



International Journal of Innovative Technologies in Social Science

e-ISSN: 2544-9435

Operating Publisher
SciFormat Publishing Inc.
ISNI: 0000 0005 1449 8214

2734 17 Avenue SW,
Calgary, Alberta, T3E0A7,
Canada
+15878858911
editorial-office@sciformat.ca

ARTICLE TITLE	TRUST, PRIVACY, AND LIABILITY: EXPLORING PATIENT ATTITUDES TOWARD ARTIFICIAL INTELLIGENCE IN HEALTHCARE
----------------------	---

DOI	https://doi.org/10.31435/ijitss.3(51).2026.6315
------------	---

RECEIVED	12 June 2026
-----------------	--------------

ACCEPTED	08 September 2026
-----------------	-------------------

PUBLISHED	17 September 2026
------------------	-------------------

LICENSE



The article is licensed under a **Creative Commons Attribution 4.0 International License**.

© The author(s) 2026.

This article is published as open access under the Creative Commons Attribution 4.0 International License (CC BY 4.0), allowing the author to retain copyright. The CC BY 4.0 License permits the content to be copied, adapted, displayed, distributed, republished, or reused for any purpose, including adaptation and commercial use, as long as proper attribution is provided.

TRUST, PRIVACY, AND LIABILITY: EXPLORING PATIENT ATTITUDES TOWARD ARTIFICIAL INTELLIGENCE IN HEALTHCARE

Kinga Głodek (Corresponding Author, Email: kingaglodek@icloud.com)

Medical University of Lublin, Lublin, Poland

ORCID ID: 0009-0006-8126-0702

Monika Kamińska

Medical University of Lublin, Lublin, Poland

ORCID ID: 0009-0005-6138-708X

Aleksandra Duda

Medical University of Lublin, Lublin, Poland

ORCID ID: 0009-0006-7081-3738

ABSTRACT

Background: The aim of this study was to evaluate the level of trust and social acceptance regarding the use of artificial intelligence (AI) in medical self-diagnosis, and to identify patients' primary concerns related to the implementation of AI algorithms in the healthcare system.

Materials and Methods: The research was carried out in July 2026 via a cross-sectional survey design utilizing an anonymous online questionnaire (the CAWI technique). The study cohort comprised 274 participants. The analysis evaluated digital health information-seeking behaviors, the acceptance of artificial intelligence in clinical and diagnostic workflows, and patients' perceptions of responsibility for medical errors caused by AI systems.

Results: Nearly all participants (96%) reported verifying their medical symptoms online. Patients demonstrated moderate acceptance of AI as a supportive tool in analyzing medical imaging; however, they strongly rejected (80% of respondents) the concept of fully automated treatment processes without physician supervision. Significant concerns regarding the privacy of medical data were identified, particularly within the 36–50 age demographic. The analysis of liability for AI-generated errors revealed that patients tend to attribute responsibility to themselves, highlighting a discrepancy between their moral intuition and current legal and ethical frameworks.

Conclusions: Public acceptance of AI in healthcare is strictly dependent on preserving the "human-in-the-loop" paradigm. There is a critical need to enhance health education regarding cybersecurity and to establish clear regulatory frameworks for AI-related errors, all while maintaining the physician's overriding authority in the therapeutic process.

KEYWORDS

Artificial Intelligence, Patient Perspectives, Healthcare Digitalization, Medical Ethics, Trust in AI

CITATION

Głodek, K., Kamińska, M., & Duda, A. (2026). Trust, privacy, and liability: exploring patient attitudes toward artificial intelligence in healthcare. *International Journal of Innovative Technologies in Social Science*, 3(51). [https://doi.org/10.31435/ijitss.3\(51\).2026.6315](https://doi.org/10.31435/ijitss.3(51).2026.6315)

COPYRIGHT

© The author(s) 2026. This article is published as open access under the Creative Commons Attribution 4.0 International License (CC BY 4.0), allowing the author to retain copyright. The CC BY 4.0 License permits the content to be copied, adapted, displayed, distributed, republished, or reused for any purpose, including adaptation and commercial use, as long as proper attribution is provided.

