

Original Article

Open Access



Exploring the role of ChatGPT in decision making for gender-affirming surgery

Mitchell R. Koss¹ , Matthew McLaughlin¹, Kayla Switalla², Israel Falade¹, Esther Kim³

¹School of Medicine, University of California San Francisco, San Francisco, CA 94143, USA.

²University of Minnesota Medical School, Minneapolis, MN 55455, USA.

³Division of Plastic and Reconstructive Surgery, Department of Surgery, University of California San Francisco, San Francisco, CA 94143, USA.

Correspondence to: Dr. Esther Kim, Division of Plastic and Reconstructive Surgery, Department of Surgery, University of California San Francisco, 350 Parnassus Ave, San Francisco, CA 94143, USA. E-mail: Esther.Kim@ucsf.edu

How to cite this article: Koss MR, McLaughlin M, Switalla K, Falade I, Kim E. Exploring the role of ChatGPT in decision making for gender-affirming surgery. *Art Int Surg*. 2025;5:116-25. <https://dx.doi.org/10.20517/ais.2024.91>

Received: 15 Oct 2024 **First Decision:** 11 Dec 2024 **Revised:** 26 Dec 2024 **Accepted:** 11 Jan 2025 **Published:** 20 Feb 2025

Academic Editor: Andrew Gumbs **Copy Editor:** Pei-Yun Wang **Production Editor:** Pei-Yun Wang

Abstract

Aim: Large language models (LLMs) like ChatGPT have transformed access to health information. For transgender individuals considering gender-affirming surgery (GAS), accurate and reliable information is essential for informed decision making. This study aimed to quantitatively assess the use of ChatGPT among individuals considering GAS and its impact on their decision-making process.

Methods: A cross-sectional survey was conducted in January 2024 on Prolific. Participants included English-speaking U.S. users over 18 whose current gender differed from their assigned gender at birth. The survey collected demographic information, evaluated interest in GAS, and examined interactions with ChatGPT. Descriptive statistics were used for analysis.

Results: The study included 207 participants (average age 30.2 years), primarily identifying as non-binary (40.6%), transgender men (29.5%), and transgender women (13%). Most expressed interest in GAS (89%). Primary information sources for GAS were online forums (24.6%), medical websites (21.3%), and social media (17.4%). While many had used ChatGPT (73%), few utilized it for GAS information (6.7%). Among those who did, the majority (70%) rated its usefulness as moderate to slight, with some reporting a positive influence on their decision making (40%). Trust in ChatGPT's information was moderate to highly rated by 80% of participants.

Conclusions: In our cohort, ChatGPT is less commonly used for GAS information than online forums and medical



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.



websites. This suggests that patients prefer platforms that offer visual content, human interaction, and relatability. These findings highlight the importance of guiding patients toward reliable health information sources, such as healthcare providers, reputable medical websites, and academic literature, to support informed decision making.

Keywords: ChatGPT, gender-affirming surgery, decision making, transgender health, health information sources

INTRODUCTION

Large language models (LLMs), such as ChatGPT, have transformed information access by providing immediate and comprehensive responses to a wide range of prompts. For transgender and gender-diverse individuals considering gender-affirming surgery (GAS), having access to accurate, reliable, and easily understandable information is crucial for making informed decisions about such complex and sensitive procedures. Patients require a thorough knowledge of their surgical options, potential risks, recovery processes, and expected outcomes.

LLMs have been shown to be effective in many areas of healthcare delivery. For example, one study found that 45% of LLM-generated clinical summaries were equivalent to, and 36% were superior to, those of medical experts, underscoring their potential to perform at or above expert levels^[1]. Additionally, ChatGPT demonstrated its competency by passing the United States Medical Licensing Examination with a score of 64.4%, further showcasing its understanding of medical knowledge^[2].

Despite these successes, concerns about misinformation, readability, and the oversimplification of complex medical concepts persist. A recent qualitative study among health informatics researchers found that while LLMs significantly benefit patient education, clinical tasks, personalized care, and patient-healthcare interactions, they also pose risks, such as spreading misinformation, biased decision making, and inaccuracies in communication^[3]. These limitations are echoed by other researchers, who acknowledge the potential of LLMs but emphasize the ease with which misinformation can spread due to a lack of transparency and accountability^[4-6]. Furthermore, the readability of LLM-generated content often exceeds the average literacy level of U.S. patients, which could impair their ability to understand essential details, particularly in a sensitive and personal subject like GAS^[7].

