

Original Paper

Laypeople's Use of and Attitudes Toward Large Language Models and Search Engines for Health Queries: Survey Study

Tamir Mendel¹, PhD; Nina Singh², MD; Devin M Mann³, MS, MD; Batia Wiesenfeld⁴, PhD; Oded Nov¹, PhD

¹Department of Technology Management and Innovation, Tandon School of Engineering, New York University, New York, NY, United States

²Department of Medicine, School of Medicine, University of California, San Francisco, San Francisco, CA, United States

³Department of Population Health, Grossman School of Medicine, New York University, New York, NY, United States

⁴Department of Management and Organizations, Stern School of Business, New York University, New York, NY, United States

Corresponding Author:

Tamir Mendel, PhD

Department of Technology Management and Innovation

Tandon School of Engineering

New York University

2 Metrotech Center, Brooklyn

New York, NY, 11201

United States

Phone: 1 8287348968

Email: tamir.mendel@nyu.edu

Abstract

Background: Laypeople have easy access to health information through large language models (LLMs), such as ChatGPT, and search engines, such as Google. Search engines transformed health information access, and LLMs offer a new avenue for answering laypeople's questions.

Objective: We aimed to compare the frequency of use and attitudes toward LLMs and search engines as well as their comparative relevance, usefulness, ease of use, and trustworthiness in responding to health queries.

Methods: We conducted a screening survey to compare the demographics of LLM users and nonusers seeking health information, analyzing results with logistic regression. LLM users from the screening survey were invited to a follow-up survey to report the types of health information they sought. We compared the frequency of use of LLMs and search engines using ANOVA and Tukey post hoc tests. Lastly, paired-sample Wilcoxon tests compared LLMs and search engines on perceived usefulness, ease of use, trustworthiness, feelings, bias, and anthropomorphism.

Results: In total, 2002 US participants recruited on Prolific participated in the screening survey about the use of LLMs and search engines. Of them, 52% (n=1045) of the participants were female, with a mean age of 39 (SD 13) years. Participants were 9.7% (n=194) Asian, 12.1% (n=242) Black, 73.3% (n=1467) White, 1.1% (n=22) Hispanic, and 3.8% (n=77) were of other races and ethnicities. Further, 1913 (95.6%) used search engines to look up health queries versus 642 (32.6%) for LLMs. Men had higher odds (odds ratio [OR] 1.63, 95% CI 1.34-1.99; $P<.001$) of using LLMs for health questions than women. Black (OR 1.90, 95% CI 1.42-2.54; $P<.001$) and Asian (OR 1.66, 95% CI 1.19-2.30; $P<.01$) individuals had higher odds than White individuals. Those with excellent perceived health (OR 1.46, 95% CI 1.1-1.93; $P=.01$) were more likely to use LLMs than those with good health. Higher technical proficiency increased the likelihood of LLM use (OR 1.26, 95% CI 1.14-1.39; $P<.001$). In a follow-up survey of 281 LLM users for health, most participants used search engines first (n=174, 62%) to answer health questions, but the second most common first source consulted was LLMs (n=39, 14%). LLMs were perceived as less useful ($P<.01$) and less relevant ($P=.07$), but elicited fewer negative feelings ($P<.001$), appeared more human (LLM: n=160, vs search: n=32), and were seen as less biased ($P<.001$). Trust ($P=.56$) and ease of use ($P=.27$) showed no differences.

Conclusions: Search engines are the primary source of health information; yet, positive perceptions of LLMs suggest growing use. Future work could explore whether LLM trust and usefulness are enhanced by supplementing answers with external references and limiting persuasive language to curb overreliance. Collaboration with health organizations can help improve the quality of LLMs' health output.

(*J Med Internet Res* 2025;27:e64290) doi: [10.2196/64290](https://doi.org/10.2196/64290)

KEYWORDS

large language model; artificial intelligence; LLMs; search engine; Google; internet; online health information; United States; survey; mobile phone

Introduction

Search engines, such as Google, democratized access to health information and changed the dynamics of the patient-provider relationship, with 72% of internet users in the United States looking on the web for health information [1-4]. Clinicians and health organizations initially had mixed reactions, including some early efforts to discourage patients from searching for health information on the web because of concerns that web-based information could make them misinformed or anxious [5-7]. Patients' use of internet-based health information enhances their understanding and their ability to manage their health conditions [8,9]. Over time, clinicians and health organizations started to partner with search engine companies to improve content reliability. For example, the World Health Organization has worked with Google to ensure that COVID-19 searches yield evidence-based information [10].

Large language models (LLMs) have the potential to replace internet searches for clinicians and patients. LLMs, such as ChatGPT, have demonstrated promising performance in clinical decision-making [11] and diagnosis [12]. Numerous studies have explored doctors' evaluation of LLMs as an information resource and diagnostic aid [13-18] and some research has begun to consider laypeople's attitudes toward and use of these tools, such as concerning ethical considerations, whether they are differentiable from doctor's responses, and the accuracy of symptom-checkers [19-21]. LLMs can potentially make information accessible to patients in more specific and personalized ways, but previous studies have yet to consider patient use of LLMs in comparison to search engines for health-related questions.

We surveyed laypeople in the United States to compare the use of LLMs and search engines for health queries, the types of health queries posed, and attitudes toward interactions with these tools.

