatbot-generated responses to patients' questions.

Methods

Question Selection

Between February 1, 2023, and February 28, 2023, after obtaining verbal and written informed consent, we surveyed patients and their caregivers admitted to Ankara University Faculty of Medicine Outpatient Oncology Clinic, Department of Medical Oncology. Ankara University Faculty of Medicine Ethical Board of Human Research approved the study protocol (application number: AUTF-KAEK 2023/34, approval date: 1.2.2023). We asked about the most common questions they had asked online about the four most common cancers worldwide (colorectal cancer, breast cancer, lung cancer, and prostate cancer) in Turkish. The in-person survey collected only the patients' primary malignancy diagnosis, the 10 most common questions they asked online, and didn't include any demographic information. The patients' questions were answered in the outpatient clinics, but were not included in the analyses. The most frequently asked questions about colorectal cancer, breast cancer, lung cancer, prostate cancer, as well as general cancer information were extracted from the survey data, translated into English, and consolidated for analysis.

Asking Questions

The questions collected in the patient survey, in English, were asked to ChatGPT following a prefatory statement such as "Can you answer the following questions as I have... cancer?" We asked ChatGPT the questions sequentially, and after each set of questions about a particular type of cancer, we cleared the browser cache and opened a new chat session. All sessions were performed on the same computer, over virtual machines created from scratch, on custom user accounts. Apart from the pre-determined questions, no additional questions or potentially interfering sentences were posed to ChatGPT. The collected answers were copied into plain unformatted text files and prepared for further analysis.

Readability and Reliability Analyses of Responses

The readability analyses were performed using the TextStat package on Python 3.11 only with simple command line

prompts using plain unformatted texts. We calculated the Flesch Reading Ease and the grade-equivalent reading scales Flesch-Kincaid (FK) Grade Level, SMOG Index, Gunning-Fog Score, Automated Readability Index (ARI), Coleman-Liau Index, and Linsear Write Formula. The mean readability score and standard deviation were calculated to obtain more accurate results.

To evaluate the readability and quality of the materials, we utilized several established scoring systems and scales. Readability was assessed using the FK Grade Level, which determines the US school grade level required to comprehend the text based on syllables per word and words per sentence. The SMOG Index estimated the years of education needed by focusing on the number of polysyllabic words, while the Gunning-Fog Scale assessed text complexity by considering average sentence length and the percentage of complex words. Additionally, the ARI and Coleman-Liau Scale provided grade level estimates based on characters per word and letters per 100 words, respectively. The Linsear Write Formula further evaluated readability by distinguishing between easy and hard words.

The DISCERN Score and Global Quality Score (GQS), although not originally developed for written AI responses, were calculated for reliability and quality analyses of all responses from each author. All authors have agreed on the DISCERN and GQS Scores. The DISCERN Score was used to measure the reliability and quality of the information, with higher scores indicating better quality. The GQS offered an overall subjective evaluation of the content's flow, ease of use, and reliability. These scoring systems collectively ensured a comprehensive analysis of both the complexity and quality of the materials, enhancing the study's overall rigor.

Statistical Analysis

Means were calculated for descriptive analysis, and Student's t-test was used for continuous variable analysis. One-sided one-sample t-tests were calculated for the analysis of the difference between the Flesch Reading Ease Score proposed value of 60 and graded readability of 6th-grade level [8]. Differences between scores of different cancer types were calculated with paired samples t-test. Statistical analyses were performed using Microsoft Excel (Microsoft Corporation, Redmond, WA) and R 4.1 (the R Foundation, Vienna, Austria).

Results

Determining the Most Common Questions About the Four Most Common Cancer Types

A total of 508 patients participated in the study, including 226 with colorectal cancer, 115 with breast cancer, 91 with lung cancer, and 76 with prostate cancer. After reviewing the collected questions, we selected 57 questions, including 11 questions on colorectal cancer, breast cancer, lung cancer, and general cancer knowledge; and 13 questions on prostate cancer. The most common questions that patients ask online are listed in Table 1. Selected questions were mostly about survival depending on stages, how are treatments done,

benefits of treatments, and they were consistent with the National Health Institute's "Questions to Ask Your Doctor about Treatment" [9].