While LLMs have risen to prominence for their broader healthcare applications, their role in specialized areas, like GAS, remains largely unexplored. GAS is a rapidly evolving field, with the number of procedures tripling in recent years^[8]. This increase in operations has heightened the need for transgender and gender-diverse individuals to be able to access accurate and reliable information as they navigate the process. Prior research has shown that online forums, social media, and medical websites often serve as primary sources of GAS information for many patients, who frequently seek firsthand experiences and community support^[9-11]. However, there is limited evidence on how this population uses LLMs like ChatGPT and whether these tools meet the unique needs of transgender and gender-diverse patients considering GAS.

Expanding our understanding of this area is crucial for improving patient education and ensuring that individuals considering GAS have access to reliable, accurate, and comprehensive information. As the number of GAS procedures continues to rise, equipping patients with trustworthy resources is key for enhancing their health outcomes and overall satisfaction with the surgical process. Therefore, we aimed to quantitatively assess the extent of ChatGPT use among individuals considering GAS and evaluate how it influences their decision-making process.

METHODS

Study design and setting

This cross-sectional study received an exemption from the institutional review board at the University of California, San Francisco. The primary aim was to understand how patients seeking GAS utilize ChatGPT. An anonymous 5-minute online survey was developed using Prolific (London, UK, 2024), an online research platform, and distributed to eligible users in January 2024. The full survey, including all questions, is provided in [Supplementary Materials](#). Prolific offers advantages as a research survey platform, including access to a diverse participant pool, built-in demographic screening tools, and a straightforward process for targeting specific populations. The platform is frequently used in academic research for its ability to deliver reliable data quickly and cost-effectively.

The survey included multiple-choice, free-response, Likert scale, and checkbox-style questions. Participants were compensated \$1.10 for completing the survey, in accordance with Prolific's recommended rate of \$12/h, ensuring fair and ethical compensation. We acknowledge that participant compensation may influence responses, as individuals may feel incentivized to complete surveys quickly rather than thoughtfully. To mitigate this, we designed the survey to be concise, clear, and straightforward, minimizing opportunities for rushed or disengaged responses. Additionally, Prolific's payment structure is designed to provide fair compensation without incentivizing low-quality data. We chose Prolific due to its transparency, demographic screening capabilities, and robust reputation for high-quality data in academic research. Alternative platforms, such as MTurk, have been criticized for issues such as lack of demographic diversity and unreliable responses. Traditional methods of recruitment, such as in-person surveys or phone interviews, were not feasible for this study, given time and resource constraints.

Study population

Individuals were included if they were English-speaking, aged 18 and older, based in the United States, and indicated that their current gender differed from their assigned gender at birth. Participants were selected based on pre-screening criteria available within Prolific, such as age range and self-reported interest in health-related topics. All participants had pre-noted demographic data from when they signed up for a Prolific account which allowed them to be included in the study. These criteria were established to ensure that participants were likely to be representative of individuals who might seek information online about gender surgery. On Prolific, participants provide demographic information when creating an account, allowing us to identify eligible individuals based on their self-reported demographics. To ensure participants were truthful about their demographic information, we utilized Prolific's pre-screening tools, which require users to provide verified demographic details before they can participate in studies. Additionally, we included attention-check questions and survey logic consistency checks to detect and exclude participants providing potentially unreliable responses. Efforts were made to obtain a racially and ethnically diverse sample by distributing the survey to 200 Prolific users, approximately half of whom identified as White and half as non-White.

Data collection

Data were collected through the online survey, which gathered information on demographics, interest in GAS, and interactions with ChatGPT. To ensure data quality, multiple attention checks were embedded throughout the survey, and participants who failed these checks were excluded from the final analysis. In the survey, GAS was defined as any of the following procedures: mastectomy, breast augmentation, facial feminization, facial masculinization, vaginoplasty, phalloplasty, and/or metoidioplasty. For participants who used ChatGPT to access information about GAS, additional questions explored the scope of information sought and how this information compared to other resources.

