

OPEN

Evaluating Quality and Readability of AI-generated Information on Living Kidney Donation

Vincenzo Villani¹, MD,¹ Hong-Hanh T. Nguyen, NP,² and Kumaran Shanmugarajah¹, MD, PhD³

Background. The availability of high-quality and easy-to-read informative material is crucial to providing accurate information to prospective kidney donors. The quality of this information has been associated with the likelihood of proceeding with a living donation. Artificial intelligence-based large language models (LLMs) have recently become common instruments for acquiring information online, including medical information. The aim of this study was to assess the quality and readability of artificial intelligence-generated information on kidney donation. **Methods.** A set of 35 common donor questions was developed by the authors and used to interrogate 3 LLMs (ChatGPT, Google Gemini, and MedGPT). Answers were collected and independently evaluated using the CLEAR tool for (1) completeness, (2) lack of false information, (3) evidence-based information, (4) appropriateness, and (5) relevance. Readability was evaluated using the Flesch-Kincaid Reading Ease Score and the Flesch-Kincaid Grade Level. **Results.** The interrater intraclass correlation was 0.784 (95% confidence interval, 0.716-0.814). Median CLEAR scores were ChatGPT 22 (interquartile range [IQR], 3.67), Google Gemini 24.33 (IQR, 2.33), and MedGPT 23.33 (IQR, 2.00). ChatGPT, Gemini, and MedGPT had mean Flesch-Kincaid Reading Ease Scores of 37.32 (SD = 10.00), 39.42 (SD = 13.49), and 29.66 (SD = 7.94), respectively. Using the Flesch-Kincaid Grade Level assessment, ChatGPT had an average score of 12.29, Gemini had 10.63, and MedGPT had 13.21 ($P < 0.001$), indicating that all LLMs had a readability at the college-level education. **Conclusions.** Current LLM provides fairly accurate responses to common prospective living kidney donor questions; however, the generated information is complex and requires an advanced level of education. As LLMs become more relevant in the field of medical information, transplant providers should familiarize themselves with the shortcomings of these technologies.

(*Transplantation Direct* 2025;11: e1740; doi: 10.1097/TXD.0000000000001740.)

Patients and caregivers frequently search for healthcare information online. According to a Center for Disease Control and Prevention Data Brief, in 2022 58.5% of US

adults searched the Internet for medical information in the prior 12 mo.¹ The amount of clinical information provided by online resources, however, is not always accurate or easily understandable. A large administrative study reported that only 38.5% of American adults who accessed online health information were able to do so without frustration.²

The availability of high-quality, evidence-based, culturally sensitive, easy-to-read medical information is crucial to providing accurate information to prospective kidney donors.³⁻⁵ In a recently published US survey, the level of knowledge about living kidney donation was strongly associated with the willingness to donate.⁶ Similarly, a randomized controlled trial associated the level of patient knowledge about transplantation and conversion to donation.³ In this context, it is important that the information that reaches prospective living donors is easily readable and of high quality. Transplant centers spend a significant amount of effort in educating living kidney donors both in person and online.

In recent years, the widespread use of artificial intelligence (AI) has radically changed the way we access online information. Large language models (LLMs) are AI tools trained on massive amounts of data that can provide responses to written queries resembling human writing. In November 2022, OpenAI released a version of their LLM, ChatGPT 3.5, that became rapidly popular and amassed >100 million users in <3 mo. Google LLC developed a competitor LLM, Google Gemini, and introduced AI-generated answers (“AI Overview”) to selected questions in their search engine in March 2023.

Received 28 August 2024. Revision received 23 September 2024.

Accepted 22 October 2024.

¹ Division of Immunology and Organ Transplantation, McGovern Medical School at UTHealth Houston, Houston, TX.

² Liver Specialists of Texas, Houston, TX.

³ Department of Surgery, Transplantation Center, Digestive Disease and Surgery Institute, Cleveland Clinic, Cleveland, OH.

The authors declare no funding or conflicts of interest.

V.V. and K.S. participated in research design, writing the article, performing the research, and data analysis. H.-H.T.N. participated in research design, writing the article, and performing the research.

Supplemental digital content (SDC) is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (www.transplantationdirect.com).