Ease of Reading

The mean Flesch Reading Ease Score for all questions was 48.18 [standard deviation (SD) ± 11.65] meaning the answers were difficult to read. In topic-specific calculation, the mean Flesch Reading Ease Score was 46.27 (SD ± 11.31) for colorectal cancer, 52.66 (SD ± 13.49) for breast cancer, 48.36 (SD ± 11.3) for lung cancer, 46.27 (SD ± 12.52) for prostate cancer, and 47.12 (SD ± 10.25) for general questions. All topics were difficult to read, except for breast cancer, with questions on breast cancer being fairly challenging in score interpretation. The Flesch Reading Ease Score ranged from 15.24 to 70.36. No significant differences were found for colorectal cancer (Student's t-test, $p=0.66$), breast cancer (Student's t-test, $p=0.16$), lung cancer (Student's t-test, $p=0.96$), prostate cancer (Student's t-test, $p=0.51$), and general questions (Student's t-test, $p=0.74$) when each was compared to the others. The mean score of Flesch Reading Ease was significantly lower

than the suggested score of at least 60 points in all topics (one-way t-test, $p<0.01$).

The mean score of graded readabilities for all questions was 13.21 (SD ± 2.49), corresponding to college-graded reading level requirements. In topic-specific calculations, the mean scores were 13.43 (SD ± 2.46) for colorectal cancer, 11.93 (SD ± 2.16) for breast cancer, 13.58 (SD ± 2.87) for lung cancer, 13.32 (SD ± 1.71) for prostate cancer, and 13.8 (SD ± 2.82) for general cancer questions. Breast cancer responses had a lower readability score, equivalent to a high school grade level, while the other topics were at a college grade level. The graded readability scores ranged from 8.55 to 18.96. The mean score was lower ($p=0.02$) for breast cancer responses than for colorectal cancer (Student's t-test, $p=0.69$), lung cancer ($p=0.56$), prostate cancer ($p=0.82$), and general questions ($p=0.28$). The mean value of graded readability was significantly higher in all topics than the suggested value for sixth graders (one-way t-test, $p<0.001$). The one-way ANOVA test showed no significance among the graded readability scores ($p=0.21$). A summary of the reading ease and graded readability scores is shown in Table 2.

Table 1. Most frequently asked questions

Colorectal cancer (n=226)	Breast cancer (n=91)	Lung cancer (n=91)	Prostate cancer (n=76)	General questions (n=506)
How is colorectal cancer staged and what is the prognosis and survival depending on stage?	Where does breast cancer spread and metastasize?	How is lung cancer staged, categorized and what is the prognosis like depending on stage and type?	What are the stages of Prostate cancer and how are they treated depending on stage?	What is an adjuvant chemotherapy and what are the benefits?
How colorectal cancer treated if I had surgery?	Which chemotherapies are used in breast cancer and what are the side effects?	Which one is better for early-stage LC, surgery, radiotherapy, or chemotherapy? What are the benefits and potential harms of these treatment modalities?	How are prostatectomy or prostate irradiation done and which one is better?	Is adjuvant chemotherapy necessary when treating patients?
What is adjuvant treatment and which chemotherapies are used in adjuvant treatment?	How is mastectomy done and what are the problems that I will suffer afterward?	What is the probability of recurrence and survival of NSCLC depending on stages?	What is the hormone therapy for prostate cancer and how long does it take?	How is cancer staged and what is the importance of cancer stage?
How long should the treatment take?	How is breast cancer staged? What is the early or late stage of breast cancer?	What are the adjuvant treatments used on NSCLC and what are the benefits and side effects?	What is castration, and what does castration-sensitive or castration-resistant mean?	How is cancer treated and what modalities are being used?
When should the PET/CT done?	When should the breast reconstruction be done, same time as mastectomy or afterward?	Which chemotherapies are used on lung cancer and what are the side effects?	How does prostate cancer spread and metastasize?	What is cancer and there are how many cancer types?
Is radiotherapy necessary?	What is the probability of losing hair on breast cancer chemotherapy and will it be permanent?	What are the molecular profiles on pathology reports of lung cancer and what do they mean?	What is Gleason's Score and what does it mean?	Is there a chance that I will recover from late-stage cancer?

