Evaluation of Chat Generative Pretrained Transformer (ChatGPT) Performance in Answering Kidney Transplant Related Questions

Yunus Çolakoğlu[1], Ali Ayten[2], Çağlar Sertkaya[3], Kağan Toksal[3], Serdar Karadağ[3]

[1] Department of Urology, Bahcelievler Memorial Hospital, Istanbul, Türkiye

[2] Department of Urology, Gaziosmanpasa Training and Research Hospital, Istanbul, Türkiye

[3] Department of Urology, Bakirkoy Dr. Sadi Konuk Training and Research Hospital, Istanbul, Türkiye

Corresponding Author;
Yunus Çolakoğlu, MD
Address: Bahcelievler Memorial Hospital, Department of Urology, Istanbul, Türkiye.
E-mail: dr.yunusc@gmail.com

ORCID
Y.Ç.   0000-0001-6432-765X
A.A.   0000-0001-8441-1293
Ç.S.   0009-0009-0096-8453
K.T.   0009-0000-8691-2710
S.K.   0000-0002-1420-4536

Abstract

Objective: Social media such as (Youtube, Facebook, Instagram, Twitter, etc.) and artificial intelligence (AI) are applications that have become popular in recent years, they are the first resources that patients turn to today. ChatGPT is an AI-powered language model developed by OpenAI and its success on health problems are demonstrated by many studies. In this study, we aimed to evaluate the adequacy of ChatGPT's answers to questions about kidney transplantation.

Material and Methods: Frequently asked questions about kidney transplantation by patients on health forums, websites and social media (YouTube, Instagram, Twitter) were analyzed. We also analyzed the recommendation tables of the Kidney Transplantation section of the 2024 European Association of Urology (EAU) guidelines. Those with strong recommendations were translated into a question form. ChatGPT version 4o questions were asked and the answers were evaluated by 3 urologists experienced in kidney transplantation.

Results: Of the 126 questions evaluated, 65 questions were continued after the exclusion criteria. 57 (87.6%) of the answers were correct and adequate. According to EAU Guideline recommendations, 77 questions were prepared. 64 (83.1%) of the questions were answered completely correctly. There were no completely wrong answers in both frequently asked questions and questions adapted from the EAU Guidelines. Reproducibility of the questions was 100%.

Conclusion: Our study confirms that ChatGPT is a reliable source for kidney transplantation. We think that it will be a platform that both patients and their relatives and healthcare professionals can frequently refer to in the future.

Keywords: kidney transplantation, artificial intelligence, ChatGPT

INTRODUCTION

End-stage renal failure patients and kidney donors are worried, fearful and curious about kidney transplantation. They research their questions on the internet and social media before meeting with the transplant team (1). Social media (Youtube, Facebook, Instagram, Twitter, etc.) and artificial intelligence (AI) applications that have become popular in recent years are the first sources that come to mind in this regard (2)defined as interactive Web applications, have been on the rise globally, particularly among adults. The objective of this study was to investigate the trend of the literature related to the most used social network worldwide (i.e. Facebook, Twitter, LinkedIn, Snapchat, and Instagram.

ChatGPT is an AI-supported language model developed by OpenAI. It is based on a large text data set that allows to provide information on a wide range of topics and enable multilingual communication (3)primarily caused by non-urgent cases overwhelming the system, have spurred a critical necessity for innovative solutions that can effectively differentiate genuine emergencies from situations that could be managed through alternative means, such as using AI chatbots. This study aims to evaluate and compare the accuracy in differentiating between a medical emergency and a non-emergency of three of the most popular AI chatbots at the moment. Methods In this study, patient questions from the online forum r/AskDocs on Reddit were collected to determine whether their clinical cases were emergencies. A total of 176 questions were reviewed by the authors, with 75 deemed emergencies and 101 non-emergencies. These questions were then posed to AI chatbots, including ChatGPT, Google Bard, and Microsoft Bing AI, with their responses evaluated against each other and the authors' responses. A criteria-based system categorized the AI chatbot answers as \"yes,\" \"no,\" or \"cannot determine.\" The performance of each AI chatbot was compared in both emergency and non-emergency cases, and statistical analysis was conducted to assess the significance of differences in their performance. Results In general, AI chatbots considered around 12-15% more cases to be an emergency than reviewers, while they considered a very low number of cases as non-emergency compared to reviewers (around 35% fewer cases. The increasing use of ChatGPT has been tested on health issues and its success has been demonstrated by many studies (4–6)hospitals, and social media about prostate cancer and BPH were evaluated. Also, strong recommendation-level data

were noted in the recommendations tables of the European Urology Association (EAU).