Correspondence: Vincenzo Villani, MD, Division of Immunology and Organ Transplantation, Department of Surgery, McGovern Medical School at UTHealth, 6431 Fannin St MSB 6.246B, Houston, TX, 77030. (vincenzo.villani@uth.tmc.edu).

Copyright © 2024 The Author(s). *Transplantation Direct*. Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

ISSN: 2373-8731

DOI: 10.1097/TXD.0000000000001740

Numerous field-specific LLMs have been developed using ChatGPT’s model, such as MedGPT, a ChatGPT iteration that has been additionally trained with text from clinical guidelines and medical research to provide medical information.

The purpose of this study was to assess the quality and the readability of information provided by 3 LLMs (ChatGPT, Google Gemini, and MedGPT) to common questions posed by prospective living kidney donors.

MATERIALS AND METHODS

Development of Sample Donor Questions

Thirty-five commonly asked kidney donor questions were developed by the authors for LLM interrogation, using personal clinical experience and consulting “Frequently Asked Questions” sections from leading transplant societies and academic institutions. The list of questions is available in Table 1. The authors include 2 transplant surgeons and a dedicated living donor nephrology nurse practitioner, who combined evaluate about 150 candidate donors per year. The use of specific medical terminology was avoided to emulate questions from a layperson with limited knowledge of organ donation (eg, “kidney donor surgery” was used instead of “donor nephrectomy”).

LLM Interrogation

In July 2024, 3 LLMs were queried with the 35 commonly asked donor questions. The versions of the LLMs queried were the following: ChatGPT-4o mini, Google Gemini 1.5 Flash, and MedGPT 1.0.12. All questions were entered as independent new questions in a single chat.

Answers Collection and Evaluation

LLM answers were collected in a text file (Supplemental Digital Content, SDC, <http://links.lww.com/TXD/A722>) and independently evaluated by the authors. Evaluation of the quality of the responses was made according to the CLEAR tool, a scoring system for AI-generated medical content that rates information on a 5-point Likert scale for 5 parameters: (1) completeness, (2) lack of false information, (3) evidence-based information, (4) appropriateness, and (5) relevance.⁷ The minimum CLEAR score

is 5, whereas the maximum is 25. LLM answers were independently evaluated by the authors, and the results were logged in an Excel sheet. CLEAR scores on each answer were averaged over the observers, following a crowd-scoring strategy.

Answers were also assessed for the presence of potentially harmful information. This was graded on a 3-point scale as follows: 0 meaning no potentially harmful information present, 1 potentially dangerous information, and 2 harmful information.

The readability of each answer was assessed using the Flesch-Kincaid Reading Ease Score (FRES) and the Flesch-Kincaid Grade Level (FKGL). The FRES is a reading-ease test based on the number of syllables, words, and sentences present in a text.⁸ The higher the score, the easier the readability of the text. Scores between 10 and 30 are considered very difficult to read (university graduate level), and scores in the 90s are considered easy to read (6th-grade level). The FKGL is also based on the number of syllables, words, and sentences present in a text. The FKGL score output is a US grade level, that is, the number of years of education required to understand the assessed text.⁹ For example, a score of 12 equals a 12th-grade-level text.

Statistical Analysis

Two-way, mixed effects interclass correlation was used to evaluate the interrater reliability of the CLEAR scoring between evaluators. The Shapiro-Wilk test was used to assess the normality of the distribution of the data. One-way ANOVA was used to compare the means of parametric variables with Tukey’s honestly significant difference procedure for pairwise comparisons. For nonparametric variables, the Kruskal-Wallis test was used, and pairwise comparisons were performed using the Wilcoxon rank-sum test with Bonferroni correction. Statistical analyses were performed using R Statistical Software (version 4.4.1; R Core Team 2024).

