Submitted: 08 October, 2023 Accepted: 29 December, 2023 Published: 30 June, 2024 DOI:10.22514/j.androl.2024.010

REV INT ANDROL
Revista Internacional de Andrología

ORIGINAL RESEARCH

Can ChatGPT help patients understand their andrological diseases?

İsmail Emre Ergin[1,*], Adem Sancı[2]

[1] Kızılcahamam State Hospital, 06890 Ankara, Turkey
[2] Etlik City Hospital, 06010 Ankara, Turkey

***Correspondence**
dremreergin@cumhuriyet.edu.tr
(İsmail Emre Ergin)

Abstract

We aimed to assess the reliability of Chat Generative Pre-training Transformer (ChatGPT)'s andrology information and its suitability for informing patients and medical students accurately about andrology topics. We presented a series of systematically organized frequently asked questions on andrology topics and sentences containing strong recommendations from the European Association of Urology (EAU) Guideline to ChatGPT-3.5 and 4.0 as questions. These questions encompassed Male Hypogonadism, Erectile Dysfunction and Sexual Desire Disorder, Disorders of Ejaculation, Penile Curvature and Penile Size Abnormalities, Priapism, and Male Infertility. Two expert urologists independently evaluated and assigned scores ranging from 1 to 4 to each response based on its accuracy, with the following ratings: (1) Completely true, (2) Accurate but insufficient, (3) A mixture of accurate and misleading information, and (4) Completely false. A total of 120 questions were included in the study. Among these questions, 50.0% received a grade of 1 (completely correct) (55.4% for 4.0 version). The combined rate of correct answers (grades 1 and 2) was 85.2% for frequently asked questions (88.8% for 4.0 version) and 81.5% for questions obtained from the guideline. The rate of completely incorrect answers (grade 4) was 1.8% for frequently asked questions (0% for 4.0 version) and 5.2% for questions based on strong recommendations. The response rate of version 4.0 to questions created from sentences containing strong recommendations from the EAU guideline was the same as version 3.5. ChatGPT provided satisfactory answers to the questions asked, although some responses lacked completeness. It may be beneficial to utilize ChatGPT under the guidance of a urologist to enhance patients' comprehension of their andrology issues.

Keywords

ChatGPT; Andrology; Artificial intelligence; Education

¿ChatGPT puede ayudar a los pacientes a comprender sus enfermedades andrológicas?

Resumen

El uso de Internet y de sistemas de información basados en inteligencia artificial está aumentando en muchos ámbitos. Nuestro objetivo era evaluar la confiabilidad de la información de andrología de Chat Generative Pre-training Transformer ChatGPT y su idoneidad para informar a pacientes y estudiantes de medicina con precisión sobre temas de andrología. Presentamos una serie de preguntas frecuentes organizadas sistemáticamente sobre temas de andrología y oraciones que contienen recomendaciones sólidas de la Guía de la Asociación Europea de Urología (EAU) para ChatGPT-3.5 y 4.0 como preguntas. Estas preguntas abarcaban hipogonadismo masculino, disfunción eréctil y trastorno del deseo sexual, trastornos de la eyaculación, curvatura del pene y anomalías del tamaño del pene, priapismo e infertilidad masculina. Dos urólogos expertos evaluaron de forma independiente y asignaron puntuaciones que van del 1 al 4 a cada respuesta en función de su precisión, con las siguientes calificaciones: (1) Completamente cierto, (2) Preciso pero insuficiente, (3) Una mezcla de información precisa y engañosa, y (4) Completamente falso. En el estudio se incluyeron un total de 120 preguntas. Entre estas preguntas, el 50.0% recibió una calificación de 1 (completamente correcta) (55.4% para la versión 4.0). La tasa combinada de respuestas correctas (grados 1 y 2) fue del 85.2% para las preguntas frecuentes (88.8% para la versión 4.0) y del 81.5% para las preguntas obtenidas de la guía. La tasa de respuestas completamente incorrectas (grado 4) fue del 1.8% para las preguntas frecuentes (0% para la versión 4.0) y del 5.2% para las preguntas basadas en recomendaciones sólidas. La tasa de respuesta de la versión 4.0 a las preguntas creadas a partir de oraciones que contienen recomendaciones sólidas de la guía EAU fue la misma que la de la versión 3.5. ChatGPT proporcionó respuestas satisfactorias a las preguntas formuladas, aunque algunas respuestas no fueron completas. Puede resultar beneficioso utilizar ChatGPT bajo la guía de un urólogo para mejorar la comprensión de los pacientes sobre sus problemas andrológicos.