Data analysis/statistics

Descriptive statistics were used to summarize participant characteristics and study variables. Frequencies and percentages were calculated for categorical variables, while means and standard deviations were calculated for continuous variables. All statistical analyses were conducted using STATA/IC software (StataCorp LP, College Station, Texas).

RESULTS

Among the 207 participants, 100% reported that their current gender differs from their assigned gender at birth. The majority identified as non-binary (40.6%), followed by transgender men (29.5%) and transgender women (13%), with an average age of 30.2 years [Table 1]. A high level of interest in GAS was observed, with 89% expressing interest [Table 2]. When seeking information about GAS, participants primarily relied on online forums or communities such as Reddit and Tumblr (24.6%) and medical websites such as WebMD and Mayo Clinic (21.3%). Social media platforms were also frequently used (17.4%), alongside healthcare professionals (15%) and personal stories (13.5%) [Figure 1].

Regarding the use of ChatGPT, 73% of participants had used the platform, though only a small subset (6.7%) used it specifically as a resource for GAS information. Participants who used ChatGPT for GAS information generally had positive perceptions, with 70% rating the information as moderately to slightly useful. Additionally, 40% of users felt ChatGPT had a positive influence on their decision-making process, and 80% expressed moderate to high levels of trust in the information provided.

When comparing ChatGPT's information to other sources, 25% of users found it comparable to advice from healthcare providers, while 60% viewed it as similar to information found on medical websites. However, 20% considered ChatGPT's information somewhat worse than that on medical websites, and 30% found it somewhat worse than scientific journals. Despite these varied opinions, a majority (60%) of ChatGPT users indicated they would recommend the tool to peers considering GAS.

Participants used ChatGPT to search for various GAS-related topics, with the most common being facial feminization, facial masculinization, and mastectomy (18.2% each). Other frequently queried topics included metoidioplasty, phalloplasty, and breast augmentation (13.6% each). 30% of participants used ChatGPT to learn about surgeon information. Many also sought information about surgical risks, techniques, and recovery processes, although some felt the platform provided insufficient detail on financial considerations (23.1%) and recovery information (15.4%). Despite these limitations, ChatGPT was seen by many as a useful supplementary resource.

Several participants shared their personal thoughts and experiences using ChatGPT for GAS information. One participant noted that "ChatGPT gave a detailed explanation of what gender-affirming surgery is", while another commented that it "can be efficient to an extent". Others appreciated its straightforwardness, with one remarking that it was "pretty accurate to what I wanted to know... and straight to the point".

DISCUSSION

This study aimed to characterize the use of ChatGPT as a resource for obtaining information about GAS among transgender and gender-diverse individuals in the United States. While most participants had used ChatGPT, its application specifically for GAS information was notably limited compared to more traditional resources such as online forums and medical websites. These findings align with prior research suggesting that patients navigating complex health decisions, such as GAS, prefer personalized, human-centered resources over automated tools^[9,10]. There are several reasons why participants may not have wanted to use

Table 1. Demographics

Characteristic	Value
Age (mean; SD, range)	30; 9.3, 18-73
Gender identity	
Agender	12 (5.8)
Cisgender man	3 (1.4)
Cisgender woman	3 (1.4)
Gender fluid	8 (3.9)
Gender queer	9 (4.3)
Non-binary	84 (40.6)
Transgender man	61 (29.5)
Transgender woman	27 (13.0)
Race/ethnicity	
American Indian or Alaskan Native	7 (3.4)
Asian/Pacific Islander	27 (13.0)
Black or African American	26 (12.6)
Hispanic	28 (13.5)
White/Caucasian	99 (47.8)
Multiple ethnicities or other	20 (9.7)
Sexual orientation	
Asexual	29 (14.0)
Bisexual	53 (25.6)
Gay	26 (12.6)
Lesbian	16 (7.7)
Pansexual	25 (12.1)
Queer	41 (19.8)
Straight (heterosexual)	14 (6.8)
Other	3 (1.4)
Geographic region	
Midwest	56 (27.1)
Northeast	41 (19.8)
South	62 (30.0)
West	48 (23.2)
Highest level of education	
Elementary school	1 (0.5)
Some high school	6 (2.9)
High school/GED	34 (16.4)
Some college	76 (36.7)
Bachelor's degree	69 (33.3)
Graduate level degree	21 (10.1)
Employment	
Full-time employment	72 (34.8)
Part-time employment	47 (22.7)
Student	29 (14.0)
Unemployed	39 (18.8)
Disabled	14 (6.8)
Other	6 (2.9)