Methods

Recruitment Procedures

US participants aged older than 18 years were recruited from Prolific, a web-based research participant platform, in February 2024. We programmed the survey questions in Qualtrics. The invitation in Prolific included the purpose and a brief description of this study, a link to the Qualtrics survey, and it informed participants that survey completion would take approximately 5 minutes for study 1 and 5-10 minutes for study 2. Participants completed study 1 in 1.21 (SD 1.04) minutes on average. Participants completed study 2 in 10.01 (SD 7.02) minutes on average. All questionnaires are presented in [Multimedia Appendix 1](#).

Ethical Considerations

This study was reviewed and approved by New York University's institutional review board (IRB-FY2024-8278).

Participants were presented with an informed consent form including a description of their role as participants in a research study to learn more about patterns of using search engines and LLMs for health queries. If they agreed to participate in this study, they were asked to complete a questionnaire about their experience using a search engine and an LLM for health questions. The research involves no more than minimal risk of harm to participants. At the end of the survey, participants were asked to provide a limited amount of personal data (Prolific ID) to provide compensation and asked about their willingness to participate in the second study if they qualified. Data were stored in password-protected computers of the faculty investigators and their assistants. Participants received compensation of US \$0.25 in study 1 and US \$2.50 in study 2.

Survey Questionnaire

A screening survey (study 1) identified the prevalence of participants' LLM and search engine use for health queries. Next, study 1 participants who reported using both LLMs and search engines to answer health questions were invited to participate in study 2, in which their use of these tools was compared. Study 2 participants were asked which types of health information they sought, and how the results affected their relationships with their providers. Likert scales from 1 (strongly disagree) to 7 (strongly agree) were used to evaluate the effect of LLMs and search engines on participants' relationships with health care providers.

Participants were asked about LLMs' versus search engines' perceived usefulness and ease of use. These two measures represent people's perception of whether a system would help them perform their tasks and whether the system is easy or difficult to use [22]. Trustworthiness is the extent to which people trust and rely on LLMs or search engines to answer health questions [23]. Relevance is the extent to which participants consider the source as offering output relevant to their health needs. We also asked whether the source's results were perceived to be biased or benefit advertisers. Finally, we asked about participants' reactions to LLM and search engine query results [24] and their perceptions of anthropomorphism [25]. Likert scales from 1 (strongly disagree) to 7 (strongly agree) were used to evaluate attitudes toward LLMs and search engines. A separate Likert scale ranging from 1 (not at all) to 7 (extremely) was used for items assessing participants' feelings about using LLMs and search engines. Finally, we asked about demographics: age, gender, education, income, use of LLMs for health, health status, and perceived tech skills. Perceived technical skills assessed participants' perceptions of their smartphone skills (1: not at all skilled to 7: expert; [Multimedia Appendix 1](#)).

Statistical Analysis

We compared the demographics of LLM users and nonusers using logistic regression. We used ANOVAs followed by a Tukey post hoc test to compare the frequency of use of LLMs and search engines for each of the types of health information

they sought. Using the Shapiro-Wilk test, we observed a nonnormal distribution for perceived usefulness, ease of use, relevance, bias, trustworthiness, feelings, perceptions of benefit to advertisers, and perceptions of anthropomorphism. Therefore, paired samples of Wilcoxon tests were used to compare the effects of LLMs and search engines on these perceptions.

Results

Overview

In total, 2002 participants responded to study 1 (see demographic information in [Table 1](#)). Furthermore, 300 randomly selected study 1 participants who reported using both LLMs and search engines for health queries completed study 2. Further, 281 participants were retained after indicating the use of LLMs for health-related queries and passing attention checks (see [Table 1](#) for demographic information).

Table 1. Demographic characteristics of respondents in study 1 and study 2, with study 2 participants being a subset of those from study 1.

Variable	Study 1	Study 2
Respondents, n	2002	281
Age (years), mean (SD)	39.4 (13)	40.43 (13.17)
Gender, n (%)		
Male	904 (45.2)	161 (57.3)
Female	1045 (52.2)	112 (39.9)
Other	53 (1.6)	8 (2.8)
Race, n (%)		
Asian	194 (9.7)	35 (12.5)
Black	242 (12.1)	46 (16.4)
Hispanic	22 (1.1)	3 (1.1)
White	1467 (73.3)	187 (66.5)
Other	77 (3.8)	10 (3.6)
Education		
High school or less	578 (28.9)	72 (25.6)
Associate degree	249 (12.4)	39 (13.9)
Bachelor's degree	847 (42.3)	125 (44.5)
Graduate degree	328 (16.4)	45 (16)
Income (US \$)		
Less than 50,000	715 (35.7)	106 (37.7)
50,000-99,999	694 (34.7)	103 (36.7)
More than 100,000	518 (25.9)	72 (25.6)
Unknown	75 (3.7)	0 (0)
Insurance, n (%)		
Private insurance	1166 (58.2)	159 (56.6)
Public insurance	592 (29.6)	92 (32.7)
No insurance	192 (9.6)	25 (8.9)
Other insurance	52 (2.6)	5 (1.8)
Using LLMs^a for health, n (%)		
Yes	642 (32.1)	281 (100)
No	1360 (67.9)	0 (0)
Using search engines for health, n (%)		
Yes	1913 (95.6)	281 (100)
No	89 (4.4)	0 (0)
Health status, n (%)		
Excellent	268 (13.4)	44 (15.7)
Good	1231 (61.5)	158 (56.2)
Only fair	426 (21.3)	66 (23.5)
Poor	67 (3.3)	12 (4.3)
Prefer not to say	10 (0.5)	1 (0.4)
Perceived tech skills, mean (SD)	5.60 (1.05)	5.85 (1.04)

^aLLM: large language model.