Introduction

Recent years have witnessed an unprecedented development of artificial intelligence (AI) technologies, including advanced large language models (LLMs) [1]. Their implementation extends across a growing number of areas of social life, including the healthcare sector, as evidenced by the high performance of algorithms on standardized medical examinations [2]. Historically, the digitalization of medicine has primarily relied on tools designed to assist medical professionals. Currently, however, a clear paradigm shift is underway: patients are gaining direct access to advanced algorithms, which opens a new chapter in medical sociology and the patient-technology relationship [3].

The phenomenon of independent online medical self-diagnosis, colloquially referred to as "Dr. Google," frequently leads to the development of cyberchondria, a heightened health anxiety driven by the obsessive and uncontrollable searching for symptom information [4]. Currently, this trend is evolving toward interactions with a "virtual physician." Chatbots and AI-based systems are capable of analyzing patient-entered symptoms in real time, suggesting potential medical conditions, and recommending further courses of action [5,6]. While these technologies hold immense potential to alleviate the burden on healthcare systems and equalize access to health education, their public acceptance remains a subject of debate [7]. Trusting machines, particularly in the highly sensitive domain of human health and life, encounters numerous psychological and ethical barriers. Issues such as the fear of misdiagnosis, the dehumanization of the treatment process, algorithmic objectivity, and liability for potential medical errors currently constitute key challenges for researchers in the social sciences and innovative technologies [8,9].

The aim of this study was to evaluate the level of trust and public acceptance regarding the use of artificial intelligence in medical self-diagnosis. The research also aimed to identify key patient concerns associated with the implementation of AI-based solutions into the mainstream healthcare system.

Methodology

Data for this cross-sectional study were collected in July 2026. Given its strictly observational nature, the research did not constitute a medical experiment. Data collection was performed utilizing a voluntary, fully anonymous online questionnaire based on the Computer-Assisted Web Interview (CAWI) methodology.

The primary research instrument was a custom-designed questionnaire administered via Google Forms. Data collection utilized a snowball sampling technique for online distribution across social media platforms and discussion forums. The survey comprised four distinct thematic sections. The first section gathered sociodemographic data to establish the baseline characteristics of the study sample. The second part explored participants' prior experiences with digital health information-seeking and their interactions with LLMs (e.g., ChatGPT) for medical inquiries. The third section assessed the respondents' trust and acceptance of AI algorithms using a standard 5-point Likert scale (ranging from 1 – "Strongly disagree" to 5 – "Strongly agree"). Finally, the fourth section investigated perceptions of security and patient concerns regarding the privacy of sensitive health data.

Ultimately, the study cohort comprised 274 participants ($N = 274$). The analyzed sample revealed a distinct predominance of women, who accounted for 73.7% ($n = 202$) of all respondents, while men constituted 25.6% ($n = 70$), and 0.7% ($n = 2$) preferred not to disclose their gender. Given the significant overrepresentation of female respondents, a detailed gender-based comparative analysis was omitted. Instead, the focus was directed toward the age variable, which exhibited a more balanced distribution across the study group. The raw data obtained from the questionnaire were analyzed using spreadsheet software and the survey platform's built-in analytical tools, with the findings presented descriptively and graphically.

Results

A total of 274 respondents ($N = 274$) participated in the study. Data collected via the questionnaire underwent quantitative analysis to assess patients' attitudes toward the integration of artificial intelligence (AI) into the diagnostic process. The subsequent findings outline the demographic profile of the study cohort, their habits regarding independent online medical self-diagnosis, the level of acceptance of AI algorithms in clinical workflows, and perceptions of liability for potential diagnostic errors. A detailed analysis of the collected research material is presented below.

Initially, patients' habits regarding independent medical information-seeking in the digital environment were analyzed. The frequency with which respondents use the Internet to verify disease symptoms prior to a medical consultation is presented in **Figure 1**.