Table 2. Mean interest in GAS by gender identity on a 10-point Likert scale

Characteristic	Interest in GAS (0-10)
Gender identity	
Agender	4.2
Cisgender man	3.7
Cisgender woman	3.3
Gender fluid	4.5
Gender queer	6.0
Non-binary	4.4
Transgender man	8.4
Transgender woman	7.6

GAS: Gender-affirming surgery.

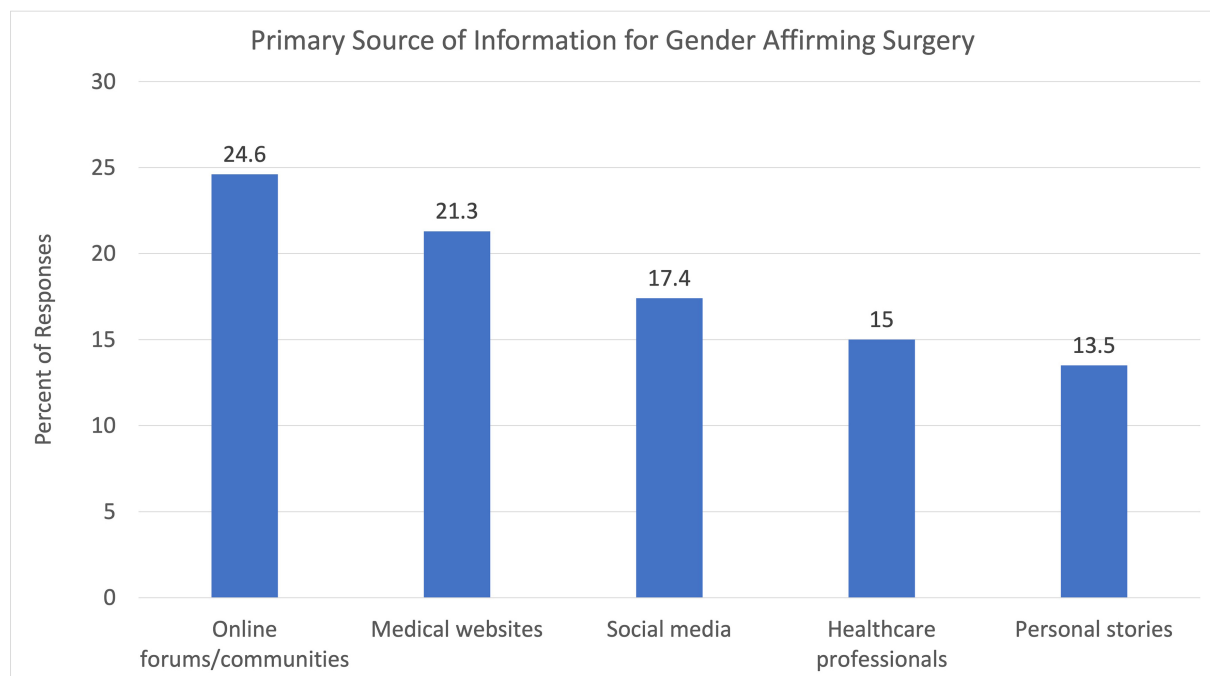


Figure 1. Primary source of information for GAS. GAS: Gender-affirming surgery.

ChatGPT as a leading information source. Participants specifically commented on the efficiency and straightforward nature of ChatGPT compared to other sources. Future studies should seek to evaluate platforms that offer a variety of perspectives beyond efficiency alone - such as those that include high-fidelity images, opportunities for human interaction, or a sense of community - elements that are currently missing from ChatGPT's functionality^[6,12].