Among the participants from study 2, a total of 218 (77.6%) participants reported using ChatGPT [OpenAI] as their only LLM tool to answer health questions, 48 (17.1%) reported using both ChatGPT and other LLM tools (such as Bard [Google AI] or Copilot [Microsoft]), and only 15 (5.3%) indicated they used other LLM tools without using ChatGPT. For health-related searches, 228 (81.1%) participants used Google, 51 used both Google and other search engines (such as Bing [Microsoft]), and only 2 (<1%) used Bing alone for finding health answers. Therefore, most participants used ChatGPT as their primary LLM tool and Google search to answer health questions.

Demographic Comparison: LLM Users Versus Nonusers for Health-Related Questions

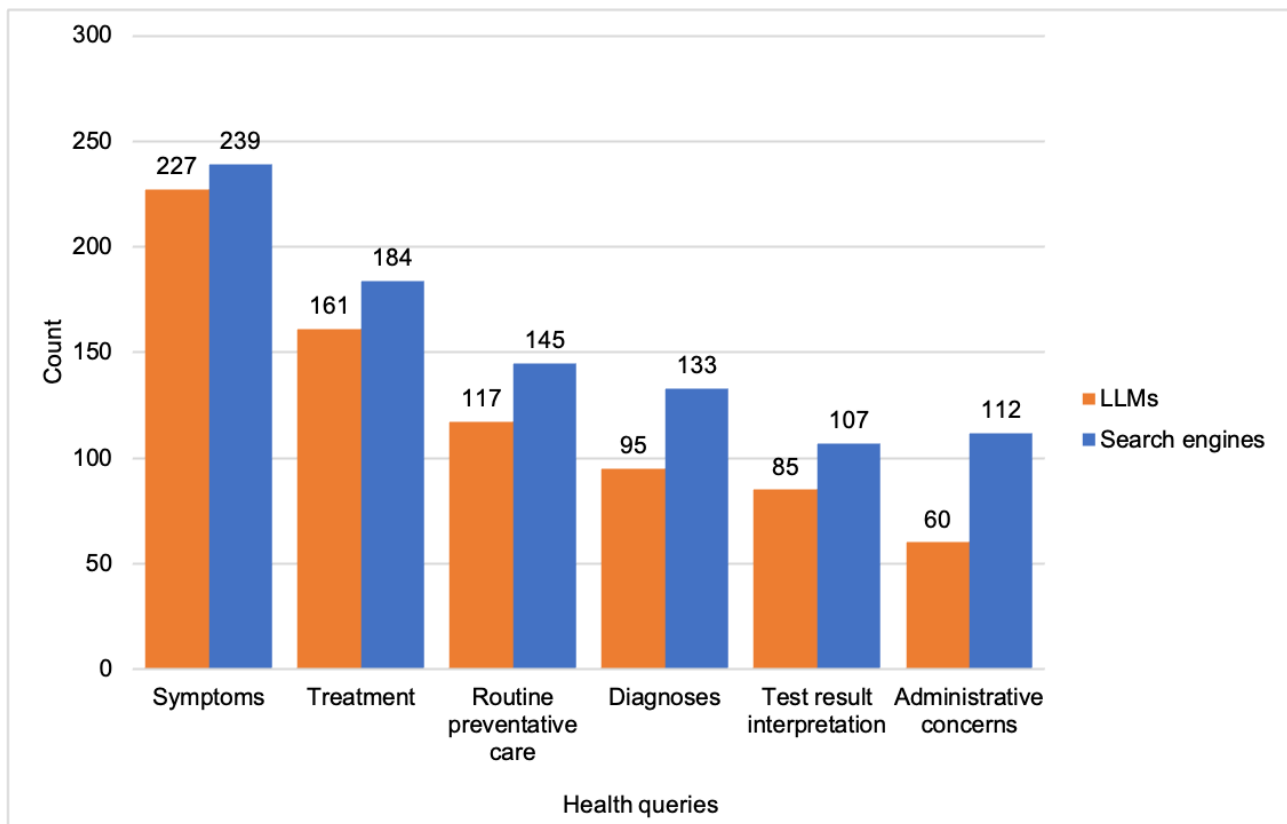
A logistic regression analysis revealed that men had higher odds (odds ratio [OR] 1.63, 95% CI 1.34-1.99; *P*<.001) of using LLMs for health-related questions compared to women. Individuals who reported their race as Black (OR 1.9, 95% CI 1.42-2.54; *P*<.001) and Asian (OR 1.66, 95% CI 1.19-2.30; *P*<.01) used LLMs for health-related questions more than participants who identified as White. Participants who perceived their health status as excellent (OR 1.46, 95% CI 1.1-1.93; *P*=.01) reported using LLMs for health-related questions more than those reporting good health status. Individuals with higher perceived technical proficiency have a higher likelihood of using

LLMs for health-related questions (OR 1.26, 95% CI 1.14-1.39, *P*<.001; see Table S1 in [Multimedia Appendix 1](#)).