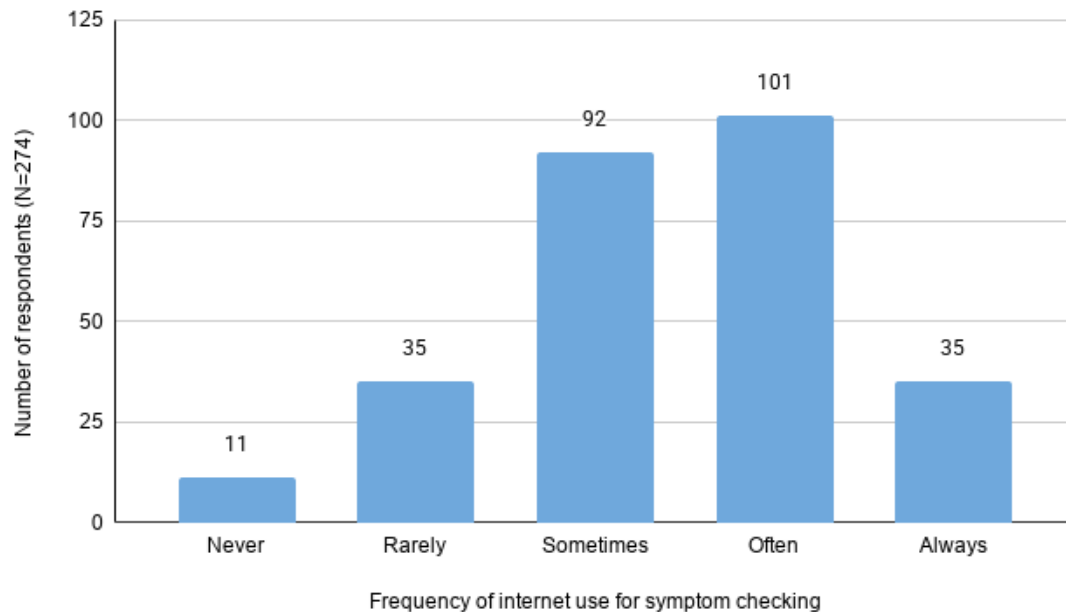


Fig. 1. Frequency of online medical information-seeking for symptom verification prior to medical consultations (N = 274).

The analyzed data indicate the widespread nature of online health information-seeking in the study cohort, with almost all respondents declaring that they verify their symptoms on the Internet. The largest subgroups consist of those who do so "often" ($n = 101$) and "sometimes" ($n = 92$), representing a combined 70% of the sample. It is noteworthy that the percentage of individuals who never undertake such searches is negligible, standing at just 4% ($n = 11$). These results corroborate that access to digital medical data is now an integral part of pre-visit patient behavior, potentially exerting a significant influence on the physician-patient relationship and the clinical communication process.

The next stage of the analysis assessed patients' acceptance of utilizing AI technologies in clinical treatment. Findings detailing participants' comfort with a treatment plan generated solely by an AI algorithm are illustrated in **Figure 2**.