Palabras Clave

ChatGPT; Andrología; Inteligencia artificial; Educación

1. Introduction

Andrological problems, with their broad subcategories, impact a significant portion of the male population [1–4]. When we consider subcategories such as male hypogonadism, erectile dysfunction (ED), ejaculation disorders, infertility, penis size abnormalities, penile curvature, and priapism, it becomes evident that they collectively affect a substantial segment of society [2–4]. Epidemiological studies published to date have consistently indicated a high worldwide incidence and prevalence of ED [1]. For example, a study conducted in Boston reported that the overall prevalence of ED in men aged 40 to 70 was 52% [2]. Regarding Premature Ejaculation (PE), epidemiological data indicates that approximately 5% of the general population has an ejaculation latency of less than 2 minutes, both for acquired PE and lifetime PE [3]. An article on penile deformity stated that the prevalence of definite and probable cases of Peyronie's disease in the USA was 0.7% and 11%, respectively [4]. Approximately 15 percent of couples seek medical assistance because they cannot achieve pregnancy in the first year [5]. Fifty percent of couples have a factor associated with male infertility, usually abnormal semen parameters [6]. As can be understood from the data above, andrological diseases constitute common health problems in society and are frequently encountered in urology practice.

The use of the internet and artificial intelligence-supported information systems is increasing in many areas. Since andrological problems are typically sensitive issues that patients may hesitate to discuss, they often prefer to seek preliminary information and satisfy their curiosity through the internet and artificial intelligence-supported information systems before consulting a healthcare professional. Medical students can also utilize these systems to enhance their education due to their speed and ease of accessing information.

Chat Generative Pre-training Transformer (ChatGPT) is an artificial intelligence program created by OpenAI that provides instant answers to questions using an Internet database [7]. It can be accessed through a smartphone application or an Internet browser. This program, which helps users access information by instantly synthesizing data on various subjects, is rapidly gaining popularity. Several articles have been published investigating ChatGPT's ability to provide accurate information in the medical field. In a study conducted by Gilson *et al.* [8], it was determined that ChatGPT provided satisfactory answers to 60% of medical school exam questions. In another study, despite its limitations, Caglar *et al.* [9] found that ChatGPT provided satisfactory answers to questions related to pediatric urology. Although there are many articles published so far investigating the accuracy and reliability of ChatGPT's answers in various medical fields [7–10], we noticed that there is no such study in the field of andrology. Therefore, in this study, we investigated the accuracy and validity of ChatGPT's answers on andrology and the usability of these answers in terms.

2. Material-methods

All questions frequently asked by patients on health websites, urology association websites, and social media (Facebook, Instagram, and Twitter) were recorded. Health website selection was made as follows: When the sub-topic to be evaluated was searched on Google, the frequently asked questions on the patient information sites that first appeared were evalu-

ated. Strongly recommended recommendations on each topic of the sexual and reproductive health section of the EAU were transformed into questions [11]. Inappropriate questions, questions related to personal health, and repetitive questions were excluded from the study. The questions obtained from the internet and the questions obtained from the EAU Guidelines were categorized as follows: Male Hypogonadism, Erectile Dysfunction and Sexual Desire Disorder, Disorders of Ejaculation, Penile Curvature and Penile Size Abnormalities, Priapism, Male Infertility.

Questions were answered using free ChatGPT-3.5 version and subscription paid version 4.0 on internet browsers with cookies and history cleared. The answers provided by ChatGPT were evaluated and scored by two urologists. The authors assigned a score to each answer on a scale of 1 to 4: If the answer correctly included all the information that a urologist should provide to a patient, it was rated as 1 (completely correct); if it was correct but insufficient, it was rated as 2; if it contained a mix of correct and misleading information, it was rated as 3; and if it contained completely incorrect answers, it was rated as 4.