Although it has been the subject of many studies in the medical field, ChatGPT has not been previously evaluated in kidney transplantation. In this study, we aimed to evaluate the adequacy of ChatGPT's answers to questions related to kidney transplantation.

MATERIAL AND METHODS

Patients' frequently asked questions about kidney transplantation on health forums, websites and social media (YouTube, Instagram, Twitter) were analyzed. Only questions in English were included in the study. We also analyzed the recommendation tables of the Kidney Transplantation section of the 2024 European Association of Urology (EAU) Guidelines (7). Those with a strong recommendation level were translated into a question form and categorized under the topic heading in the guideline. All questions were asked in English in ChatGPT version 4o. The answers generated by the AI were noted. All questions were asked twice at different times during the day to assess reproducibility of answers.

The answers were reviewed by 3 urologists experienced in kidney transplantation. The reviewers scored the answers compared to how they would have answered if asked this question by a patient. Responses were scored by each reviewer on a scale of 1-4.

4: Correct and adequate answer (no further information to add)

3: Correct answer but insufficient (more detailed explanation required)

2: Accurate and misleading information in one

1: Wrong or irrelevant answer

For questions where not all raters gave the same score, the median score was recorded. The agreement analysis between raters was also subjected to statistical analysis to assess the responses to the ChatGPT. Repeatability was defined as the consistency of the answers given to the same question at different times. Responses generated at different times were considered reproducible if they received the same score. Exclusion criteria were repetitive questions with similar

meanings, questions that did not comply with language rules, non-medical questions, cost-related questions, and questions about transplantation that were not considered ethical. Ethics committee approval was not required since patient data were not used in the study.

Statistical Analysis

Excel version 16.0 (Microsoft Corp.; Washington, USA) was used for statistical analyses. The scores of the responses were expressed as n (%). Reproducibility of responses was expressed as %. Inter-rater agreement was analyzed using K statistics. Landis and Koch's classification system was used to interpret Fleiss' Kappa coefficient: 0.0-0.20: Poor agreement, 0.21-0.40: Low agreement, 0.41-0.60: Moderate agreement, 0.61-0.80: High agreement, 0.81-1.00: Excellent agreement. The analysis was performed using R software. For this purpose, the categorical responses given by the evaluators for each evaluation topic were organized in a data matrix and Fleiss's Kappa coefficient was calculated using the irr package. The results of the analysis were interpreted to assess whether there was significant agreement between the raters.

RESULTS

The flowchart of the questions included in the study is shown in Figure 1. Of the 126 questions evaluated, 61 were excluded from the study after the exclusion criteria. Answers to 65 questions were included in the study (Table 1). Of the answers, 57 (87.6%) were correct and adequate, 7 (10.7%) were correct but inadequate, and 1 (1.5%) was a combination of correct and misleading information. No question was answered incorrectly.
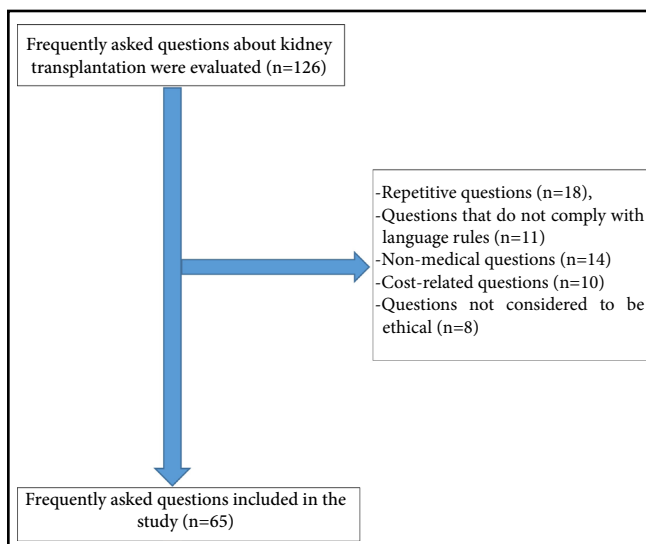
Figure 1. Flowchart of the questions in the study