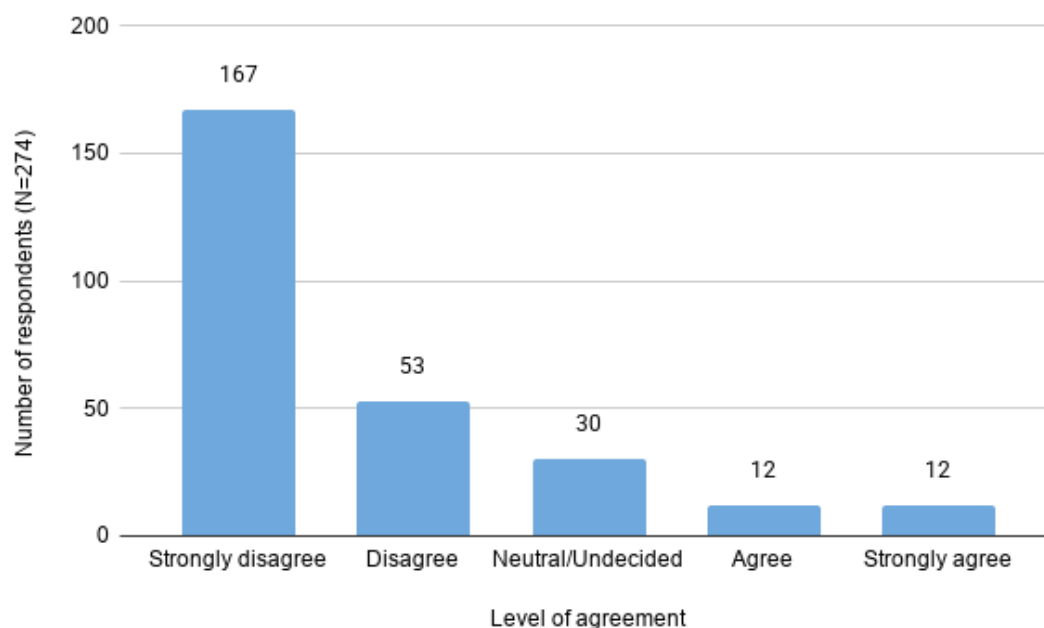


Fig. 2. Patient comfort levels with a treatment plan proposed exclusively by an artificial intelligence algorithm (N = 274).

The analysis of the data reveals a substantial reluctance among respondents toward the full automation of the treatment process. The prevailing attitude is a distinct lack of acceptance, reflected in the figures: 60.9% of respondents ($n = 167$) selected "strongly disagree," and 19.3% ($n = 53$) answered "disagree." Collectively, 80% of the surveyed individuals expressed a negative stance toward this solution. These findings suggest that, from the patients' perspective, the role of AI in medicine should remain restricted to diagnostic support, whereas final therapeutic decisions strictly require human oversight.

Another area subject to analysis was patients' readiness to accept the use of AI algorithms for analyzing the results of medical imaging, such as magnetic resonance imaging (MRI) or ultrasound (USG). The findings regarding the level of acceptance for this solution, which aims to expedite the diagnostic process, are presented in **Figure 3**.

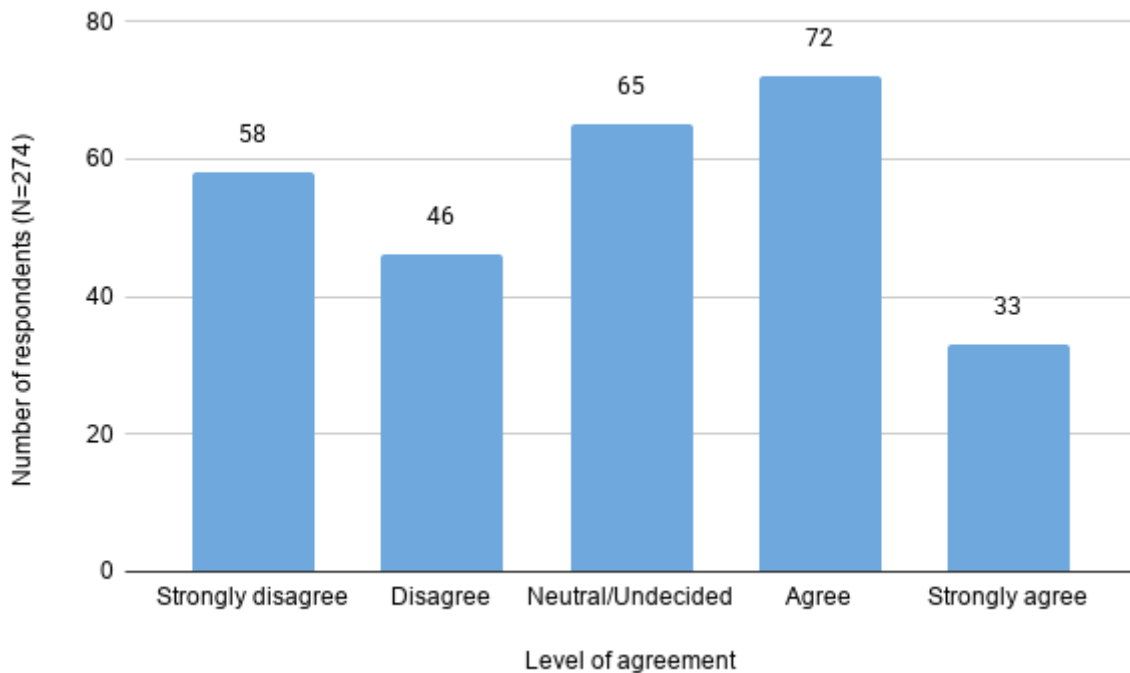


Fig. 3. Acceptance levels for the analysis of medical test results by artificial intelligence systems to expedite the diagnostic process ($N = 274$).