LLMs Versus Search Engines

In the follow-up study (study 2), most participants used search engines first (n=174, 61.9%) to answer health questions, but the second most common first source consulted was LLMs (n=39, 13.9% of participants), followed by doctor’s appointment (n=19, 6.8%), calling their doctor (n=17, 6%), messaging their doctor (n=14, 5%), asking friends (n=7, 2.5%), going to the emergency department (n=6, 2.1%), and other options to answer health questions (n=5, 1.8%). They reported using search engines 6.91 (SD 2.67) times over the past year for health queries, compared to using LLMs 4.7 (SD 2.79) times, and other services fewer times (Figure S1 in [Multimedia Appendix 1](#), *P*<.05). Finally, as shown in [Figure 1](#), there was no difference in the likelihood that participants would use LLMs versus search engines to ask about symptoms (LLM: n=227; search engines: n=239), treatment (LLM: n=161; search engines: n=184), routine preventative care (LLM: n=117; search engines: n=145), diagnoses (LLM: n=95; search engines: n=133), and interpretation of test results (LLM: n=85; search engines: n=107), but they were more likely to use search engines than LLMs for administrative queries (LLM: n=60; search engines: n=112).

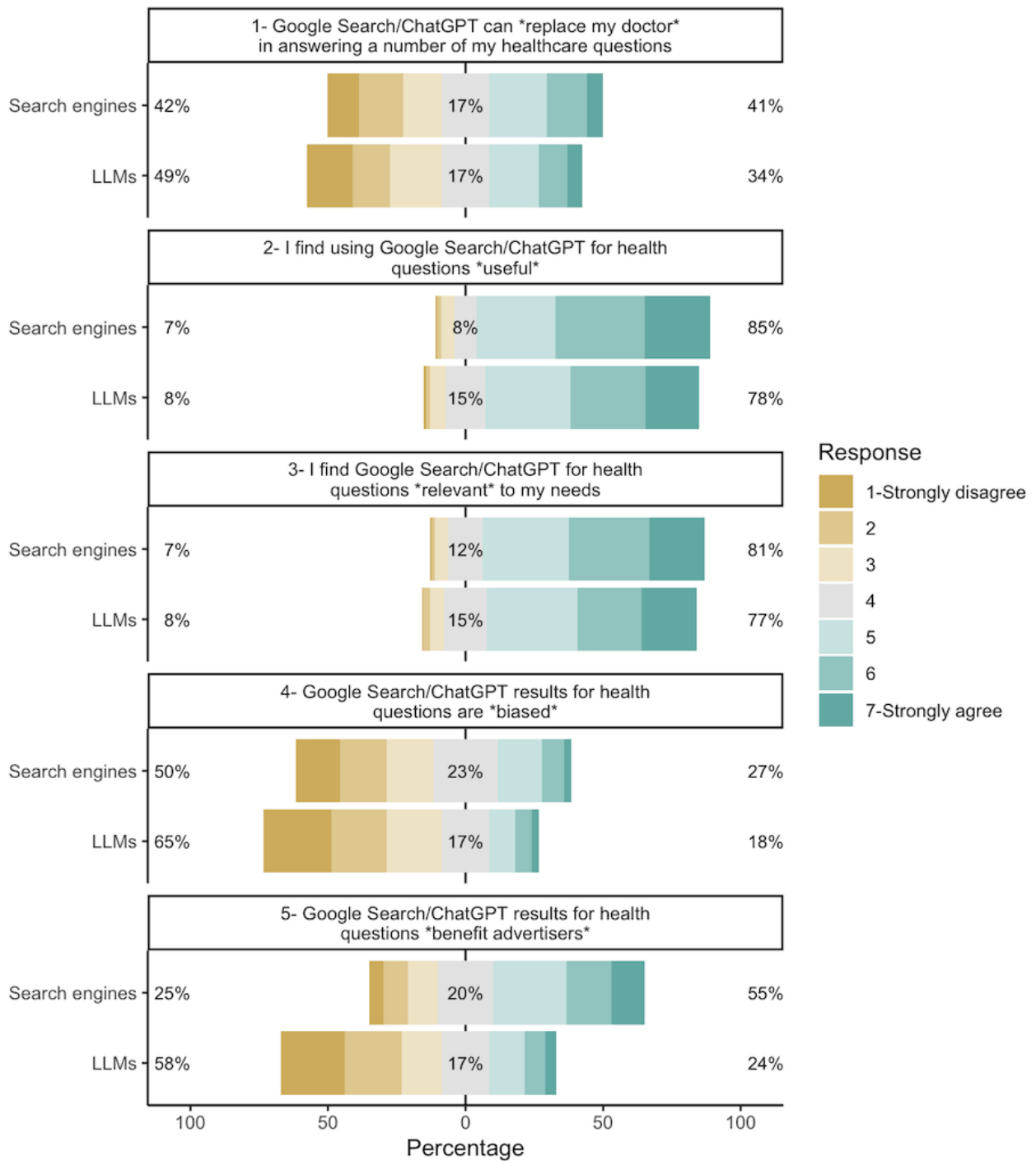
Figure 1. Distribution of health queries between LLMs and search engines among 281 participants who reported using both for health queries. The question asked was “When consulting [Google Search/ChatGPT] for health questions, which questions do you use it to answer?” LLM: large language model.