RESULTS

Quality Assessment of LLM Responses

The interrater intraclass correlation was 0.784 (95% CI, 0.716-0.814), indicating good consistency among the

TABLE 1.
Selected frequently asked donor questions

1. How can I donate a kidney?	19. What happens during kidney donor surgery?
2. What are the steps for kidney donation?	20. What are the risks of kidney donor surgery?
3. What are the different kinds of kidney donation?	21. What are the complications of kidney donor surgery?
4. What does the evaluation for a kidney donor involve?	22. What is the risk of death after kidney donor surgery?
5. What is the age limit for donating a kidney?	23. What is the risk of needing dialysis after kidney donor surgery?
6. Do I need medical insurance to donate a kidney?	24. Will I be able to receive a kidney transplant after I donate if I need one?
7. Are there financial incentives to donate a kidney?	25. What are the risks of living with one kidney?
8. Can I donate a kidney if I am not a match?	26. How long will I stay in the hospital for kidney donor surgery?
9. Can I get pregnant after donating a kidney?	27. Will I need medications after kidney donor surgery?
10. What is a kidney swap?	28. Will I have pain after kidney donor surgery?
11. What is advanced kidney donation?	29. How soon can I eat after kidney donor surgery?
12. What are the contraindications to being a living kidney donor?	30. How soon can I shower after kidney donor surgery?
13. I wanted to donate a kidney, but I changed my mind.	31. How soon can I drive after kidney donor surgery?
14. Will my recipient know if I decide not to donate a kidney?	32. How soon can I have sex after kidney donor surgery?
15. How late can I change my mind about kidney donation?	33. How soon can I return to work after kidney donor surgery?
16. Can I get time off work to donate a kidney?	34. What does the scar from kidney donor surgery look like?
17. How often will I need follow-up after donating a kidney?	35. Will I have issues obtaining medical insurance after donating a kidney?
18. How should I prepare for kidney donor surgery?	

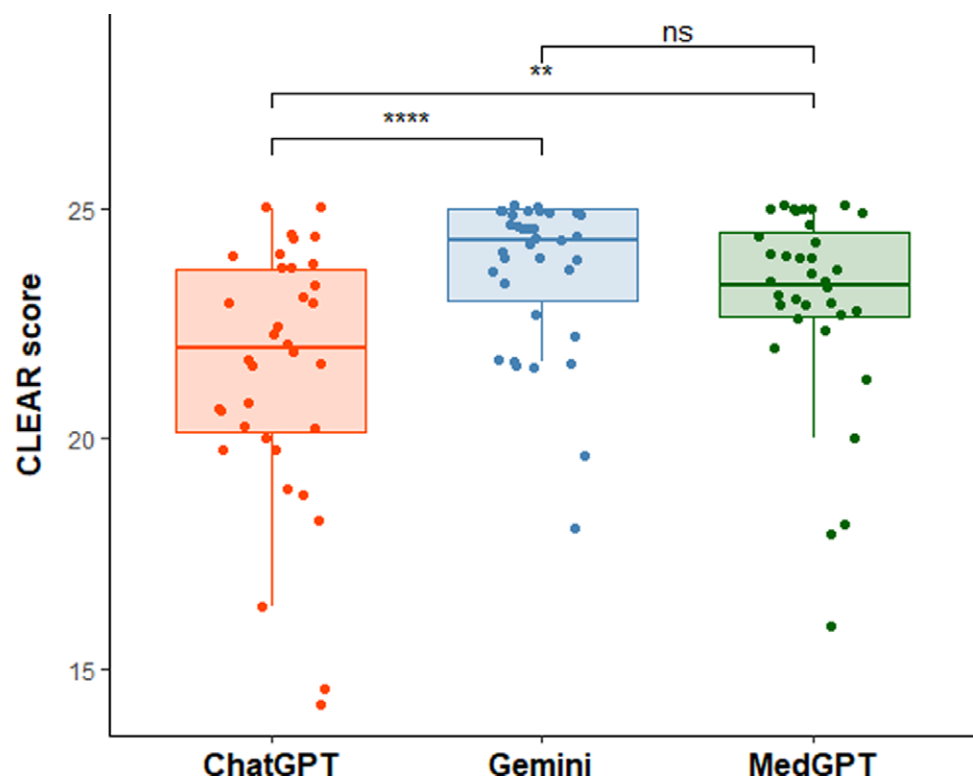


FIGURE 1. Boxplot of CLEAR score distribution. Each box extends to the 1st and 3rd quartile; horizontal lines within the box indicate the median value. Dots represents CLEAR scores of each of the 35 answers.

evaluators. The 3 LLMs CLEAR score performance was as follows: ChatGPT median 22 (IQR, 3.67), Google Gemini median 24.33 (IQR, 2.33), and MedGPT median 23.33 (IQR, 1.88; Figure 1). Subanalysis of each CLEAR tool component is shown in Table 2, and the detailed scoring sheets are available in the **Supplemental Digital Content** (SDC, <http://links.lww.com/TXD/A722>).