The obtained data reveal a more nuanced patient attitude toward AI technology in technical diagnostics compared to therapeutic decision-making. The highest frequencies were recorded in the "Agree" ($n = 72$) and "Neutral/Undecided" ($n = 65$) categories, which cumulatively account for half of the surveyed population. It is worth noting that while respondents maintain a degree of reservation regarding AI in clinical treatment, they exhibit greater receptivity toward its use in imaging analysis, recognizing the technology as a means to enhance diagnostic efficiency. This highlights the potential of AI as a technical support tool in the radiologist's workflow, provided that the physician's pivotal role in the final interpretation is preserved.

Furthermore, the analysis investigated patient apprehensions concerning data privacy amidst the integration of AI, taking age-related disparities into account. A breakdown of the respondents' perspectives across different age cohorts is illustrated in **Figure 4**.

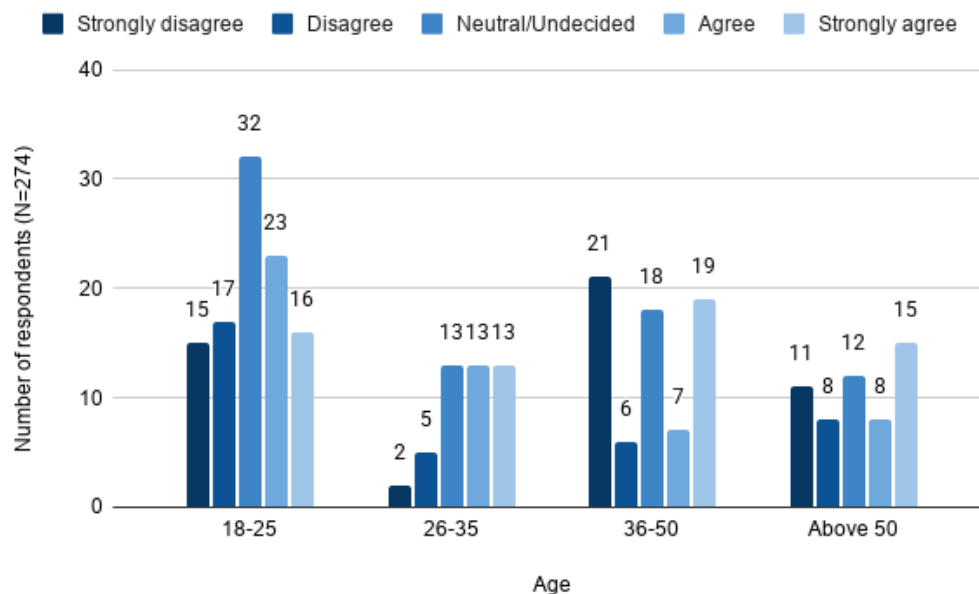


Fig. 4. Respondent concerns regarding medical data security upon the implementation of AI solutions, stratified by age group ($N = 274$).

Data analysis reveals varying attitudes toward the risk of medical data privacy loss, contingent upon the respondents' age. Within the 18–25 age group, “Neutral/Undecided” responses predominate ($n = 32$), followed most frequently by “Agree” ($n = 23$). In the 26–35 cohort, the highest frequencies are shared equally among “Neutral/Undecided”, “Agree,” and “Strongly Agree” responses ($n = 13$ each). Only a small minority expressed a dissenting stance ($n = 7$), highlighting a distinct subset of individuals with acute concerns regarding this threat. Conversely, in the 36–50 and over-50 age brackets, the response distribution is highly heterogeneous, lacking a single dominant attitude and resulting in dispersed values across all categories. This variance indicates the absence of a unified perception of cybersecurity within the study population, underscoring the necessity of tailoring data protection communication strategies to specific age groups.