As can be seen in [Figure 2](#), LLMs were perceived to be less useful than search engines for answering health-related questions (*W*=7214.5, *P*<.01). However, LLMs were also perceived as less biased (*W*=9533.5, *P*<.001) and less beneficial to advertisers

than search engine results (*W*=20,708, *P*<.001). Moreover, participants were less likely to view LLMs as being able to replace doctors in answering health care questions than search engines (*W*=9223, *P*=.003).

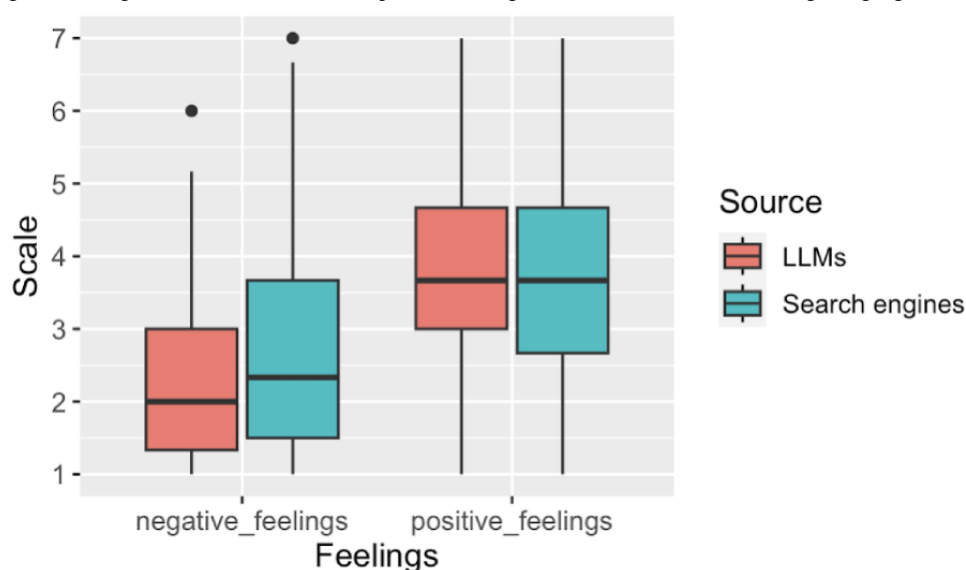
Figure 2. Comparison of LLMs and search engines among 281 participants who reported using both for health queries, using the Wilcoxon signed rank test for replacing "my doctor" ($W=9223, P=.003$), usefulness ($W=7214.5, P<.01$), relevance ($W=7117, P=.07$), result bias ($W=9533.5, P<.001$), and benefits to advertisers ($W=20708, P<.001$). LLM: large language model.



When searching for health information, results can reassure users but they can also make users more concerned [26]. As shown in Figure 3, LLMs elicited fewer negative feelings in response to the results of their queries than did search engines ($W=19,496, P<.001$). The LLMs also elicited more positive feelings in response to results than search engines did ($W=9773.5, P=.01$), but the median was lower than the midpoint of 4 ($W=977.5, P<.001$). On the other hand, while falling short of statistical significance, search engines were perceived to offer

responses more relevant to users' needs than LLMs ($W=7117, P=.07$). There were no significant differences between ratings of LLMs and search engines in ease of use or trustworthiness of answers ($P's>.05$). All paired samples of Wilcoxon tests are available in Table S2 in Multimedia Appendix 1. Finally, most participants ($n=160, 57%$) perceived LLMs as more human than search engines, but a minority perceived search engines as either more human ($n=32, 11%$) or equivalent to LLMs ($n=32, 12%$).

Figure 3. Comparison between LLMs and search engines among 281 participants who reported using both for health queries, based on the Wilcoxon signed rank test for negative feelings ($W=19,496, P<.001$) and positive feelings ($W=9773.5, P=.01$). LLM: large language model.



Discussion

Principal Findings

We compared the use of LLMs, search engines, and other sources for health queries and found that while the most common first source of information participants consulted was search engines, 13.9% of respondents consulted LLMs before any other source. Moreover, they were as likely to use LLMs as search engines for the most common health queries (symptoms, treatment, routine preventive care, diagnoses, and interpretation of test results), though they were more likely to use search engines than LLMs for less common administrative queries. In addition, LLMs elicited fewer negative feelings, such as less distress in response to the information provided, and were perceived as less useful and relevant than search engines. LLMs appeared more human-like, less biased, and less favorable toward advertisers. No significant differences were found between trust in LLMs and trust in search engines.

At this early stage of LLM adoption, participants indicated that LLMs were less useful and relevant than search engines for answering health-related questions. One key reason that search engines provide more useful and relevant health information compared to LLMs could include their greater transparency, such as providing external references to information sources that convey credibility and trustworthiness to users—features that are absent or inaccurate in many popular LLMs. The low perceived usefulness and relevance of LLMs can be attributed to issues such as incorrect statements, hallucinations, and ethical concerns [21,27,28], which can be mitigated through the involvement of domain experts, refining input prompts, and enhancing the fine-tuning of LLMs for specific tasks or domains [29].