Discrepancies in the scoring between the two auditors were resolved by taking the arithmetic average, and consecutively scored answers were reviewed collaboratively by the auditors, evaluating them based on primary sources.

ChatGPT's responses to all questions were generated after confirmation by reposing the same questions from different IP addresses, thus confirming the repeatability of the answers. We asked ChatGPT about the difficulty level of each question and classified them based on its responses. According to anwers, all questions were categorized as easy, medium or difficult.

3. Statistical analysis

The data collected from the answers were imported into Excel (Version 16, Microsoft, USA). The accuracy of the answers was assessed, and the accuracy percentages were calculated based on andrology topics. All data from both groups were compared using SPSS for Windows (Version 25.0, IBM Corp., Armonk, NY, USA).

4. Results

A total of 120 questions were initially included in the study. However, 28 questions were excluded based on exclusion criteria. Among these excluded questions, 13 were repetitive, 11 were grammatically inadequate, and 4 pertained to personal health. After applying the exclusion criteria, 92 questions remained and were included in the study. Fig. 1 illustrates the questions included and excluded in the research.

The answer rates provided by ChatGPT for all questions are presented in Table 1. Among the questions, 50.0% received a rating of "completely correct" (grade 1) (57.6% for 4.0 version). Approximately 33.6% of the questions had missing information (grade 2) (32.2% for 4.0 version), 13.3% were a mix of accurate and misleading information (grade 3) (11.1% for 4.0 version), and 3.3% were answered completely incorrectly (grade 4) (2.2% for 4.0 version). Out of the 92 questions, 77 received ratings of "completely correct" or "correct but insufficient" (80 for 4.0 version).

The topic with the fewest completely correct answers in both versions was "Ejaculation Disorders" 33.3% for 3.5 version (44.4% for 4.0 version). Questions that received completely incorrect answers were related to "Penile Curvature and Size Abnormalities" and "Priapism" issues.

The grading of answers provided for frequently asked questions and questions obtained from the EAU Guidelines is presented in Table 2. The rate of completely correct answers to frequently asked questions on the subjects of Male Hypogonadism, Erectile Dysfunction and Sexual Desire Disorder, Disorders of Ejaculation, Penile Curvature and Size Abnormalities, Priapism, and Male Infertility was 80.0%, 60.0%, 42.8%, 50.0%, 44.4% and 38.4%, respectively. These rates were 50%, 66%, 0%, 30%, 77.7% and 42.8%, respectively, in the questions generated from the EAU guidelines. For version 4.0, the rates of answers to the EAU guidelines questions were the same, but the rates of answers to the frequently asked questions were different. The rate of completely correct answers were 80.0%, 80.0%, 57.1%, 60.0%, 44.4% and 46.1% respectively.

All questions were categorized as easy, medium, and difficult. Among the frequently asked questions, 38 questions were