The final area of investigation addressed the attribution of legal and ethical liability in the event of a diagnostic error generated by artificial intelligence systems. The results indicating respondents' preferences in this regard, stratified by age group, are presented in **Figure 5**.

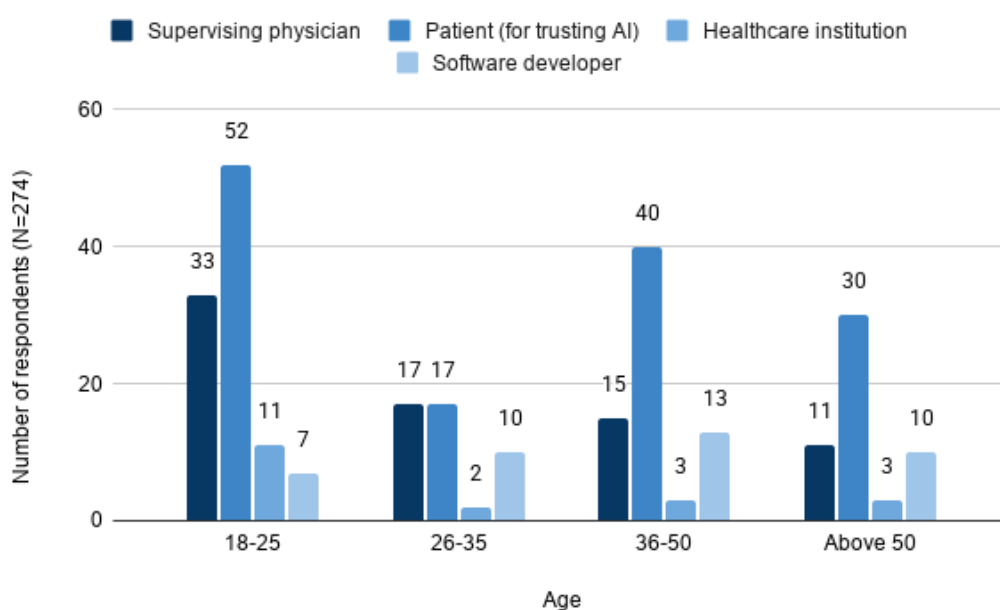


Fig. 5. Respondent attribution of liability for AI-generated diagnostic errors, stratified by age group ($N = 274$).

Data analysis reveals minor variations in the perception of liability for diagnostic errors across different age cohorts. In the 18–25 ($n = 52$ for the patient, $n = 33$ for the physician), 36–50 ($n = 40$ for the patient, $n = 15$ for the physician), and above 50 ($n = 30$ for the patient, $n = 11$ for the physician) age groups, the patient is most frequently identified as the responsible party, followed by the supervising physician. In the 26–35 cohort, the patient ($n = 17$) and the supervising physician ($n = 17$) are held equally accountable. The marginal attribution of blame to medical institutions across all age demographics suggests that, in the respondents' view, liability for the functioning of AI technologies in medicine lies predominantly with end-users and clinicians, rather than with organizational entities.

Discussion

The results of this study indicate a complex and multidimensional patient attitude toward the implementation of AI in healthcare. The ubiquity of online symptom verification (over 70% of respondents) demonstrates a high level of digital activity among patients. This phenomenon is widely reflected in the research by Lupton and Jutel [10], who argue that the growing availability of mobile diagnostic applications is irreversibly altering the power dynamics in medicine. This shift positions the patient as an active, albeit frequently vulnerable to misinformation, participant in the clinical process. Despite this openness to information technologies, the full implementation of AI algorithms in medical care encounters distinct social resistance, the intensity of which strictly depends on the proposed application of the algorithm itself.