Table 1. Continued

Colorectal cancer (n=226)	Breast cancer (n=91)	Lung cancer (n=91)	Prostate cancer (n=76)	General questions (n=506)
What should I eat to prevent my cancer from recurring?	What are estrogen receptor, progesterone receptor, HER2, and Ki67 on my breast cancer pathology report and what do they mean?	How can I know if I'm a candidate for smart drugs?	Does prostate cancer curable in later stages?	What are the main side effects of chemotherapy and how can I overcome the side effects?
Are metastases operable? Is metastatic disease curable?	How long does adjuvant radiotherapy take?	Is SCLC curable? If it is not curable, which problems will I encounter if I don't take any chemotherapy?	When is a PET scan taken on prostate cancer?	How should my diet be included to prevent my cancer from recurring?
I have colon cancer. What is the probability of developing cancer in my family?	What are the hormonal therapies, how long do they take and what are the benefits and side effects?	When can I receive immunotherapy for lung cancer, and will it cure my cancer?	What targeted therapies are mostly used on prostate cancer, and what are the benefits and side effects?	When are the targeted therapies used in cancer treatment?
When is immunotherapy used for colon cancer?	What are the targeted therapies, when are they used, and what are the benefits and side effects?	Where does lung cancer spread and metastasize and how metastases are treated?	What will I encounter in my sexual life while on prostate cancer treatment?	Is surgery necessary when treating early-stage cancer?
Is FOLFOX or XELOX better?	What is the expected survival of breast cancer depending on stages?	Should I quit smoking?	When is chemotherapy used for prostate cancer?	Can tumor markers predict cancer recurrence or do they have any use in diagnosis?
-	-	-	Is only PSA monitorization sufficient in prostate cancer treatment?	-
-	-	-	What are the survival rates of prostate cancer depending on stages?	-

PET/CT: Positron emission tomography/computed tomography, SCLC: Small-cell lung cancer, NSCLC: Non-SCLC, PSA: Prostate-specific antigen, HER2: Human epidermal growth factor receptor 2, LC: Lung cancer

Table 2. Readability scales

Question topic	Flesch Reading Ease		Graded scales	
	Mean (\pm SD)	p*	Mean (\pm SD)	p*
Colorectal cancer	46.84 (\pm 11.31)	0.66	13.43 (\pm 2.46)	0.69
Breast cancer	52.66 (\pm 13.49)	0.16	11.93 (\pm 2.16)	0.02
Lung cancer	48.36 (\pm 11.3)	0.96	13.53 (\pm 2.87)	0.56
Prostate cancer	46.27 (\pm 12.52)	0.51	13.32 (\pm 1.71)	0.82
Common questions	47.12 (\pm 10.25)	0.74	13.8 (\pm 2.82)	0.28
Total	48.18 (\pm 11.65)	-	13.21 (\pm 2.49)	-

*Paired samples t-test, the analyses were used to determine the difference between one topic and all the others.
SD: Standard deviation

Mean scores for graded readability were 13.02 (SD \pm 2.19) on the FK Scale, 13.86 (SD \pm 1.90) on the SMOG Scale, 12.87 (SD \pm 2.01) on the Gunning-Fog Scale, 14.72 (SD \pm 2.45) in the ARI

Scale, 12.65 (SD \pm 1.63) in the Coleman-Liau Scale, and 13.13 (SD \pm 3.47) in the Linsear Scale. The difference between Scales was significant ($p < 0.05$) in the FK, SMOG, Gunning-Fog, ARI,

and Coleman-Liau Scales, whereas it wasn't significant in the Linsear Scale (paired-samples t-test, $p=0.77$).