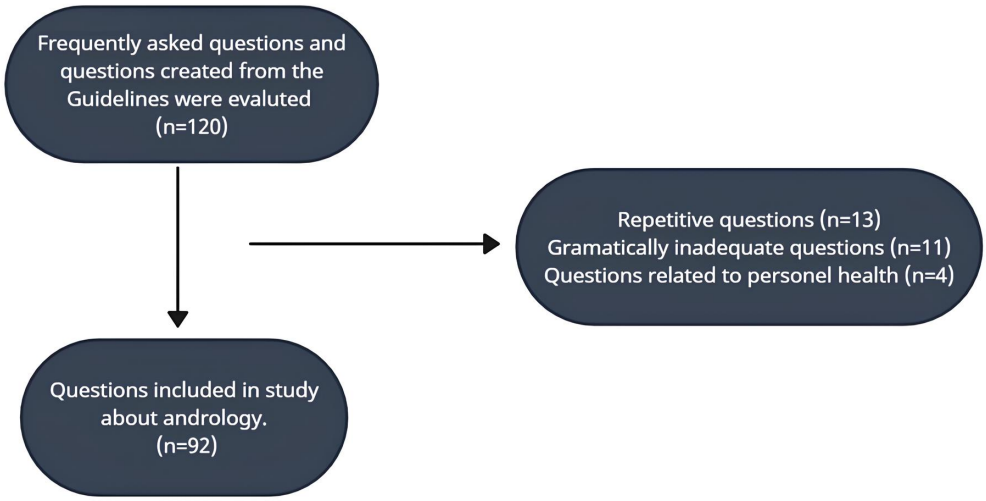

FIGURE 1. Study inclusion flow chart.

TABLE 1. Accuracy rates of ChatGPT answers according to andrology topics.

		Grade 1	Grade 2	Grade 3	Grade 4
Male hypogonadism (n = 9)	Version 3.5	6 (66.6%)	1 (11.1%)	2 (22.2%)	–
	Version 4.0	6 (66.6%)	2 (22.2%)	1 (11.1%)	
Erectile dysfunction and sexual desire disorder (n = 16)	Version 3.5	10 (62.5%)	5 (31.2%)	1 (6.25%)	–
	Version 4.0	12 (75.0%)	3 (18.75%)	1 (6.25%)	
Disorders of ejaculation (n = 9)	Version 3.5	3 (33.3%)	5 (55.5%)	1 (11.1%)	–
	Version 4.0	4 (44.4%)	4 (44.4%)	1 (11.1%)	
Penile curvature and penile size abnormalities (n = 20)	Version 3.5	8 (40.0%)	7 (35.0%)	4 (20.0%)	1 (5.0%)
	Version 4.0	9 (45.0%)	7 (35.0%)	3 (15.0%)	1 (5.0%)
Priapism (n = 18)	Version 3.5	11 (61.1%)	3 (16.6%)	2 (11.1%)	2 (11.1%)
	Version 4.0	11 (61.1%)	4 (22.2%)	2 (11.1%)	1 (5.5%)
Male infertility (n = 20)	Version 3.5	8 (40.0%)	10 (50.0%)	2 (10.0%)	–
	Version 4.0	9 (45.0%)	9 (45.0%)	2 (10.0%)	

Grade 1: completely correct, Grade 2: correct but insufficient, Grade 3: mix of correct and misleading information, Grade 4: completely incorrect answers.

TABLE 2. Accuracy rates of answers according to question sources.

		Grade 1	Grade 2	Grade 3	Grade 4
Male hypogonadism					
FAQ	Version 3.5	4 (80.0%)	1 (20.0%)	–	–
	Version 4.0	4 (80.0%)	1 (20.0%)		
EAU guidelines	Version 3.5	2 (50.0%)	–	2 (50.0%)	–
	Version 4.0	2 (50.0%)	1 (25.0%)	1 (25.0%)	
Erectile dysfunction and sexual desire disorder					
FAQ	Version 3.5	6 (60.0%)	3 (30.0%)	1 (10.0%)	–
	Version 4.0	8 (80.0%)	1 (10.0%)	1 (10.0%)	
EAU guidelines	Version 3.5	4 (66.6%)	2 (33.3%)	–	–
	Version 4.0	4 (66.6%)	2 (33.3%)		
Disorders of ejaculation					
FAQ	Version 3.5	3 (42.8%)	3 (42.8%)	1 (14.2%)	–
	Version 4.0	4 (57.1%)	2 (28.5%)	1 (14.2%)	
EAU guidelines	Version 3.5	–	2 (100.0%)	–	–
	Version 4.0		2 (100.0%)		
Penile curvature and penile size abnormalities					
FAQ	Version 3.5	5 (50.0%)	3 (30.0%)	2 (20.0%)	–
	Version 4.0	6 (60.0%)	3 (30.0%)	1 (10.0%)	
EAU guidelines	Version 3.5	3 (30.0%)	4 (40.0%)	2 (20.0%)	1 (10.0%)
	Version 4.0	3 (30.0%)	4 (40.0%)	2 (20.0%)	1 (10.0%)
Priapism					
FAQ	Version 3.5	4 (44.4%)	2 (22.2%)	2 (22.2%)	1 (11.1%)
	Version 4.0	4 (44.4%)	3 (33.3%)	2 (22.2%)	–
EAU guidelines	Version 3.5	7 (77.7%)	1 (11.1%)	–	1 (11.1%)
	Version 4.0	7 (77.7%)	1 (11.1%)		1 (11.1%)
Male infertility					
FAQ	Version 3.5	5 (38.4%)	7 (53.8%)	1 (7.6%)	–
	Version 4.0	6 (46.1%)	6 (46.1%)	1 (7.6%)	
EAU guidelines	Version 3.5	3 (42.8%)	3 (42.8%)	1 (14.2%)	–
	Version 4.0	3 (42.8%)	3 (42.8%)	1 (14.2%)	