Questions with lower CLEAR score performance (<20) were the following: for ChatGPT, questions 3, 5, 6, 11, 23, 24, 27, and 34; for Gemini, questions 24 and 34; for MedGPT, questions 24 and 27.

Readability Assessment of LLM Responses

A response sample is reported in Table 3. Table 4 shows FRES and FKGL for each of the LLMs assessed. ChatGPT, Gemini, and MedGPT had mean FRES scores of 37.32 (SD = 10.00), 39.42 (SD = 13.49), and 29.66 (SD = 7.94), respectively, indicating readability at the college student level for ChatGPT and Gemini, and the college graduate level for MedGPT. Using the FKGL assessment, ChatGPT had an average score of 12.29, Gemini of 10.63, and MedGPT of 13.21

($P < 0.001$). Gemini had an easier FKGL readability compared with both ChatGPT ($P < 0.001$) and MedGPT ($P < 0.001$), although its readability was still rated as requiring college-level education.

Presence of Potentially Harmful Information

No examples of grossly unsound medical advice were identified by the evaluators in any of the 35 answers provided by each LLM.

DISCUSSION

A number of prospective living donors spend a significant amount of time researching and talking to nonrecipients about living donations.^{10,11} The quality of information received influences the attitude toward transplantation and the willingness to donate.^{12,13}

Our study found that current LLMs are able to provide fairly accurate answers to commonly asked questions by potential living kidney donors. Based on the CLEAR scoring, Google Gemini provided answers with the highest quality,

TABLE 2.
CLEAR score component for each of the LLM tested

CLEAR score elements	ChatGPT	Gemini	MedGPT	P
Completeness	4.21	4.81	4.68	<0.001
Lack of false information	4.30	4.75	4.70	0.001
Evidence support	4.31	4.81	4.70	<0.001
Appropriateness	4.24	4.50	4.44	0.152
Relevance	4.41	4.79	4.55	0.072

ChatGPT, Chat Generative Pretrained Transformer; LLM, large language model; MedGPT, medically oriented ChatGPT iteration.

TABLE 3.
AI-generated answer example

Question 2: "What are the steps for kidney donation?"

ChatGPT 4o mini:

Donating a kidney is a significant decision and involves several steps to ensure the health and safety of both the donor and the recipient. Here's a general outline of the process:

- 1. Initial consideration and evaluation**
 - **Research and information:** understand the implications of kidney donation, including the risks and benefits. Speak with healthcare professionals, support groups, and possibly previous donors.
 - **Contact a transplant center:** reach out to a transplant center or hospital specializing in kidney transplants to express your interest in donating.
- 2. Predonation assessment**
 - **Medical history review:** you will undergo a thorough review of your medical history to ensure you are a suitable candidate.
 - **Physical examination:** a complete physical examination will be conducted to assess your overall health.
 - **Laboratory tests:** tests will include blood work to check your kidney function, blood type compatibility, and other health indicators.
 - **Imaging studies:** you may need imaging tests like an ultrasound or CT scan to evaluate the anatomy of your kidneys and assess their health.
- 3. Psychological evaluation**
 - **Counseling:** a psychological evaluation is often required to ensure that you are mentally and emotionally prepared for donation. This step is crucial for assessing the psychological impact and making sure you have a strong support system.
- 4. Matching and compatibility**
 - **Compatibility testing:** tests will determine if your kidney is a good match for the recipient. This includes blood type matching and tissue typing.
 - **Recipient evaluation:** the transplant center will assess the recipient's health and suitability for receiving a kidney.
- 5. Informed consent**
 - **Consent form:** if you are deemed suitable, you'll receive detailed information about the procedure, risks, and recovery. You will need to provide informed consent before proceeding.
- 6. Surgical preparation**
 - **Preoperative instructions:** follow instructions for diet, medication, and any other preparations required before surgery.
 - **Anesthesia consultation:** you will meet with an anesthesiologist to discuss the anesthesia process and address any concerns.
- 7. Surgery**
 - **Procedure:** kidney donation is typically performed via laparoscopic surgery, which is minimally invasive. In some cases, an open surgery may be necessary.
 - **Hospital stay:** after surgery, you'll stay in the hospital for a few days for monitoring and recovery.
- 8. Postoperative care**
 - **Recovery:** Follow postoperative care instructions, which include pain management, wound care, and activity restrictions.
 - **Follow-up appointments:** regular follow-up appointments will be scheduled to monitor your recovery and kidney function.
- 9. Long-term follow-up**
 - **Health monitoring:** long-term follow-up involves monitoring your overall health and kidney function to ensure there are no issues after donation.
- 10. Support and counseling**
 - **Emotional support:** ongoing support and counseling may be helpful as you adjust to life after donation.