A central finding of the conducted analyses is the pronounced dissonance between the acceptance of AI in the diagnostic process and its rejection in the therapeutic setting. While a substantial proportion of respondents exhibit a neutral or positive attitude toward the use of algorithms in medical imaging analysis to expedite diagnosis, the full automation of treatment planning faces non-acceptance from 80% of the surveyed cohort. These observations are highly consistent with the survey results of Lennartz et al. [11], who demonstrated that patients most readily accept AI in administrative and analytical roles during the initial stages of diagnosis, yet categorically demand that physicians retain control over final therapeutic decisions. These conclusions are further corroborated by a comprehensive scoping review [12], which indicates that patient trust in AI remains strictly contingent upon ultimate human decision-making, aligning with the "human-in-the-loop" paradigm. As aptly pointed out by Amann et al. [13], a primary barrier here is the lack of algorithmic explainability (Explainable AI). Patients are willing to accept the impenetrable nature of the "black box" when the algorithm serves merely to preselect imaging data; however, they demand complete transparency when it comes to therapeutic interventions. Furthermore, He et al. [14] emphasize that a seamless transition from diagnostics to the implementation of AI-driven therapy will require the design of systems capable of justifying their decisions in a clear, human-comprehensible manner.

Another crucial implementation barrier identified in our study involves concerns regarding medical data privacy and security. The highest level of apprehension was recorded in the 36–50 age bracket, which may stem from a greater awareness of cyber threats among adults compared to the youngest cohort. This finding strongly correlates with the reports by Esmailzadeh [15], who demonstrated that the fear of sensitive health information leaks constitutes one of the most significant factors impeding consumer trust in digital medicine. This challenge is analyzed at a systemic level by Kelly et al. [16], who indicate that gaps in data governance and the absence of standardized anonymization protocols for patient information currently act as the primary "bottleneck" limiting the clinical impact of AI. Mitigating these concerns will therefore necessitate the transparent implementation of rigorous cybersecurity standards across healthcare institutions.

The evaluation of perceived liability for AI-generated medical errors provides highly intriguing and novel conclusions. Within the current literature, the bioethical discourse, as comprehensively examined by Vayena et al. [17], focuses on the apportionment of legal and moral accountability among the supervising clinician, the software developer, and the healthcare institution. In stark contrast, our findings indicate that respondents most predominantly designated the patient themselves as the liable entity for any occurring error, predicated on their provision of consent and entrustment of the machine. From a clinical standpoint, this scenario is additionally confounded by the assertion of Grote and Berens [18] that responsibility for algorithmic decision-making cannot exist in an ethical vacuum. It is imperative that artificial intelligence systems strictly adhere to and are grounded in officially sanctioned clinical guidelines. The final verification of the algorithm's recommendations against established normative standards will perpetually fall upon the supervising physician, a reality that, under current law, prohibits the full automation of liability. The profound dichotomy between

theoretical medico-legal frameworks and the empirical convictions of patients uncovered in this research constitutes a paramount challenge for forthcoming health literacy initiatives.

Finally, a limitation of this study is the gender imbalance within the cohort, characterized by a significant overrepresentation of female respondents (73.7%). While this is a common occurrence in online survey-based research, it may influence the generalizability of the findings, particularly regarding perceptions of digital health technologies and data privacy concerns. Future research should prioritize achieving a more balanced gender distribution to ascertain whether these attitudes differ significantly across gender lines.

Conclusions

The conducted study highlights the complex nature of societal attitudes toward the implementation of artificial intelligence in the healthcare system. Despite the high level of patient digitalization and the widespread use of Internet resources for self-diagnosis, the full automation of medical processes encounters distinct resistance. The results unequivocally indicate a polarization of trust: patients accept AI as a useful tool supporting the analysis of medical tests (e.g., imaging), yet they categorically reject the concept of autonomous treatment planning by algorithms, demanding continuous medical supervision within the "human-in-the-loop" paradigm.

Further critical barriers to implementation include apprehensions regarding the cybersecurity of sensitive medical data, particularly prominent among middle-aged adults, alongside uncertainties concerning the assignment of liability for diagnostic errors. The finding that respondents primarily held themselves accountable for AI-induced mistakes underscores an urgent requirement for educational strategies aimed at clearly defining and safeguarding patient rights within the digital transformation of healthcare.