According to EAU Guideline recommendations, 77 questions were prepared (Table 2). 64 (83.1%) of the questions were answered completely correctly. Nine (11.6%) questions received 3 points and 4 (5.1%) questions received 2 points. Similar to the frequently asked questions, there were no completely wrong answers in the guideline recommendations. Inter-rater agreement was generally good ($Ҟ = 0.84$), 95% CI: 0.65-0.93), with only 18 questions with inter-rater disagreement. Inter-rater agreement was excellent for all three ($Ҟ > 0.92$).

The reproducibility and similarity rate of the answers to the questions was 100% for both the frequently asked questions and the questions prepared according to the EAU Guideline recommendations.

Table 1. Frequently Asked Questions About Kidney Transplantation

		4 Points	3 Points	2 Points	1 Point
1.	What is a kidney transplant?	*			
2.	Who can get a kidney transplant?	*			
3.	Why is a kidney transplant necessary?	*			
4.	How is kidney transplant surgery performed?	*			
5.	Where to get a kidney for transplantation?	*			
6.	What is a living donor?	*			
7.	What is a cadaver donor?	*			
8.	Is it safe to be a donor?	*			
9.	Is there an age limit to become a donor?	*			

10.	Who can be a kidney donor?	*			
11.	Is kidney transplantation risky?	*			
12.	What are the risks of kidney transplant surgery?	*			
13.	What happens if the transplanted kidney is rejected?		*		
14.	How to prevent organ rejection after transplantation?	*			
15.	What are immunosuppressive drugs?	*			
16.	Do these drugs have side effects?	*			
17.	How long will I need to take medication after the transplant?	*			
18.	What will change in my lifestyle after transplantation?	*			
19.	When can I return to work after transplantation?	*			
20.	What should my diet be after kidney transplantation?	*			
21.	How long will I recover after transplantation?	*			
22.	How long does kidney transplant surgery take?	*			
23.	Is there pain after a kidney transplant?	*			
24.	Is there a risk of infection after transplantation?	*			
25.	How do I know if the kidney was rejected after transplantation?	*			
26.	Are regular check-ups necessary after kidney transplantation?	*			
27.	How many years will I live after transplantation?	*			
28.	How long does the kidney function after transplantation?		*		
29.	Can a kidney transplant recipient play sports?	*			
30.	Do smoking and alcohol affect kidney transplantation?	*			
31.	Is it possible to get pregnant after a kidney transplant?	*			
32.	How will my sex life be affected after transplantation?	*			
33.	Can a kidney transplant recipient have dental treatment?	*			
34.	Do allergic reactions affect kidney transplantation?	*			
35.	Can I travel after transplant?	*			
36.	What happens if the kidney transplant fails?	*			
37.	Can a kidney transplant be repeated?	*			
38.	Can a kidney transplant be performed in emergencies?		*		
39.	How important is tissue matching in kidney transplantation?		*		
40.	What happens if tissue compatibility is not achieved?	*			
41.	How long is the waiting period for a kidney transplant?		*		
42.	How to get on the waiting list?	*			
43.	What happens if a living donor cannot be found?	*			
44.	How is organ donation done?	*			
45.	What should I pay attention to after kidney transplantation?	*			
46.	Which vaccinations should I get after a kidney transplant?	*			
47.	In which cases should I consult a doctor after kidney transplantation?	*			
48.	Are chronic diseases affected after kidney transplantation?	*			