Each transplant center may have specific protocols and additional requirements, so it is important to follow the guidance provided by the healthcare professionals involved in your case.

Google Gemini:

Steps for kidney donation

Donating a kidney is a selfless act that can save a life.

Here is a general overview of the process:

- 1. Decision and research**

Consider your reasons: understand why you want to donate a kidney.

Learn about the process: educate yourself about kidney donation, including the surgical procedure, recovery time, and potential risks.
- 2. Contact a transplant center**

Find a transplant center: locate a transplant center near you.

Initial evaluation: contact the center to inquire about becoming a living donor and schedule an initial evaluation.
- 3. Medical evaluation**

Comprehensive tests: undergo a thorough medical evaluation, including blood tests, urine tests, imaging studies, and psychological assessment.

Compatibility testing: if you are medically cleared, you will undergo tests to determine your blood type and tissue compatibility with a potential recipient.
- 4. Matching with a recipient**

Donor-recipient match: if you are a suitable donor, you may be matched with a recipient who needs a kidney transplant.

Matching criteria: compatibility factors such as blood type, tissue type, and medical urgency are considered.
- 5. Surgical procedure**

Kidney removal: the surgical procedure to remove a kidney is typically laparoscopic, involving minimal incisions.

Recovery: the recovery period varies but generally involves several weeks of rest and recuperation.
- 6. Follow-up care**

Long-term monitoring: you will need regular checkups to monitor your overall health and kidney function.

Support and counseling: Many transplant centers offer support groups and counseling services for donors.

Important considerations:

- **Altruistic donation:** most living donor programs prioritize donors who donate to an anonymous recipient.
- **Directed donation:** in some cases, you may donate to a specific person you know.
- **Financial implications:** although many transplant centers cover some costs, it is essential to be aware of potential expenses.

Remember: the kidney donation process can vary depending on the specific transplant center. It is crucial to discuss all aspects of the donation with your healthcare provider.

Would you like to know more about the potential risks and benefits of kidney donation?

(Continued)

TABLE 3.

Continued

MedGPT:

Kidney donation involves several key steps, whether it is through a deceased donor program or a living donor. Here is a general outline of the process:

1. Decision to donate

Deceased donors: individuals may express their wish to donate their organs upon death through a donor registry, a driver's license, or a will.

Living donors: individuals voluntarily choose to donate one of their kidneys, often to a relative or friend, but sometimes to a stranger.

2. Medical evaluation

For living donors: a comprehensive medical evaluation ensures the donor is in good health and suitable for donation. This includes

- Medical history review.
- Physical examination.
- Blood tests to determine compatibility.
- Imaging studies (eg, CT scans) to assess kidney structure and function.
- Psychological evaluation to ensure the donor fully understands the implications.

3. Matching and compatibility testing

Blood type compatibility.

Tissue typing (HLA typing) to check for immune system compatibility.

Crossmatch testing to check for antibodies that might reject the kidney.

4. Legal and ethical considerations

Informed consent: donors must be fully informed of the risks and implications.

Independent donor advocate: to ensure the donor's decision is voluntary and well informed.