The secure and ethical integration of algorithms into routine clinical workflows demands more than just comprehensive public education; it necessitates, fundamentally, rigorous procedural compliance. Artificial intelligence systems must operate unequivocally within the parameters of validated, standardized clinical protocols. Adhering to the principles of evidence-based medicine, the final clinical judgment, including the validation of algorithmic recommendations against these established norms, must invariably remain the responsibility of a qualified clinician.

Acknowledgments: The authors would like to express their gratitude to all the respondents who participated in the study and dedicated their time to complete the questionnaire. This research received no external financial support or grants.

REFERENCES

1. Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine. *Nature Medicine*, 29(8), 1930–1940. <https://doi.org/10.1038/s41591-023-02448-8>
2. Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, 2(2), e0000198. <https://doi.org/10.1371/journal.pdig.0000198>
3. Topol E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44–56. <https://doi.org/10.1038/s41591-018-0300-7>
4. Starcevic, V., & Berle, D. (2013). Cyberchondria: Towards a better understanding of excessive health-related Internet use. *Expert Review of Neurotherapeutics*, 13(2), 205–213. <https://doi.org/10.1586/ern.12.162>
5. Cocco, A. M., Zordan, R., Taylor, D. M., Weiland, T. J., Dilley, S. J., Kant, J., Dombagolla, M., Hendarto, A., Lai, F., & Hutton, J. (2018). Dr Google in the ED: Searching for online health information by adult emergency department patients. *The Medical Journal of Australia*, 209(8), 342–347. <https://doi.org/10.5694/mja17.00889>
6. Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A. Y. S., & Coiera, E. (2018). Conversational agents in healthcare: A systematic review. *Journal of The American Medical Informatics Association : JAMIA*, 25(9), 1248–1258. <https://doi.org/10.1093/jamia/ocy072>

7. Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2019). Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. *Digital Health*, 5, 2055207619871808. <https://doi.org/10.1177/2055207619871808>
8. Richardson, J.P., Smith, C., Curtis, S., Watson, S., Zhu, X., & Zhu, X. (2021). Patient apprehensions about the use of artificial intelligence in healthcare. *npj Digital Medicine* 4, 140. <https://doi.org/10.1038/s41746-021-00509-1>
9. World Health Organization (2021). Ethics and governance of artificial intelligence for health: WHO guidance. World Health Organization. <https://iris.who.int/handle/10665/341996>
10. Lupton, D., & Jutel, A. (2015). 'It's like having a physician in your pocket!' A critical analysis of self-diagnosis smartphone apps. *Social Science & Medicine* (1982), 133, 128–135. <https://doi.org/10.1016/j.socscimed.2015.04.004>
11. Lennartz, S., Dratsch, T., Zopfs, D., Persigehl, T., Maintz, D., Große Hokamp, N., & Pinto Dos Santos, D. (2021). Use and control of artificial intelligence in patients across the medical workflow: Single-center questionnaire study of patient perspectives. *Journal of Medical Internet Research*, 23(2), e24221. <https://doi.org/10.2196/24221>
12. Moy, S., Irannejad, M., Manning, S. J., Farahani, M., Ahmed, Y., Gao, E., Prabhune, R., Lorenz, S., Mirza, R., & Klinger, C. (2024). Patient perspectives on the use of artificial intelligence in health care: A scoping review. *Journal of Patient-Centered Research and Reviews*, 11(1), 51–62. <https://doi.org/10.17294/2330-0698.2029>
13. Amann, J., Blasimme, A., Vayena, E., Frey, D., Madai, V. I., & Precise4Q consortium (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, 20(1), 310. <https://doi.org/10.1186/s12911-020-01332-6>
14. He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine*, 25(1), 30–36. <https://doi.org/10.1038/s41591-018-0307-0>
15. Esmailzadeh P. (2020). Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. *BMC Medical Informatics and Decision Making*, 20(1), 170. <https://doi.org/10.1186/s12911-020-01191-1>
16. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17(1), 195. <https://doi.org/10.1186/s12916-019-1426-2>
17. Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLoS Medicine*, 15(11), e1002689. <https://doi.org/10.1371/journal.pmed.1002689>
18. Grote, T., & Berens, P. (2020). On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, 46(3), 205–211. <https://doi.org/10.1136/medethics-2019-105586>