Prior studies have emphasized the importance of personal experience and relatability in health decision making, particularly for marginalized groups such as transgender and gender-diverse individuals^[13,14]. Social media platforms and online forums have become essential spaces for peer support and sharing firsthand experiences, offering a level of emotional validation and specificity that LLMs like ChatGPT are unable to replicate^[15-17]. Although ChatGPT can deliver broad, data-driven responses, its utility in addressing the personal, psychological, and emotional aspects of GAS decision making appears limited^[4,18]. This may

explain why it is not yet seen as a primary resource for making such important medical decisions.

The underutilization of ChatGPT for GAS information could also be attributed to broader concerns about the accuracy and trustworthiness of LLMs in healthcare settings. Several studies have raised alarms about the potential for LLMs to disseminate misinformation or oversimplify complex medical concepts, which can lead to patient confusion or misinformed decisions^[4,19]. Furthermore, the ability of LLMs to account for patient-specific factors, such as individual medical histories or co-occurring health conditions, remains limited. For transgender patients, whose healthcare needs are often highly specialized and require tailored interventions, this lack of personalization could further diminish trust in ChatGPT as a reliable resource.

Interestingly, despite the limited use of ChatGPT for information on GAS, some participants did report that it positively influenced their decision making. This indicates that while ChatGPT may not currently serve as a primary resource, it may have greater utility as a supplementary tool, especially as LLMs evolve to integrate more specialized medical data and provide real-time, accurate patient feedback. Furthermore, research in fields such as ophthalmology and urology has shown that, while LLMs can provide reasonably accurate and comprehensive information, they often fall short of addressing the full range of patient needs^[20,21], underscoring ChatGPT's potential to complement, rather than replace, traditional sources of medical information for GAS^[22].

This study additionally identified several areas where ChatGPT's content could be improved to better meet the needs of individuals seeking information on GAS. Participants specifically noted that ChatGPT provided insufficient details on financial considerations, surgical techniques, and recovery processes - key elements required in the decision-making process for GAS. This mirrors findings from other research where LLMs have been criticized for their inability to provide comprehensive, context-specific medical information, particularly in areas that require detailed, patient-centered guidance^[18]. These limitations are particularly concerning in fields like transgender healthcare, where access to accurate, personalized, and affirming medical information is often limited^[23-25].

From a broader perspective, the findings of this study emphasize the ongoing need to guide patients toward trusted, reputable sources of medical information, especially for GAS, where misinformation can have serious and life-long consequences. Healthcare providers should guide patients toward high-quality resources, including peer-reviewed medical websites such as GAS websites produced by academic institutions, such as our institution's transgender care website <https://genderaffirmingsurgicalcare.ucsf.edu/>, and consultations with trained professionals. It is also equally important to approach the integration of LLMs into healthcare with caution, emphasizing that these technologies should complement, rather than replace, human expertise and empathy.

Our study is not without limitations. The small number of participants who used ChatGPT specifically for information on GAS limits the generalizability of our results. Additionally, the reliance on self-reported data introduces potential biases, such as over-reporting or under-reporting the use of ChatGPT or other information sources. Self-reported demographics such as gender identity and sexual orientation may also introduce bias into the results given differences in participant understanding of specific terms, which could differ from widely accepted definitions. There are limitations to the use of Prolific, including reliance on self-reported data, the potential for participant bias due to financial incentives, and the fact that the participant pool may not fully represent the general population. While Prolific participants may not fully represent the broader population, prior research has shown that data quality from Prolific is comparable to or better than other commonly used platforms such as MTurk. Efforts were also made to reduce the risk of

inaccurate data through attention-check questions and Prolific's required demographic verification. Our survey was not exhaustive and did not offer the opportunity for detailed responses such as reasons for ranking ChatGPT lower than other sources. However, the diversity of our sample, in terms of gender identity and racial background, provides a valuable snapshot of the current landscape of information-seeking behaviors among transgender and gender-diverse individuals considering GAS.