Strengths of LLMs relative to search engines that emerged in our study include that LLMs elicited fewer negative feelings, such as less distress in response to the information provided. They also appeared more human-like, less biased, and less favorable toward advertisers. LLMs compared to search engines

provide people with a feeling of human-like and personalized attention and understanding, enhancing digital communication that makes information delivery more intuitive. Evidence outside of health care suggests that these features make LLMs as persuasive as humans [30]. LLMs have advanced state-of-the-art performance in generating human-like text based on user health questions [19].

With respect to ease of use, there was no significant difference between LLMs and search engines for health queries. This is notable because search engines have been in regular use for decades while LLMs are a relatively new lay user-facing technology. The fact that respondents perceived that new and unfamiliar technology is as easy to use as an extensively used and highly familiar technology suggests that users may shift their use to LLMs further and relatively quickly.

Prior research has found that trust evaluations shape web-based health information-seeking behavior and compliance with health advice [31-33]. No significant differences were found between trust in LLMs and trust in search engines. A possible explanation for the lack of difference in trustworthiness between search engines and LLMs is the counterbalance between the two dimensions of trust: cognitive trust based on perceived accuracy and competence, and affective trust which depends on believing one's interests are protected [32]. People perceived that search engines, compared to LLMs, provide more helpful and relevant health information, reinforcing perceived accuracy and competence in health care information. However, search engines were also perceived to prioritize advertisers' interests more than LLMs did, likely diminishing users' trust in them. While LLMs excel at engaging with users, their output often lacks interpretation within a medical context [34] and they may sometimes fabricate facts or present incorrect information more convincingly and believably [35]. Future work can further analyze the effects of each trust dimension to better understand perceptions of LLMs.

Limitations

Our study has some key limitations. All of the participants were recruited from Prolific. Prolific participants tend to be technically proficient and may be more likely to use new technologies such as LLMs than the general population. Additionally, we screened for US residency so our findings may not generalize to other countries and languages. Future studies are needed to evaluate whether these findings generalize to broader populations. Finally, the current inferential analysis aggregates attitudinal reactions across varied types of health queries. A valuable future study would separately inquire about perceptions of technology (eg, usefulness, ease of use, or trust) for different types of health queries to enable more nuanced assessments.

Conclusion

Overall, most individuals still turn to search engines as their primary source for health-related questions, considering them more useful than LLMs. In the future, people may incorporate LLMs more routinely in their search for answers to health questions. In the case of LLMs, incorporating references and external sources of information would be a beneficial practice to increase trustworthiness. Just as clinicians and health organizations have partnered with search engine companies to enhance the reliability of health-related content [12], similar collaborations could enhance the quality of the health-related information provided by LLMs.

Acknowledgments

The authors received financial support from the US National Science Foundation (grants 1928614 and 2129076) for the submitted work.

Data Availability

The anonymized data analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional materials.

[[DOCX File, 106 KB-Multimedia Appendix 1](#)]

References

1. Fox S, Duggan M, Purcell K. Family caregivers are wired for health. Pew Research Center. Washington, D.C. Pew Research Center's Internet & American Life Project; Jun 20, 2013. URL: <https://www.pewresearch.org/internet/2013/06/20/family-caregivers-are-wired-for-health/> [accessed 2025-01-15]
2. Luo A, Qin L, Yuan Y, Yang Z, Liu F, Huang P, et al. The effect of online health information seeking on physician-patient relationships: systematic review. *J Med Internet Res*. 2022;24(2):e23354. [FREE Full text] [doi: [10.2196/23354](https://doi.org/10.2196/23354)] [Medline: [35142620](https://pubmed.ncbi.nlm.nih.gov/35142620/)]
3. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. *J Med Internet Res*. 2014;16(12):e262. [FREE Full text] [doi: [10.2196/jmir.3706](https://doi.org/10.2196/jmir.3706)] [Medline: [25470306](https://pubmed.ncbi.nlm.nih.gov/25470306/)]
4. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. *JAMA*. 2002;287(20):2691-2700. [doi: [10.1001/jama.287.20.2691](https://doi.org/10.1001/jama.287.20.2691)] [Medline: [12020305](https://pubmed.ncbi.nlm.nih.gov/12020305/)]
5. Ahmad F, Hudak PL, Bercovitz K, Hollenberg E, Levinson W. Are physicians ready for patients with internet-based health information? *J Med Internet Res*. 2006;8(3):e22. [FREE Full text] [doi: [10.2196/jmir.8.3.e22](https://doi.org/10.2196/jmir.8.3.e22)] [Medline: [17032638](https://pubmed.ncbi.nlm.nih.gov/17032638/)]
6. Ahluwalia S, Murray E, Stevenson F, Kerr C, Burns J. 'A heartbeat moment': qualitative study of GP views of patients bringing health information from the internet to a consultation. *Br J Gen Pract*. 2010;60(571):88-94. [FREE Full text] [doi: [10.3399/bjgp10X483120](https://doi.org/10.3399/bjgp10X483120)] [Medline: [20132702](https://pubmed.ncbi.nlm.nih.gov/20132702/)]
7. DDB. Don't google it. Ads of the World by CLIOS. 2014. URL: <https://www.adsoftheworld.com/campaigns/don-t-google-it> [accessed 2024-04-22]
8. Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. *Arch Intern Med*. 2003;163(14):1727-1734. [doi: [10.1001/archinte.163.14.1727](https://doi.org/10.1001/archinte.163.14.1727)] [Medline: [12885689](https://pubmed.ncbi.nlm.nih.gov/12885689/)]
9. Kassirer JP. Patients, physicians, and the Internet. *Health Aff (Millwood)*. 2000;19(6):115-123. [doi: [10.1377/hlthaff.19.6.115](https://doi.org/10.1377/hlthaff.19.6.115)] [Medline: [11192394](https://pubmed.ncbi.nlm.nih.gov/11192394/)]
10. Germani F, Pattison AB, Reinfelde M. WHO and digital agencies: how to effectively tackle COVID-19 misinformation online. *BMJ Glob Health*. 2022;7(8):e009483. [FREE Full text] [doi: [10.1136/bmjgh-2022-009483](https://doi.org/10.1136/bmjgh-2022-009483)] [Medline: [35918074](https://pubmed.ncbi.nlm.nih.gov/35918074/)]