EAU: European association of urology, FAQ: Frequently asked questions, Grade 1: completely correct, Grade 2: correct but insufficient, Grade 3: mix of correct and misleading information, Grade 4: completely incorrect answers.

categorized as easy, and 16 questions were at the medium level. There were no difficult questions among the frequently asked questions.

Examples of questions are shown in Table 3. Among the questions derived from the guidelines, 13 questions were easy, 15 questions were medium, and 12 were considered difficult.

The rates of correct answers to the questions (grades 1 and 2 combined) were 85.2% for frequently asked questions and 81.5% for questions obtained from the guidelines (88.9% for 4.0 version). The rate of completely incorrect answers (grade 4) for the questions was 1.8% in the group of frequently asked questions and 5.2% in the group of guideline-based questions (0% and 5.2% for 4.0 version). Despite the guideline questions being more challenging and detailed, the success rate in answering the questions was reasonably similar.

5. Discussion

Artificial intelligence (AI) is increasingly being used widely today. The ChatGPT program, which can be easily used by patients in their own language, is an evaluation by the OpenAI company. While the source of information is not specified, previous studies have highlighted its high success rates, particularly in providing accurate responses to health-related questions.

Numerous studies have previously been published, indicating the prevalence of controversial and inaccurate health information shared on social media platforms such as YouTube, Facebook, Twitter, Reddit and Pinterest [12]. For instance, in a study conducted by Alsyouf *et al*. [13] that assessed articles shared on social media regarding urological malignancies, it was found that the quantity of false and misleading articles was notably higher than the number of accurate articles. However,

TABLE 3. The most frequently asked questions.

Number	The questions
1	What are the types of hypogonadism?
2	How is hypogonadism treated?
3	What are the causes of erectile dysfunction?
4	How common is erectile dysfunction?
5	What are the differences between penile prosthesis types?
6	How long does the penile implant last?
7	Is it true that antidepressants reduce sexual desire?
8	What can I do to stop myself from ejaculating so quickly?
9	Is premature ejaculation a sign of a serious health problem?
10	What causes delayed ejaculation?
11	What Causes a Curved Penis (Peyronie)?
12	Which Patients Can Have Penile Curvature Surgery?
13	How is the Surgical Treatment of Penile Curvature?
14	Can prosthesis be used for penile curvature?
15	What are the Treatments for Acute Penile Curvature?
16	What is the normal penis length and thickness?
17	Is penis size important for women?
18	What symptoms occur in patients with priapism?
19	Can priapism cause permanent damage?
20	What medications can cause priapism?
21	When should couples who cannot have children consult a doctor?
22	What increases a man’s risk of infertility
23	What should I do to improve sperm quality?
24	What are normal sperm parameters?
25	How Do I Prepare for a Sperm Analysis?
26	No sperm was found after microTESE. Can I try mikroTESE again?
27	What is the rate of finding sperm in a repeat micro TESE procedure?
28	I have varicocele, is surgery necessary?
29	Does varicocele cause pain?
30	Do chemotherapy and radiotherapy administered during cancer treatment disrupt sperm production?

such negative findings have not been reported for systems utilizing Artificial Intelligence (AI). In a study by Bulck and Moons, it was observed that the patient information articles selected were summarized and presented at a level that patients could easily comprehend [14].