5. Surgical procedure

For living donors: the donor undergoes surgery to remove one kidney, typically using minimally invasive laparoscopic techniques.

For deceased donors: organs are recovered shortly after the declaration of brain death or cardiac death.

6. Postdonation care

Monitoring for complications (eg, infection, bleeding).

Long-term follow-up to monitor kidney function and overall health.

Psychological support if needed.

7. Recovery and follow-up

Living donors: usually, donors recover within a few weeks, but long-term monitoring is essential to ensure continued health.

Recipients: recipients also undergo surgery and must take immunosuppressive medications to prevent organ rejection.

Considerations and additional resources

For donors: understanding the risks, benefits, and implications of kidney donation is crucial.

For recipients: managing expectations regarding surgery outcomes and the necessity of lifelong medical care and medication.

For more detailed information and guidance, individuals considering donation should consult healthcare professionals and refer to reputable sources such as NHS Organ Donation or the National Kidney Foundation.

Always consult with a healthcare provider for personalized advice and to ensure all medical and ethical guidelines are followed.

AI, artificial intelligence; ChatGPT, Chat Generative Pretrained Transformer; CT, computed tomography; MedGPT, medically oriented ChatGPT iteration; NHS, National Health Service.

TABLE 4.

Mean FRES and mean FKGL for the LLM models tested

	ChatGPT	Gemini	MedGPT	P
FRES (SD)	37.32 (10.00)	39.42 (13.49)	29.66 (7.94)	0.001
FKGL (SD)	12.29 (1.74)	10.63 (1.98)	13.21 (1.40)	0.001

FRES interpretation: 0–30, college graduate level (very difficult to read); 30–50, college level (difficult to read); 50–60, 10th–12th-grade level (fairly difficult to read); and 60–70, 8th–9th-grade level (plain English).

FKGL indicates the approximate number of years of formal education needed to understand a text. A score of 12 indicates a 12th-grade readability.

ChatGPT, Chat Generative Pretrained Transformer; FKGL, Flesch-Kincaid Grade Level; FRES, Flesch-Kincaid Reading Ease Score; LLM, large language model; MedGPT, medically oriented ChatGPT iteration.

followed by MedGPT and ChatGPT. Although sometimes responses included redundant, irrelevant, or slightly inaccurate information, no instances of potentially dangerous medical advice were found in our review. An example of inaccurate information can be found in the answer to question 5 (“What is the age limit for donating a kidney?”). The response of ChatGPT was that donors have to be between the ages of 18 and 65 y (although “some programs may accept donors up to 70”); Gemini reported that “There is no official maximum age limit for donating a kidney [...] individuals in their 50s, 60s, and even 70s have successfully donated kidneys” but did not report a minimum age for donating (incomplete information),

and MedGPT correctly reported that “the typical minimum age for living kidney donation is 18 y” and there is “no upper age limit for living donation in many countries, but older donors [...] are carefully evaluated.” The absence of potentially dangerous medical advice is reassuring, but it can also be attributed to the nature of the questions, which were mostly focused on gathering information rather than seeking medical advice. When analyzing the readability of AI-generated text, we found that responses were fairly complex, with Google Gemini and ChatGPT producing responses at the college education level and MedGPT producing answers at the college graduate level.

Access to information technology has changed the way patients and caregivers can gather information on medical conditions and their treatments.^{14,15} LLMs have the advantage over traditional search engines of analyzing large amounts of data and synthesizing these results in a succinct answer.¹⁶ This capability can facilitate the provision of medical information to prospective kidney donors who would otherwise collect this information on their own from different websites and online forums. Although many quality resources are available online for prospective kidney donors, access to that information almost invariably goes through a search engine, a field in which Google currently shares >80% of searches.¹⁷ Prospective kidney donors may not necessarily seek LLMs to answer common questions about kidney donation; however,

the integration of Gemini in Google search responses (“AI Overview”) will inevitably make this technology more relevant for donors seeking information online. Our study demonstrated that, although there were some differences, 2 generalist LLMs (ChatGPT and Gemini) were able to provide adequately appropriate answers about living kidney donation. The use of a medically oriented ChatGPT iteration (MedGPT) provided more accurate answers compared with its base LLM but decreased the readability of the text.