Quality of Responses

The mean DISCERN score was 51.98 (SD ± 7.27) for all questions, indicating responses of good quality. In topic-specific calculation, the mean DISCERN scores were 54.27 (SD ± 8) for colorectal cancer, 52.91 (SD ± 9.19) for breast cancer, 49.27 (SD ± 6.75) for lung cancer, 53 (SD ± 4.6) for prostate cancer, and 50.27 (SD ± 7.59) for another cancer type. The DISCERN scores ranged between 33 and 65. Mean DISCERN values did not differ for any cancer type ($p=0.17-0.57$) and all topics had good-quality responses. For question 4 of DISCERN "Is it clear what sources of information were used to compile the publication (other than the author or producer)?" only one question received 2 points, the others 1 point, and for question 5 of DISCERN "Is it clear when the information used or reported in the publication was produced?" all questions

received 1 point because ChatGPT didn't specify the sources it used to prepare the responses. The means and standard deviations of the DISCERN questions are shown in Table 3.

The mean GQS for all questions was 3.91 (SD ± 0.69); indicating that the responses were of nearly good quality. In the topic-specific calculations, the mean GQS scores were 3.73 (SD ± 0.65) for colorectal cancer, 4.27 (SD ± 0.65) for breast cancer, 3.64 (SD ± 0.81) for lung cancer, 3.85 (SD ± 0.55) for prostate cancer, and 4.09 (SD ± 0.7) for general questions. The breast cancer, and general questions were rated as good quality in GQS scoring; whereas colorectal cancer, prostate cancer, and lung cancer were rated as medium quality. The GQS scores ranged between 2 and 5 for all topics. The mean GQS score was significantly higher for breast cancer than for the other cancers, with a statistically significant difference indicated by $p<0.05$. Details of the GQS scores can be found in Table 4.

None of the responses included references, and most responses (7/57) included a disclaimer and sentences with the

Table 3. DISCERN questionnaire

Question number	DISCERN questions	Mean	SD
1	Are aims clear?	4.72	0.53
2	Does it achieve its aims?	4.32	0.81
3	Is it relevant?	4.30	0.80
4	Is it clear what sources of information were used to compile the publication (other than the author or producer)?	1.02	0.13
5	Is it clear when the information used or reported in the publication was produced?	1.00	0.00
6	Is it balanced and unbiased?	4.35	0.79
7	Does it provide details of additional sources of support and information?	1.58	0.53
8	Does it refer to areas of uncertainty?	2.21	0.98
9	Does it describe how each treatment works?	3.53	1.38
10	Does it describe the benefits of each treatment?	3.37	1.42
11	Does it describe the risks of each treatment?	3.04	1.36
12	Does it describe what would happen if no treatment is used?	2.28	1.22
13	Does it describe how the treatment choices affect the overall quality of life?	3.51	1.21
14	Is it clear that there may be more than one possible treatment choice?	4.44	1.10
15	Does it provide support for shared decision-making?	4.37	1.22
16	Based on the answers to all of the above questions, rate the overall quality of the publication as a source of information about treatment choices	3.96	0.73
-	Total DISCERN score	51.98	7.27

SD: Standard deviation

Table 4. Statistics of GQS by topics

Question topic	Mean	SD	p*
Colorectal cancer	3.73	0.65	0.33
Breast cancer	4.27	0.65	0.05
Lung cancer	3.64	0.81	0.14
Prostate cancer	3.85	0.55	0.69
Common questions	4.09	0.7	0.34
Total	3.91	0.69	-

*Paired samples t-test, the analyses were used to determine the difference between one topic and all the others.

GQS: Global Quality Score, SD: Standard deviation

meaning that it is recommended to consult the health care provider.

Discussion

The quality and readability of online medical information have long been a subject of debate, and as technology continues to evolve, more issues will be added. LLM AI chatbots have been in use for several years, and concerns about their reliability have been growing since their public introduction. ChatGPT is one of the first commercially successful LLM AI chatbots, and more than 400 scientific articles or editorials have been written about it from its release in November 2022 to May 2023.