Our article represents the first study to assess the accuracy and reliability of ChatGPT's responses to andrology-related inquiries. In our study, we observed that ChatGPT provided responses rated as completely correct or correct but insufficient (grade 1 and grade 2) for 83.6% (89.8% for 4.0 version) of the andrology-related questions. Unlike other internet resources such as social media, ChatGPT acquires and presents comprehensive information from numerous sources. It is evident that the breadth of the database it utilizes contributes to the higher percentage of accurate responses to the questions.

Guidelines provide physicians with universally applicable and typically detailed information in their respective fields. It was initially thought that ChatGPT might face challenges in accessing such information, but in the study conducted by Cakır *et al*. [15], it was discovered that ChatGPT provided satisfactory answers to 80% of the questions based on the EAU urolithiasis guidelines. In our study, we observed that ChatGPT provided 83.6% correct answers (89.8% for 4.0 version). This indicates the high effectiveness of ChatGPT in addressing questions that require detailed information, such as those derived from strong recommendations in guidelines.

In the present study, we observed that the answer success rates for frequently asked questions and guideline-based questions were quite similar, a finding consistent with other studies [8]. This indicates that ChatGPT is adept at retrieving specific information with a high degree of success. However, it's worth noting that the Grade 2 response category (33.6%), which also includes missing information, is relatively substantial. As a result, we believe that ChatGPT may not be entirely sufficient for medical education on its own but could serve as a valuable supplementary tool in the learning process.

The program offers advantages in various aspects. We believe that such easy access to information will enhance the health literacy of society. When patients search for their symptoms on the internet, they often encounter topics related to malignancies, which can increase their anxiety levels. We anticipate that artificial intelligence methods can provide more accurate information, thus reducing anxiety levels.

Although the use and popularity of ChatGPT continue to rise, it is essential to be mindful of its potential drawbacks. For instance, in a study conducted by Rafeh *et al*. [16], it was found that ChatGPT increased the perceived urgency of the disease in 44% of the cases queried. This increase in urgency may lead to higher healthcare costs and heightened patient anxiety levels [16]. In another study by Hofmman *et al*. [17], it was determined that the responses provided by ChatGPT 3.5 and 4.0 versions to orthopedic board exam questions exhibited statistically significant differences in success rates. In their study, the 4.0 version's answers were found to be on par with those of a 3-year orthopedic specialist, whereas the 3.5 version's responses were comparable to those of an orthopedic intern [17].

An advantage of our study was that the paid version of ChatGPT 4.0 was used. We had the opportunity to evaluate both versions. The fact that versions 3.5 and 4.0 answered the questions obtained from the EAU 2023 guideline with the same success rate was due to the fact that both versions used the EAU 2023 guideline as a database. Another advantage of our study is that it addressed a deficiency seen in previous studies, which were predominantly conducted in English and lacked research in other languages. The fact that we obtained similar results to studies conducted with English questions indicates that we could readily access information with ChatGPT's high translation success.

However, a limitation of our study is the number of the evaluator urologists. It may result in observer fatigue bias.

6. Conclusions

In conclusion, ChatGPT offered satisfactory responses to frequently asked questions and questions derived from the EAU Guidelines. However, the category of responses with missing information was substantial. Therefore, we suggest that the implementation of ChatGPT under the supervision of urologists in urology clinics could enhance patients' understanding of andrology issues. Additionally, we believe it can serve as a valuable educational support tool for medical students.

AVAILABILITY OF DATA AND MATERIALS

The data presented in this study are available on reasonable request from the corresponding author.

AUTHOR CONTRIBUTIONS

İEE and AS—designed the research study. İEE—performed the research; analyzed the data. AS—wrote the manuscript. All authors read and approved the final manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Ethical approval was not obtained for this study because the use of ChatGPT and European Association of Urology (EAU) guideline information is publicly accessible.

ACKNOWLEDGMENT

Not applicable.