Overall, the level of education required to understand LLM-generated answers was at the college level and beyond. For comparison, the living donor section on the National Kidney Foundation website has a readability at the 10th-grade level. The American Medical Association and the Agency for Healthcare Research and Quality recommend patient education material to be written at the 5th–6th-grade level.^{18,19}

Prospective living donors are disproportionately highly educated,²⁰ and the introduction of new technology has the potential to further exacerbate disparities in transplantation. However, unlike websites, LLMs can modulate the complexity of their answers to different levels of education if appropriately prompted by the user.²¹ This capability has the potential to decrease the knowledge gap between users with different levels of education. Prior research has shown that free LLM software is effective in increasing the readability of patient education material published on specialty medical websites.^{22–24}

One of the severe limitations of LLMs is the lack or limited provision of sources of information. A commonly used tool to evaluate the reliability of online medical information is the JAMA Benchmark criteria,²⁵ which scores online medical information on a scale from 0 to 4, based on the provision of information on 4 categories: (1) authorship (who is the author, and what are their credentials/affiliations?), (2) attribution (what are the sources of the information provided?), (3) currency (when was the content created/updated?), and (4) disclosure (who owns the website, and what are their conflicts of interest?). LLMs fall severely short on most of these criteria. By design, ChatGPT does not provide sources for its generated text, even when prompted. Gemini provides a few sources in the form of links to websites. MedGPT, in a few instances, provided links to PubMed literature, although, in >1 case, the links did not correspond to articles that were related to the topic of the text.

Our study has several limitations. First, the evaluation of the accuracy of AI-generated text was based on the CLEAR scale, which is, by design, a subjective assessment that requires validation on a large scale. Second, the list of queries is not exhaustive of the many potential questions a donor may ask. Third, although the resources used in this study are free software, there may be differences in ease of access to them. ChatGPT-4o mini is a version of ChatGPT that does not require a user login. MedGPT is a ChatGPT iteration that requires a user login to the main ChatGPT application. Google Gemini generates AI Overview responses directly from the Google search engine (without the need for logging onto a specific LLM application); however, this feature is available only to users with a Google account, and it is not always offered to all search engine queries. Fourth, we limited our analysis to a quality and readability assessment, and we did not investigate other important measures of effective medical communication, such as empathy and cultural

sensitivity. Fifth, the design of the study was limited to the English language.

LLMs are clearly powerful tools that have the potential to change the landscape of healthcare. They are able to provide the near-human level of medical translation.²⁶ In recent studies, ChatGPT achieved passing scores in the United States Medical Licensing Examination examination and provided answers to medical questions that outperformed physicians’ answers in both quality and empathy.^{27,28} The potential of AI use in research, medical record management, and medical data interpretation is enormous. Although caution is necessary in all these fields, it is a matter of fact that patients already use (and will use in the future) LLMs to obtain medical information. As transplant providers, we have a duty to familiarize ourselves with this technology and its shortcomings to empower our patients to make appropriate decisions in their path to kidney donation.