		4 Points	3 Points	2 Points	1 Point
49.	Does diabetes affect kidney transplantation?	*			
50.	Does hypertension affect kidney transplantation?	*			
51.	Can kidney transplantation be performed on children?	*			
52.	Can kidney transplantation be performed in the elderly?	*			
53.	Can HIV positive patients receive a kidney transplant?	*			
54.	Does blood group incompatibility prevent kidney transplantation?	*			
55.	Are medications used for life after kidney transplantation?	*			
56.	Which medications are not used after transplantation?		*		
57.	How much water should I drink after kidney transplantation?	*			
58.	Which foods should I avoid after transplantation?	*			
59.	Is psychological support important during kidney transplantation?	*			
60.	Does stress affect the kidney after kidney transplantation?	*			
61.	What about sleep patterns after kidney transplantation?		*		
62.	Can I work after a kidney transplant?	*			
63.	Will my kidneys recover completely after transplantation?			*	
64.	How is the immune system affected after a kidney transplant?	*			
65.	Can a person with cancer have a kidney transplant?	*			

Table 2. Questions Related to the European Association of Urology (EAU) Guideline Recommendations

Organ retrieval and transplantation surgery		4 Points	3 Points	2 Points	1 Point
1.	Which technique should be preferred for living donor nephrectomy?	*			
2.	Which technique can be used for living donor nephrectomy in centers where endoscopic methods are not accessible?	*			
3.	When can laparo-endoscopic single site (LESS) surgery, robotic and natural orifice transluminal endoscopic surgery-assisted (NOTES) living-donor nephrectomy be preferred?	*			
Organ preservation					
4.	In which solutions can a donor kidney be stored for cold storage?	*			
5.	Where should the donor kidney be stored if the University of Wisconsin or histidine tryptophan ketoglutarate preservation solutions for cold storage are not available?	*			
Methods of kidney preservation: static and dynamic reservation					
6.	Is the duration of ischemia important and how should it be?	*			
7.	What should be done to reduce delayed graft function in cadaveric donor kidneys?	*			
8.	Can hypothermic machine perfusion (HMP) be performed in cadaveric donor kidneys with standard criteria?	*			
9.	What should be the pressure level in HMP maintenance?		*		
10.	Should HMP be intermittent or continuous and is pressure or flow more important in HMP?			*	
Donor Kidney Biopsies					
11.	Is the decision to accept a donor kidney based only on histology? Are there other important parameters?	*			

12.	Should paraffin histology or frozen sections be used for histomorphology in donor kidney biopsy?	*			
13.	Who should evaluate procurement biopsies?			*	
Peri-operative antibiotics in renal transplant					
14.	Should perioperative antibiotic prophylaxis in kidney transplant recipients be in multiple doses or single doses?	*			
Specific fluid regimes during renal transplantation					
15.	How should pre-, intra- and postoperative hydration be adjusted to improve kidney transplant function?	*			
16.	How should intraoperative hydration be managed to reduce rates of delayed graft function and optimize early graft function?	*			
Surgical approaches for first, second, third and further transplants					
- Single kidney transplant – living and deceased donors					
17.	What should be done before starting immunosuppression and anesthesia for cadaveric kidney transplantation?	*			
18.	What should be considered in the donor and recipient arteries before starting arterial anastamosis?		*		
19.	How should the preoperative surgical approach be planned for third or subsequent transplants?	*			
Ureteric implantation in normal urinary trac					
20.	Which ureteral anastamosis should be preferred in kidney transplant recipients with normal urologic anatomy?	*			
21.	Which type of anastamosis can be used especially in very short or poorly vascularized transplant ureters?		*		
22.	Should a transplant ureteric stent be used and is it beneficial?	*			
23.	Is the surgical principle different in double ureters and how can anastamosis be performed?	*			
Donor complications					
24.	In which centers should living donor nephrectomies be performed?	*			
25.	How long should kidney donors be followed up?	*			
Arterial thrombosis					
26.	What should be done when graft thrombosis is suspected?	*			
27.	What should be done if ultrasonography shows poor graft perfusion?		*		
28.	Non viable greft varlığında ne yapılmalıdır?	*			
Venous thrombosis					
29.	What should be done when graft thrombosis is suspected?	*			
30.	Should pharmacologic prophylaxis be routinely used to prevent transplant renal vein thrombosis?		*		
Transplant renal artery stenosis					
31.	Which test should be performed primarily for the detection of arterial stenosis and what are the other diagnostic methods in case of doubt?	*			
32.	What should be the first-line treatment of arterial stenosis in the transplanted kidney?		*		
33.	What should be done in case of recent transplantation, multiple, long and narrow stenoses or failure of angioplasty?	*			