The implications of this study extend beyond the immediate context of GAS. As LLMs become more integrated into the larger healthcare landscape, it is essential for researchers and clinicians to continually assess their utility in various fields, particularly those that require high levels of personalization and patient-centered care. At the time of the study, ChatGPT was not a primary resource for participants seeking information about GAS. However, among those who utilized it, the platform was generally well-recommended. As AI technology continues to gain popularity, it is likely that its use for sourcing information about GAS will increase. Future research should investigate how ChatGPT and other LLMs can be adapted to better address the needs of patients considering complex medical procedures. This may involve incorporating multimedia features, such as video explanations or interactive Q&A formats, to improve user engagement and build trust. Additionally, future studies should aim to increase sample sizes and include more diverse populations to provide a more comprehensive understanding of patient education within complex surgical domains. Including definitions of key terms for participants in future surveys could also help minimize variability in understanding and improve consistency in responses.

In conclusion, this study found that ChatGPT was underutilized as a resource for information on GAS, though it had a positive impact on patient education and provided useful information for those who did use it. Participants rated ChatGPT as moderately helpful but identified gaps in areas such as financial considerations, surgical techniques, and recovery details. These findings highlight that while ChatGPT shows promise as a supplementary resource for patients considering GAS, it is not yet perceived as a primary source of reliable information. Given the critical need for trustworthy and comprehensive guidance during the decision-making process for GAS, this emphasizes the importance of steering patients toward more trusted, human-centered resources, while continuing to refine language models to better meet the needs of diverse patient populations.

DECLARATIONS

Authors' contributions

Made substantial contributions to data analysis, manuscript drafting, and manuscript revision: Koss MR

Made substantial contributions to the conception and design of the study, data acquisition, and manuscript revision: McLaughlin M

Made substantial contributions to data analysis: Switalla K

Made substantial contributions to the conception and design of the study, manuscript revision, and technical support: Falade I

Made substantial contributions to the conception and design of the study, as well as providing administrative support: Kim E

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Financial support and sponsorship

The study was funded by the personal funds of the Principal Investigator. These funds were used to compensate participants for their time via the Prolific research platform. The authors declared no external

funding sources and no conflicts of interest related to this financial arrangement.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

The study was performed in accordance with the ethical principles as stated in the Declaration of Helsinki. This study was reviewed and determined to be exempt by the Institutional Review Board (IRB) at the University of California, San Francisco. The study met the criteria for exemption as it involved anonymous survey data collection with minimal risk to participants. Participation was entirely voluntary, and all participants provided informed consent electronically before beginning the survey. Participants were informed about the purpose of the study, the approximate time commitment, and their right to withdraw at any time without penalty.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2025.