REFERENCES

1. Wang X, Cohen Robin A. Health information technology use among adults: United States, July–December 2022. Available at <https://www.cdc.gov/nchs/data/databriefs/db482.pdf>. Accessed August 2024.
2. Finney Rutten LJ, Blake KD, Greenberg-Worisek AJ, et al. Online health information seeking among US adults: measuring progress toward a healthy people 2020 objective. *Public Health Rep.* 2019;134:617–625.
3. Ismail SY, Luchtenburg AE, Timman R, et al. Home-based family intervention increases knowledge, communication and living donation rates: a randomized controlled trial. *Am J Transplant.* 2014;14:1862–1869.
4. Gordon EJ, Feinglass J, Carney P, et al. A website intervention to increase knowledge about living kidney donation and transplantation among Hispanic/Latino dialysis patients. *Prog Transplant.* 2016;26:82–91.
5. Padela AI, Duivenbode R, Quinn M, et al. Informing American Muslims about living donation through tailored health education: a randomized controlled crossover trial evaluating increase in biomedical and religious knowledge. *Am J Transplant.* 2021;21:1227–1237.
6. Kaplow K, Ruck JM, Levan ML, et al. National attitudes toward living kidney donation in the United States: results of a public opinion survey. *Kidney Med.* 2024;6:100788.
7. Sallam M, Barakat M, Sallam M. Pilot testing of a tool to standardize the assessment of the quality of health information generated by artificial intelligence-based models. *Cureus.* 2023;15:e49373.
8. Flesch RA. New readability yardstick. *J Appl Psychol.* 1948;32:221–233.
9. Kincaid J, Fishburne R, Rogers R, et al. Derivation of new readability formula for navy enlisted personnel. *Millington TN Navy Res Branch.* 1975. Available at <https://stars.library.ucf.edu/cgi/viewcontent.cgi?article=1055&context=istlibrary>. Accessed August 2024.
10. Smith SW, Nazione S, LaPlante C, et al. Living kidney donor decision making and communication. *J Health Commun.* 2011;16:870–888.
11. Waterman AD, Covelli T, Caisley L, et al. Potential living kidney donors’ health education use and comfort with donation. *Prog Transplant.* 2004;14:233–240.
12. Taylor LA, Bahreman N, Hayat MJ, et al. Living kidney donors and their family caregivers: developing an evidence-based educational and social support website. *Prog Transplant.* 2012;22:119–128.
13. Ebony Boulware L, Ratner LE, Ann Sosa J, et al. Determinants of willingness to donate living related and cadaveric organs: identifying opportunities for intervention. *Transplantation.* 2002;73:1683–1691.
14. Hartzband P, Groopman J. Untangling the web — patients, doctors, and the internet. *N Engl J Med.* 2010;362:1063–1066.
15. Eysenbach G, Powell J, Kuss O, et al. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *JAMA.* 2002;287:2691–2700.
16. Moor M, Banerjee O, Abad ZSH, et al. Foundation models for generalist medical artificial intelligence. *Nature.* 2023;616:259–265.
17. Bianchi T. Global market share of leading desktop search engines 2015–2024. Available at <https://www.statista.com/statistics/216573/>

- worldwide-market-share-of-search-engines/. Accessed August 3, 2024.
18. Weis, B. D. How to write easy to read health materials: national library of medicine website. In: *Health Literacy: A Manual for Clinicians*. American Medical Association, American Medical Foundation, 2003.
 19. Brega AG, Barnard J, Mabachi NM, et al; Agency for Healthcare Research and Quality. Assess, select, and create easy-to-understand materials. In: *AHRQ Health Literacy Universal Precautions Toolkit*. 2nd Edition. Agency for Healthcare Research and Quality; 35–38. Available at <https://archive.ahrq.gov/health-literacy/precautions/healthlittoolkit2.pdf>. Accessed August, 2024.
 20. Dageforde LA, Petersen AW, Feurer ID, et al. Health literacy of living kidney donors and kidney transplant recipients. *Transplantation*. 2014;98:88–93.
 21. Swanson K, He S, Calvano J, et al. Biomedical text readability after hypenym substitution with fine-tuned large language models. *PLOS Digit Health*. 2024;3:e0000489.
 22. Kirchner GJ, Kim RY, Weddle JB, et al. Can artificial intelligence improve the readability of patient education materials? *Clin Orthop Relat Res*. 2023;481:2260–2267.
 23. Abreu AA, Murimwa GZ, Farah E, et al. Enhancing readability of online patient-facing content: the role of AI chatbots in improving cancer information accessibility. *J Natl Compr Canc Netw*. 2024;22:e237334.
 24. Gupta M, Gupta P, Ho C, et al. Can generative AI improve the readability of patient education materials at a radiology practice? *Clin Radiol*. 2024;79:e1366–e1371.
 25. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the internet: Caveant Lector et Viewor—let the reader and viewer beware. *JAMA*. 1997;277:1244–1245.
 26. He Z, Liang T, Jiao W, et al. Exploring human-like translation strategy with large language models. *Trans Assoc Comput Linguist*. 2024;12:229–246.
 27. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health*. 2023;2:e0000198.
 28. Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med*. 2023;183:589–596.