FUNDING

This research received no external funding.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] De Rose AF, Gallo F, Bini PM, Gattuccio I, Chiriacò V, Terrone C. Epidemiology of sexual disorders in general medical practice: an italian survey. Urologia. 2019; 86: 79–85.

[2] Eardley I. The incidence, prevalence, and natural history of erectile dysfunction. Sexual Medicine Reviews. 2013; 1: 3–16.

[3] Feldman HA, Goldstein I, Hatzichristou DG, Krane RJ, McKinlay JB. Impotence and its medical and psychosocial correlates: results of the massachusetts male aging study. Journal of Urology. 1994; 151: 54–61.

[4] Althof SE, McMahon CG, Waldinger MD, Serefoglu EC, Shindel AW, Adaikan PG, *et al*. An update of the international society of sexual medicine's guidelines for the diagnosis and treatment of premature ejaculation (PE). Sexual Medicine. 2014; 2: 60–90.

[5] Stuntz M, Perlaky A, des Vignes F, Kyriakides T, Glass D. The prevalence of peyronie's disease in the United States: a population-based study. PLOS ONE. 2016; 11: e0150157.

[6] Rowe PJ, Comhaire FH, Hargreave TB, Mahmoud AMA. WHO manual for the standardized investigation and diagnosis of the infertile couple. 1st edn. Cambridge University Press: Cambridge. 2000.

[7] OpenAI. ChatGPT: optimizing language models for dialogue. 2023. Available at: https://help.openai.com/en/articles/6825453-chatgpt-release-notes (Accessed: 03 August 2023).

[8] Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, *et al*. How does ChatGPT perform on the United States medical licensing examination (USMLE)? The implications of large language models for medical education and knowledge assessment. JMIR Medical Education. 2023; 9: e45312.

[9] Caglar U, Yildiz O, Meric A, Ayranci A, Gelmis M, Sarilar O, *et al*. Evaluating the performance of ChatGPT in answering questions related to pediatric urology. Journal of Pediatric Urology. 2024; 20: 26.e1–26.e5.

[10] Yeo YH, Samaan JS, Ng WH, Ting PS, Trivedi H, Vipani A, *et al*. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. Clinical and Molecular Hepatology. 2023; 29: 721–732.

[11] EAU Guidelines. Edn. Presented at the EAU Annual Congress Milan (Period: 10 March 2023–13 March 2023). 2023. Available at: https://uroweb.org/guidelines/sexual-and-reproductive-health (Accessed: 01 April 2023).

[12] Ergul A. Quality and reliability of YouTube Videos on surgical treatment of Uterine Leiomyomas. Cureus. 2021; 13: e20044.

[13] Alsyouf M, Stokes P, Hur D, Amasyali A, Ruckle H, Hu B. 'Fake News' in urology: evaluating the accuracy of articles shared on social media in genitourinary malignancies. BJU International. 2019; 124: 701–706.

[14] Moons P, Van Bulck L. Using ChatGPT and Google Bard to improve the readability of written patient information: a proof of concept. European Journal of Cardiovascular Nursing. 2024; 23: 122–126.

[15] Cakir H, Caglar U, Yildiz O, Meric A, Ayranci A, Ozgor F. Evaluating the performance of ChatGPT in answering questions related to urolithiasis. International Urology and Nephrology. 2024; 56: 17–21.

[16] Abi-Rafeh J, Hanna S, Bassiri-Tehrani B, Kazan R, Nahai F. Complications following facelift and neck lift: implementation and assessment of large language model and artificial intelligence (ChatGPT) performance across 16 simulated patient presentations. Aesthetic Plastic Surgery. 2023; 47: 2407–2414.

[17] Hofmann HL, Guerra GA, Le JL, Wong AM, Hofmann GH, Mayfield CK, *et al*. The rapid development of artificial intelligence: gpt-4's performance on orthopedic surgery board questions. Orthopedics. 2024; 47: e85–e89.

How to cite this article: İsmail Emre Ergin, Adem Sancı. Can ChatGPT help patients understand their andrological diseases? Revista Internacional de Andrología. 2024; 22(2): 14-20. doi: 10.22514/j.androl.2024.010.