REFERENCES

1. Van Veen D, Van Uden C, Blankemeier L, et al. Adapted large language models can outperform medical experts in clinical text summarization. *Nat Med.* 2024;30:1134-42. [DOI](#) [PubMed](#) [PMC](#)
2. Gilson A, Safranek CW, Huang T, 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 Med Educ.* 2023;9:e45312. [DOI](#) [PubMed](#) [PMC](#)
3. Denecke K, May R, Rivera Romero O; LLMHealthGroup. Potential of large language models in health care: delphi study. *J Med Int Res.* 2024;26:e52399. [DOI](#) [PubMed](#) [PMC](#)
4. Clusmann J, Kolbinger FR, Muti HS, et al. The future landscape of large language models in medicine. *Commun Med.* 2023;3:141. [DOI](#) [PubMed](#) [PMC](#)
5. Karabacak M, Margetis K. Embracing large language models for medical applications: opportunities and challenges. *Cureus.* 2023;15:e39305. [DOI](#) [PubMed](#) [PMC](#)
6. Liu J, Wang C, Liu S. Utility of ChatGPT in clinical practice. *J Med Int Res.* 2023;25:e48568. [DOI](#) [PubMed](#) [PMC](#)
7. Amin K, Doshi R, Forman HP. Large language models as a source of health information: are they patient-centered? A longitudinal analysis. *Healthc.* 2024;12:100731. [DOI](#) [PubMed](#)
8. Wright JD, Chen L, Suzuki Y, Matsuo K, Hershman DL. National estimates of gender-affirming surgery in the US. *JAMA Netw Open.* 2023;6:e2330348. [DOI](#) [PubMed](#) [PMC](#)
9. Hirpara MM, Amin L, Aloyan T, Shilleh N, Lewis P. Does the internet provide quality information on metoidioplasty? Using the modified ensuring quality information for patients tool to evaluate artificial intelligence-generated and online information on metoidioplasty. *Ann Plast Surg.* 2024;92:S361-5. [DOI](#) [PubMed](#)
10. Berry CE, Fazilat AZ, Churukian AA, et al. Quality assessment of online resources for gender-affirming surgery. *Plast Reconstr Surg Glob Open.* 2023;11:e5306. [DOI](#) [PubMed](#) [PMC](#)
11. Vargas CR, Ricci JA, Lee M, Tobias AM, Medalie DA, Lee BT. The accessibility, readability, and quality of online resources for gender affirming surgery. *J Surg Res.* 2017;217:198-206. [DOI](#) [PubMed](#)
12. Kalam KT, Rahman JM, Islam MR, Dewan SMR. ChatGPT and mental health: friends or foes? *Health Sci Rep.* 2024;7:e1912. [DOI](#) [PubMed](#) [PMC](#)
13. Gerritse K, Hartman LA, Bremmer MA, Kreukels BPC, Molewijk BC. Decision-making approaches in transgender healthcare: conceptual analysis and ethical implications. *Med Health Care Philos.* 2021;24:687-99. [DOI](#) [PubMed](#) [PMC](#)
14. Lambert A, Pratt A, Conard LAE, et al. Supporting gender-related medical decision making for transgender and gender-diverse individuals: a scoping review. *Transgend Health.* 2023;8:113-23. [DOI](#) [PubMed](#) [PMC](#)
15. Song H, Xia Y, Luo Z, et al. Evaluating the performance of different large language models on health consultation and patient education in urolithiasis. *J Med Syst.* 2023;47:125. [DOI](#)
16. Entwistle VA, France EF, Wyke S, et al. How information about other people's personal experiences can help with healthcare decision-making: a qualitative study. *Patient Educ Couns.* 2011;85:e291-8. [DOI](#)

17. Bussey LG, Sillence E. The role of internet resources in health decision-making: a qualitative study. *Digit Health.* 2019;5:2055207619888073. [DOI](#) [PubMed](#) [PMC](#)
18. Wilhelm TI, Roos J, Kaczmarczyk R. Large language models for therapy recommendations across 3 clinical specialties: comparative study. *J Med Internet Res.* 2023;25:e49324. [DOI](#) [PubMed](#) [PMC](#)
19. Anibal J, Huth H, Gunkel J, Gregurick S, Wood B. Simulated misuse of large language models and clinical credit systems. *medRxiv.* 2024. [DOI](#)
20. Pushpanathan K, Lim ZW, Er Yew SM, et al. Popular large language model chatbots' accuracy, comprehensiveness, and self-awareness in answering ocular symptom queries. *iScience.* 2023;26:108163. [DOI](#) [PubMed](#) [PMC](#)
21. Davis R, Eppler M, Ayo-Ajibola O, et al. Evaluating the effectiveness of artificial intelligence-powered large language models application in disseminating appropriate and readable health information in urology. *J Urol.* 2023;210:688-94. [DOI](#)
22. Stroop A, Stroop T, Zawy Alsofy S, et al. Large language models: are artificial intelligence-based chatbots a reliable source of patient information for spinal surgery? *Eur Spine J.* 2024;33:4135-43. [DOI](#)
23. Bhatt N, Cannella J, Gentile JP. Gender-affirming care for transgender patients. 2022;19:23-32. [PubMed](#) [PMC](#)
24. Kraschel KL, Chen A, Turban JL, Cohen IG. Legislation restricting gender-affirming care for transgender youth: politics eclipse healthcare. *Cell Rep Med.* 2022;3:100719. [DOI](#) [PubMed](#) [PMC](#)
25. Pastor Bravo MDM, Linander I. Access to healthcare among transgender and non-binary youth in Sweden and Spain: a qualitative analysis and comparison. *PLoS One.* 2024;19:e0303339. [DOI](#) [PubMed](#) [PMC](#)