The information obtained by ChatGPT remains controversial. There is a possibility that it provides fabricated information (hallucinations) and gives non-existent references. Additionally, since it is not connected to the Internet and trained on Q4 2021 information, it may deliver outdated or wrong information about anything, and the fabricated abstracts can even fool experienced researchers [6,7,10].

Although literacy is increasing in the world, experts suggest that medical articles published online should be easy to read, and materials should ideally be written at a 6th-grade level [8]. While more patients are reading medical information about their disease online; older adults tend to prefer a direct doctor visit when they have questions, so the 6th-grade reading level may not fully reflect the actual information needs of patients. Even patients with college degrees may encounter misleading information or be unable to distinguish between fake or fabricated medical articles and real, trustworthy medical information. The Health on the Net recognition seal is an important tool for assessing the reliability of an online article. It is recommended to check reliability and quality using a higher score from DISCERN or JAMA [11]. However, due to the nature of LLM AI chatbots, their evaluation is challenging. Therefore, people should be cautious when using these tools for their health-related questions. Additionally, there is a gap in online cancer information, and individuals tend to seek answers to their questions not only in online articles or search engines but also in videos on platforms such as YouTube [4,12].

Most cancer patients spend about 5% of their remaining lifetime in the health care system. Since most oncologists spend less than 25 minutes per visit with their patients, their families are frustrated and try to find answers to their questions online [2,13]. Guy and Richardson's [13] study suggests that the pay-for-performance system leads to an increase in patient volume and a decrease in visit time, which in turn may lead to unanswered questions. AI chatbots, especially ChatGPT, are expected to provide great convenience for cancer patients to access online information [10].

Our results showed that readability on the Flesch Reading Ease score was lower than the recommended score of 60. College-level reading was required to understand the responses, which is well above the recommended educational level of a sixth grader. A study by Li et al. [14] showed that most online

information about four common cancers had a grade-level readability score of 10.9, which is consistent with our findings. Stevens et al. [15] showed that even the online information on a narrow topic such as neoadjuvant treatment of pancreatic cancer had a readability of grade level 10.96; Ozduran and Büyükçoban [16] showed that the information about post-Coronavirus disease-2019 pain had a readability grade level 9.83-10.9. Because ChatGPT was trained to use online information up until Q4 2021, the lower readability could be a result of the training data collected online.

Although not designed for scoring LLM responses, the average DISCERN score for questions was 51, which can be considered a good score. ChatGPT only answers the question asked and does not provide further details. DISCERN scoring asks questions about alternative treatments, such as what happens if you apply them or do not apply them. If ChatGPT asked more questions about DISCERN, the average score of DISCERN would be higher. However, this might lead to a decrease in overall quality and an increase in the risk of hallucinations. The GQS assessment yielded a mean score of 3.91, which is close to good quality and could reflect higher quality if more comprehensive questions were included.

Study Limitations

The main limitations of our study are the survey didn't include demographic data such as age, or stage of the cancer patients, and ChatGPT can give a personalized answer and therefore cannot be easily evaluated. Additionally, since the survey did not include data on educational level, our analysis could not examine patients' educational status or assess whether it differed from the proposed 6th-grade reading level. The patients' questions were translated into English, and the responses were also retrieved in English, thus, the evaluation in Turkish is not available. The patients or their caregivers did not ask questions about the specific survival of the patient. This may be caused by Turkish cultural aspects because in Turkish culture it is considered inappropriate to talk about survival and death. Since the questions we asked were short and concise, lacking in detail, the quality of the response might have been affected. Other AI chatbots, besides ChatGPT, weren't included in our study, and the results cannot be generalized to all AI chatbots. Although the study used multiple graded scoring systems, it examined only the mean scores, limiting its generalizability and the manuscript's readability. Despite these limitations, ChatGPT appears to be a good information-gathering tool for cancer patients. There is also a need for better tools to evaluate the quality of information provided by AI chatbots, as well as newer AI chatbots that are specialized in medical topics. Further studies are needed to confirm our results.