	Arteriovenous fistulae and pseudo-aneurysms after renal biopsy				
34.	Which test should be performed if arteriovenous fistula or pseudoaneurysm is suspected?	*			
35.	What should be the first-line treatment for symptomatic arteriovenous fistula or pseudoaneurysm?		*		
	Lymphocele				
36.	What is the primary treatment of large and symptomatic lymphocele?	*			
37.	What should be done if percutaneous treatments fail?	*			
	Urinary leak				
38.	How should urine leakage be managed in kidney transplantation?		*		
39.	What should be done when conservative treatment fails?	*			
	Ureteral stenosis				
40.	In the case of ureteral stricture, what should be done to diagnose stricture by both renal decompression and antegrade pyelogram?	*			
41.	How should strictures < 3 cm in length be managed?	*			
42.	What is the primary treatment for late recurrent and/or strictures longer than 3 cm?	*			
	Kidney stones				
43.	Should causes of urolithiasis be evaluated in a kidney transplant recipient?	*			
44.	How should stone-induced ureteral obstruction be treated?	*			
45.	Which treatment methods should be used for stones smaller than 15 mm?	*			
	Malignancy after renal transplantation				
46.	Should the presence of a transplant kidney in the pelvis and the possibility of subsequent transplants be considered when planning treatment for prostate cancer?	*			
47.	Which centers should kidney transplant patients with prostate cancer be referred to?	*			
	Matching of donors and recipients				
48.	Should ABO blood group and human leukocyte antigen (HLA) A, B, C and DR phenotypes be determined for all kidney transplant candidates?	*			
49.	Should the donor and recipient be tested for HLA DQ and can susceptible patients be tested for HLA DP?	*			
50.	Should comprehensive HLA testing be performed before transplantation?	*			
51.	Should cross-match testing be performed before each kidney and combined kidney/pancreas transplantation to prevent hyperacute rejection?	*			
	General immunosuppression after kidney transplantation				
52.	Which drugs should be used for first rejection prophylaxis?	*			
	Calcineurin inhibitors				
53.	Should calcineurin inhibitors be used in rejection prophylaxis?	*			
54.	Which drug should be preferred as a first-line calcineurin inhibitor?	*			
55.	What should be done to ensure appropriate dose adjustment of calcineurin inhibitors?	*			
	Mycophenolates				
56.	Should mycophenolate be administered as part of an initial immunosuppressive regimen?	*			

Steroids					
57.	Should steroid therapy be part of immunosuppression in the perioperative and early posttransplant period?	*			
Inhibitors of the mammalian target of rapamycin (m-TOR)					
58.	What should be the calcineurin inhibitor dose in combination regimen with m-TOR inhibitors to prevent nephrotoxicity?	*			
59.	Can m-TOR inhibitors be used in patients with proteinuria and poor renal function?	*			
60.	What should be done to ensure appropriate dose adjustment of sirolimus and everolimus?	*			
Immunological complications					
61.	How long should transplant recipients be monitored after transplantation for acute rejection?	*			
62.	What should be the follow-up to detect graft dysfunction during hospitalization?	*			
63.	How to exclude other causes of graft dysfunction in suspected acute rejection and which tests should be performed first?	*			
64.	According to which criteria should renal biopsy be graded in patients with suspected acute rejection episodes?	*			
65.	When can 'blind' steroid bolus therapy be administered when immunologic complications are suspected after kidney transplantation?			*	
66.	Should patients with acute rejection be tested for anti-HLA antibodies against grafts, and if so, what should be the timing?	*			
67.	How should patients be evaluated, especially in late rejection?	*			
Hyper-acute rejection					
68.	Is adequate ABO blood group and HLA matching in donors and recipients important in preventing hyperacute rejection?	*			
Treatment of T-cell mediated acute rejection					
69.	What should be the first-line treatment of T-cell mediated rejection?			*	
70.	Which agents can be used in severe or steroid-resistant T-cell mediated rejection?	*			
Treatment of antibody mediated rejection					
71.	Should treatment of antibody-mediated rejection include antibody elimination?	*			
Follow-up after transplantation					
72.	How often, how long in total and by whom should post-transplant follow-up be performed?		*		
73.	What advice should be given to patients during follow-up?	*			
74.	Which parameters should be evaluated during post-transplant follow-up? What should be done in case of abnormalities in these parameters?	*			
75.	In case of graft dysfunction, what tests should be performed to rule out obstruction and renal artery stenosis?	*			
76.	What should be done in patients on calcineurin inhibitor therapy and/or with interstitial fibrosis and tubular atrophy with histologic findings suggestive of calcineurin inhibitor toxicity (e.g. arteriolar hyalinosis, striated fibrosis)?	*			
77.	Should the treatment of diseases such as hypertension, diabetes, proteinuria, cardiac risk factors encountered in post-transplant follow-up be initiated appropriately according to current guidelines?	*			