11. Liu S, Wright AP, Patterson BL, Wanderer JP, Turer RW, Nelson SD, et al. Using AI-generated suggestions from ChatGPT to optimize clinical decision support. *J Am Med Inform Assoc.* 2023;30(7):1237-1245. [FREE Full text] [doi: [10.1093/jamia/ocad072](https://doi.org/10.1093/jamia/ocad072)] [Medline: [37087108](https://pubmed.ncbi.nlm.nih.gov/37087108/)]
12. Goodman RS, Patrinely JR, Stone CA, Zimmerman E, Donald RR, Chang SS, et al. Accuracy and reliability of chatbot responses to physician questions. *JAMA Netw Open.* 2023;6(10):e2336483. [FREE Full text] [doi: [10.1001/jamanetworkopen.2023.36483](https://doi.org/10.1001/jamanetworkopen.2023.36483)] [Medline: [37782499](https://pubmed.ncbi.nlm.nih.gov/37782499/)]
13. Rao A, Pang M, Kim J, Kamineni M, Lie W, Prasad A, et al. Assessing the utility of ChatGPT throughout the entire clinical workflow. medRxiv. Preprint posted online on February 26, 2023. [FREE Full text] [doi: [10.1101/2023.02.21.23285886](https://doi.org/10.1101/2023.02.21.23285886)] [Medline: [36865204](https://pubmed.ncbi.nlm.nih.gov/36865204/)]
14. Takita H, Walston SL, Tatekawa H, Saito K, Tsujimoto Y, Miki Y, et al. Diagnostic performance of generative AI and physicians: a systematic review and meta-analysis. medrxiv. Preprint posted online on March 18, 2024. [doi: [10.1101/2024.01.20.24301563](https://doi.org/10.1101/2024.01.20.24301563)]
15. Sorin V, Barash Y, Konen E, Klang E. Large language models for oncological applications. *J Cancer Res Clin Oncol.* 2023;149(11):9505-9508. [doi: [10.1007/s00432-023-04824-w](https://doi.org/10.1007/s00432-023-04824-w)] [Medline: [37160626](https://pubmed.ncbi.nlm.nih.gov/37160626/)]
16. Benary M, Wang XD, Schmidt M, Soll D, Hilfenhaus G, Nassir M, et al. Leveraging large language models for decision support in personalized oncology. *JAMA Netw Open.* 2023;6(11):e2343689. [FREE Full text] [doi: [10.1001/jamanetworkopen.2023.43689](https://doi.org/10.1001/jamanetworkopen.2023.43689)] [Medline: [37976064](https://pubmed.ncbi.nlm.nih.gov/37976064/)]
17. Kao HJ, Chien TW, Wang WC, Chou W, Chow JC. Assessing chatGPT's capacity for clinical decision support in pediatrics: a comparative study with pediatricians using KIDMAP of Rasch analysis. *Medicine (Baltimore).* 2023;102(25):e34068. [FREE Full text] [doi: [10.1097/MD.00000000000034068](https://doi.org/10.1097/MD.00000000000034068)] [Medline: [37352054](https://pubmed.ncbi.nlm.nih.gov/37352054/)]
18. Small WR, Wiesenfeld B, Brandfield-Harvey B, Jonassen Z, Mandal S, Stevens ER, et al. Large language model-based responses to patients' in-basket messages. *JAMA Netw Open.* 2024;7(7):e2422399. [FREE Full text] [doi: [10.1001/jamanetworkopen.2024.22399](https://doi.org/10.1001/jamanetworkopen.2024.22399)] [Medline: [39012633](https://pubmed.ncbi.nlm.nih.gov/39012633/)]
19. Nov O, Singh N, Mann D. Putting chatGPT's medical advice to the (Turing) test: survey study. *JMIR Med Educ.* 2023;9:e46939. [FREE Full text] [doi: [10.2196/46939](https://doi.org/10.2196/46939)] [Medline: [37428540](https://pubmed.ncbi.nlm.nih.gov/37428540/)]
20. Kopka M, von KN, Feufel M. Accuracy of online symptom-assessment applications, large language models, and laypeople for self-triage decisions: a systematic review. medRxiv. Preprint posted online on September 14, 2024. [doi: [10.1101/2024.09.13.24313657](https://doi.org/10.1101/2024.09.13.24313657)]
21. Haltaufderheide J, Ranisch R. The ethics of chatGPT in medicine and healthcare: a systematic review on large language models (LLMs). *NPJ Digit Med.* 2024;7(1):183. [FREE Full text] [doi: [10.1038/s41746-024-01157-x](https://doi.org/10.1038/s41746-024-01157-x)] [Medline: [38977771](https://pubmed.ncbi.nlm.nih.gov/38977771/)]
22. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* 1989;13(3):319-340. [doi: [10.2307/249008](https://doi.org/10.2307/249008)]
23. Gulati S, Sousa S, Lamas D. Design, development and evaluation of a human-computer trust scale. *Behav Inf Technol.* 2019;38(10):1004-1015. [doi: [10.1080/0144929x.2019.1656779](https://doi.org/10.1080/0144929x.2019.1656779)]
24. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J Pers Soc Psychol.* 1988;54(6):1063-1070. [doi: [10.1037//0022-3514.54.6.1063](https://doi.org/10.1037//0022-3514.54.6.1063)] [Medline: [3397865](https://pubmed.ncbi.nlm.nih.gov/3397865/)]
25. Li X, Sung Y. Anthropomorphism brings us closer: the mediating role of psychological distance in user-AI assistant interactions. *Comput Hum Behav.* 2021;118:106680. [doi: [10.1016/j.chb.2021.106680](https://doi.org/10.1016/j.chb.2021.106680)]
26. Brown RJ, Skelly N, Chew-Graham CA. Online health research and health anxiety: a systematic review and conceptual integration. *Clin Psychol: Sci Pract.* 2020;27(2):e12299. [doi: [10.1111/cpsp.12299](https://doi.org/10.1111/cpsp.12299)]
27. He Z, Bhasuran B, Jin Q, Tian S, Hanna K, Shavor C, et al. Quality of answers of generative large language models versus peer users for interpreting laboratory test results for lay patients: evaluation study. *J Med Internet Res.* 2024;26:e56655. [FREE Full text] [doi: [10.2196/56655](https://doi.org/10.2196/56655)] [Medline: [38630520](https://pubmed.ncbi.nlm.nih.gov/38630520/)]
28. Guo Z, Lai A, Thygesen JH, Farrington J, Keen T, Li K. Large language models for mental health applications: systematic review. *JMIR Ment Health.* 2024;11:e57400. [FREE Full text] [doi: [10.2196/57400](https://doi.org/10.2196/57400)] [Medline: [39423368](https://pubmed.ncbi.nlm.nih.gov/39423368/)]
29. Ahmad MA, Yaramis I, Roy TD. Creating trustworthy LLMs: dealing with hallucinations in healthcare AI. arXiv:2311.01463. Preprint posted online on September 26, 2023. [doi: [10.48550/arXiv.2311.01463](https://doi.org/10.48550/arXiv.2311.01463)]
30. Carrasco-Farré C. Large language models are as persuasive as humans, but how. About the cognitive effort and moral-emotional language of LLM arguments. arXiv:2404.09329. Preprint posted online on April 14, 2024. 2024. [doi: [10.48550/arXiv.2404.09329](https://doi.org/10.48550/arXiv.2404.09329)]
31. Sbaifi L, Rowley J. Trust and credibility in web-based health information: a review and agenda for future research. *J Med Internet Res.* 2017;19(6):e218. [FREE Full text] [doi: [10.2196/jmir.7579](https://doi.org/10.2196/jmir.7579)] [Medline: [28630033](https://pubmed.ncbi.nlm.nih.gov/28630033/)]
32. Lu X, Zhang R, Wu W, Shang X, Liu M. Relationship between internet health information and patient compliance based on trust: empirical study. *J Med Internet Res.* 2018;20(8):e253. [FREE Full text] [doi: [10.2196/jmir.9364](https://doi.org/10.2196/jmir.9364)] [Medline: [30120087](https://pubmed.ncbi.nlm.nih.gov/30120087/)]
33. Mendel T, Nov O, Wiesenfeld B. Advice from a doctor or AI? Understanding willingness to disclose information through remote patient monitoring to receive health advice. *Proc ACM Hum-Comput Interact.* 2024;8(CSCW2):1-34. [doi: [10.1145/3686925](https://doi.org/10.1145/3686925)]
34. GPT-4 technical report. OpenAI. 2023. URL: <https://cdn.openai.com/papers/gpt-4.pdf> [accessed 2025-01-10]

35. Fogg B, Marshall J, Laraki O, Osipovich A, Varma C, Fang N, et al. What makes web sites credible?: a report on a large quantitative study. 2001. Presented at: CHI '01: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2001 March 01:61-68; Seattle Washington USA. [doi: [10.1145/365024.365037](https://doi.org/10.1145/365024.365037)]

Abbreviations

LLM: large language model

OR: odds ratio

Edited by A Mavragani; submitted 14.07.24; peer-reviewed by K You, R Keshava Murthy; comments to author 27.09.24; revised version received 11.11.24; accepted 25.12.24; published 13.02.25

Please cite as:

Mendel T, Singh N, Mann DM, Wiesenfeld B, Nov O

Laypeople's Use of and Attitudes Toward Large Language Models and Search Engines for Health Queries: Survey Study

J Med Internet Res 2025;27:e64290

URL: <https://www.jmir.org/2025/1/e64290>

doi: [10.2196/64290](https://doi.org/10.2196/64290)

PMID: [39946180](https://pubmed.ncbi.nlm.nih.gov/39946180/)

©Tamir Mendel, Nina Singh, Devin M Mann, Batia Wiesenfeld, Oded Nov. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 13.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.