Conclusion

ChatGPT, by answering unresolved questions, can be a useful source of information for people undergoing cancer treatment. However, the answers generated require a higher level of education than the recommended 6th-grade level, making

them more difficult to understand. Despite this limitation, the quality of the responses can be good when assessed against both the DISCERN and GQS Scales. As the responses produced depend largely on the questions asked, it is important to be cautious when relying on AI chatbots. In addition, further research is needed to develop updated scales for assessing the quality of responses generated by chatbots.

Ethics

Ethics Committee Approval: Ankara University Faculty of Medicine Ethical Board of Human Research approved the study protocol (application number: AUTF-KAEK 2023/34, approval date: 01.02.2023).

Informed Consent: Verbal and written informed consent was obtained.

Acknowledgments

No artificial intelligence tools of any kind were used for preparation of manuscript, only the selected answer of selected questions was prompted into OpenAI's ChatGPT. None of the authors has any affiliation with artificial intelligence companies.

Presented in: This manuscript is orally presented in 12th Ege Hematology and Oncology Congress which takes part between 14-16 March 2024 in Radisson Blu Hotel Çeşme-İzmir, Türkiye.

Footnotes

Authorship Contributions

Concept: E.C.E., G.U., Design: E.C.E., G.U., Data Collection or Processing: E.C.E., E.B.K., G.U., Analysis or Interpretation: E.C.E., E.B.K., G.U., Literature Search: E.C.E., E.B.K., Writing: E.C.E.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin.* 2022;72:7-33.

2. Medscape Oncologist Compensation Report 2018. Last Accessed Date: 21.02.2025. Available from: <https://www.medscape.com/slideshow/2018-compensation-oncologist-6009663>.
3. Budenz A, Sleight AG, Klein WMP: A qualitative study of online information-seeking preferences among cancer survivors. *J Cancer Surviv.* 2022;16:892-903.
4. Ayoub G, Chalhoub E, Sleilaty G, Kourie HR. YouTube as a source of information on breast cancer in the Arab world. *Support Care Cancer.* 2021;29:8009-8017.
5. U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. National Action Plan to Improve Health Literacy. Washington, DC; 2010. Last Accessed Date: 21.02.2025. Available from: https://odphp.health.gov/sites/default/files/2019-09/Health_Literacy_Action_Plan.pdf
6. Else H. Abstracts written by ChatGPT fool scientists. *Nature.* 2023;613:423.
7. Gao CA, Howard FM, Markov NS, et al. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digit Med.* 2023;6:75.
8. Rooney MK, Santiago G, Perni S, et al. Readability of patient education materials from high-impact medical journals: a 20-year analysis. *J Patient Exp.* 2021;8:10.1177/2374373521998847.
9. National Cancer Institute: questions to ask your doctor about treatment. Last Accessed Date: 22.08.2023. Available from: <https://www.cancer.gov/about-cancer/treatment/questions>
10. Hopkins AM, Logan JM, Kichenadasse G, Sorich MJ. Artificial intelligence chatbots will revolutionize how cancer patients access information: ChatGPT represents a paradigm-shift. *JNCI Cancer Spectr.* 2023;7:10.
11. Battineni G, Baldoni S, Chintalapudi N, et al. Factors affecting the quality and reliability of online health information. *Digit Health.* 2020;6:2055207620948996.
12. Baskin AS, Wang T, Mott NM, Hawley ST, Jaggi R, Dossett LA. Gaps in online breast cancer treatment information for older women. *Ann Surg Oncol.* 2021;28:950-957.
13. Guy GP Jr, Richardson LC. Visit duration for outpatient physician office visits among patients with cancer. *J Oncol Pract.* 2012;8:2-8.
14. Li JZH, Kong T, Killow V, et al. Quality assessment of online resources for the most common cancers. *J Cancer Educ.* 2023;38:34-41.
15. Stevens L, Guo M, Brown ZJ, Ejaz A, Pawlik TM, Cloyd JM. Evaluating the quality of online information regarding neoadjuvant therapy for pancreatic cancer. *J Gastrointest Cancer.* 2023;54:890-896.
16. Ozduran E, Büyükçoban S. Evaluating the readability, quality and reliability of online patient education materials on post-covid pain. *PeerJ.* 2022;10:13686.