DISCUSSION

Social media has come to the forefront as the place where people primarily turn to for information, especially in recent years (8). It has been shown in the literature that there is a lot of misinformation and misdirection, product marketing as well as accurate information accessible on YouTube, Instagram and TikTok. It is also noteworthy that people without medical training easily publish content on these platforms (9).

As AI has become popular in many areas of life, it is becoming more and more prominent in the field of health. ChatGPT is an AI model developed by OpenAI. Many studies have investigated to what extent ChatGPT accurately answers the questions that patients are curious about (6). Caglar et al. found that ChatGPT gave satisfactorily accurate answers in the field of andrology and benign prostatic hyperplasia (4,5). Samaan et al. demonstrated the program's superior success on questions related to bariatric surgery. In these studies, the model provided approximately 90% correct answers to the questions (10). Although there are many studies showing the success of ChatGPT on urological diseases, this deficiency continues in the literature on kidney transplantation. In our study, we tested the accuracy and reliability of ChatGPT in answering questions related to kidney transplantation.

ChatGPT answers questions with information based on previously published articles and books. This suggests that ChatGPT provides quality, accurate information more frequently than other social media platforms (11). In their new study, Mankowski et al. tested how ChatGPT can be used in kidney transplantation by comparing it with human participants. They posed 12 multiple-choice questions about kidney transplantation on the American Society of Nephrology fellowship exam to ChatGPT versions 3.5, 4, 4 Visual (4 V) and nephrology residents and nephrology fellowship program directors. According to the results of the study, the 4V version performed as well as nephrology residents and training program directors (Mankowski et al. 2024). This result shows that ChatGPT is a promising tool that can help experts in kidney transplantation(12). Our study showed that 87.6% of the answers given by ChatGPT were correct. The ability of AI software to access the literature and its capacity to continuously improve itself are among the important factors in the high rate of correct answers.

Our results showed that ChatGPT provided a high percentage of correct answers to questions adapted from the EAU Guidelines and frequently asked by patients. It was remarkable that it gave correct answers even to a text as dense and high quality information as the EAU Guidelines. Kung et al. demonstrated that the model can pass a serious exam such as the United States Medical Licensing Exam (USMLE) (13). In 2024, a meta-analysis of 45 studies also revealed the high success of ChatGPT in medical licensing exams. Another important result in the meta-analysis was that ChatGPT surpassed the average score of medical students (14).

Reproducibility is an issue to be considered in AI-supported programs. Yeo et al. showed that ChatGPT's answers to frequently asked questions about hepatocellular carcinoma were about 90% reproducible (15). High reproducibility was also observed in the answers to questions asked in the field of andrology. In addition, the answers were in an easy-to-understand language (5). Our results showed that ChatGPT's answers to questions related to kidney transplantation were reproducible.

The limitations of our study include the fact that ChatGPT has no experience in examining individual patients and therefore cannot determine subjective procedures related to patients, the questions asked may not cover all topics related to kidney transplantation, and the questions were asked only in English. Although the answers were evaluated by a team experienced in transplantation, it is obvious that some of the answers may contain differences on an individual basis. We tried to minimize these differences by working with more than one experienced expert.

CONCLUSION

Our study confirms that ChatGPT is a reliable and preferable resource for kidney transplantation. The evolving structure of AI can be used in patient consultations in the future, as well as becoming an auxiliary control mechanism for experts. With its ever-evolving structure, we think that it will be a platform that both patients and their relatives and healthcare professionals can frequently refer in the future.

Declaration of Interests: The authors have no conflict of interest to declare.

Ethics Committee Approval: Since no patient data was used in our study, ethics committee approval was not required. Informed Consent: Since no patient data was used in our study, informed consent was not required.

REFERENCES

1. Selen, T., & Merhametsiz, O. (2024). YouTube™ as a source of information on autosomal dominant polycystic kidney disease: A quality analysis. *Digital health*, *10*, 20552076241248109. https://doi.org/10.1177/20552076241248109

2. Zyoud SH, Sweileh WM, Awang R, Al-Jabi SW. Global trends in research related to social media in psychology: Mapping and bibliometric analysis. Int J Ment Health Syst [Internet]. 2018;12(1):1-8. Available from: https://doi.org/10.1186/s13033-018-0182-6

3. Zúñiga Salazar, G., Zúñiga, D., Vindel, C. L., et al. (2023). Efficacy of AI Chats to Determine an Emergency: A Comparison Between OpenAI's ChatGPT, Google Bard, and Microsoft Bing AI Chat. *Cureus*, *15*(9), e45473. https://doi.org/10.7759/cureus.45473

4. Caglar, U., Yildiz, O., Meric, A., et al. (2023). Evaluating the performance of ChatGPT in answering questions related to benign prostate hyperplasia and prostate cancer. *Minerva urology and nephrology*, *75*(6), 729-733. https://doi.org/10.23736/S2724-6051.23.05450-2

5. Caglar, U., Yildiz, O., Ozervarli, M. et al. (2023). Assessing the Performance of Chat Generative Pretrained Transformer (ChatGPT) in Answering Andrology-Related Questions. *Urology research & practice*, *49*(6), 365-369. https://doi.org/10.5152/tud.2023.23171

6. Secinaro S, Calandra D, Secinaro A, Muthurangu V, Biancone P. The role of artificial intelligence in healthcare: a structured literature review. BMC Med Inform Decis Mak [Internet]. 2021;21(1):1–23. Available from: https://doi.org/10.1186/s12911-021-01488-9

7. Faba OR, Boissier R, Budde K, et al. European Association of Urology Guidelines on Renal Transplantation: Update 2024. Eur Urol Focus.

8. Stagg BC, Gupta D, Ehrlich JR, et al. HHS Public Access. 2022;4(1):71-7.

9. Dubin, J. M., Aguiar, J. A., Lin, J. S., et al. (2024). The broad reach and inaccuracy of men's health information on social media: analysis of TikTok and Instagram. *International journal of impotence research*, *36*(3), 25-6260. https://doi.org/10.1038/s41443-022-00645-6

10. Samaan JS, Yeo YH, Rajeev N, et al. Assessing the Accuracy of Responses by the Language Model ChatGPT to Questions Regarding Bariatric Surgery. Obes Surg [Internet]. 2023;33(6):1790–6. Available from: https://doi.org/10.1007/s11695-023-06603-5

11. Antaki F, Touma S, Milad D, El-Khoury J, Duval R. Evaluating the Performance of ChatGPT in Ophthalmology: An Analysis of Its Successes and Shortcomings. Ophthalmol Sci [Internet]. 2023;3(4):100324. Available from: https://doi.org/10.1016/j.xops.2023.100324

12. Mankowski, M. A., Jaffe, I. S., Xu, J., et al. (2024). ChatGPT Solving Complex Kidney Transplant Cases: A Comparative Study With Human Respondents. *Clinical transplantation*, *38*(10), e15466. https://doi.org/10.1111/ctr.15466

13. Kung, T. H., Cheatham, M., Medenilla, A., et al. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS digital health*, *2*(2), e0000198. https://doi.org/10.1371/journal.pdig.0000198

14. Liu, M., Okuhara, T., Chang, X., et al. (2024). Performance of ChatGPT Across Different Versions in Medical Licensing Examinations Worldwide: Systematic Review and Meta-Analysis. *Journal of medical Internet research*, *26*, e60807. https://doi.org/10.2196/60807

15. Yeo, Y. H., Samaan, J. S., Ng, W. H., et al. (2023). Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. *Clinical and molecular hepatology*, *29*(3), 721-732. https://doi.org/10.3350/cmh.2023.0089