Medicor : Journal of Health Informatics and Health Policy
E-ISSN : 3030-9166
Volume. 2 Issue 4 October 2024
Page No: 232-244

Balancing Innovation and Ethics: A Narrative Review of Artificial Intelligence in Healthcare

Resva Meinisasti
Poltekkes Kemenkes Bengkulu, Indonesia
Correspondent: resva@poltekkesbengkulu.ac.id

Received : August 25, 2024
Accepted : October 17, 2024
Published : October 31, 2024

Citation: Meinisasti, R. (2024). Balancing Innovation and Ethics: A Narrative Review of Artificial Intelligence in Healthcare. Medicor: Journal of Health Informatics and Health Policy, 2(4), 232-244.

https://doi.org/10.61978/medicor.v2i4.1079

ABSTRACT: Artificial intelligence (AI) is increasingly applied in healthcare, improving diagnostic accuracy, personalized treatment, and data management. However, its adoption raises ethical challenges related to patient privacy, algorithmic bias, clinical autonomy, and governance. This narrative review synthesizes peer-reviewed studies (2020–2025) identified through Scopus, PubMed, Google Scholar, and Web of Science. Findings indicate persistent risks of data breaches, algorithmic inequities, and loss of clinical autonomy. While techniques such as differential privacy and explainable AI offer solutions, their implementation remains uneven. Effective governance requires multi-stakeholder engagement and strong regulation. We conclude that responsible AI integration depends on transparent governance and inclusive model development to ensure equitable and trustworthy healthcare outcomes.

Keywords: Artificial Intelligence in Healthcare, Ethical Considerations, Data Privacy in Health Systems, Algorithmic Bias in Medical AI, Explainable AI in Healthcare, Governance and Accountability in AI, Global Health Equity.

This is an open access article under the CC-BY 4.0 license

INTRODUCTION

The integration of artificial intelligence (AI) into healthcare has emerged as one of the most transformative developments in contemporary medical science. Recent advances in AI technologies, including machine learning, natural language processing, and deep learning, have been widely recognized for their capacity to enhance clinical decision-making, optimize resource allocation, and improve patient outcomes (Amini et al., 2023; Das & Khatua, 2024). These technologies are increasingly applied across multiple domains, ranging from diagnostic imaging and personalized treatment planning to predictive analytics in population health. However, as AI systems become more deeply embedded in healthcare infrastructures, the ethical implications of their adoption have garnered heightened scrutiny. Ethical concerns, particularly those relating to data privacy, algorithmic transparency, and bias, have the potential to influence both the effectiveness of these systems and the trust patients and healthcare professionals place in them (Badawy et al., 2024; Graaf, 2025; Sargsyan et al., 2024). As such, navigating the balance between technological innovation and ethical responsibility has become a pressing challenge in global healthcare systems.

The promise of AI lies in its ability to revolutionize healthcare delivery by offering diagnostic precision, scalable solutions for underserved populations, and efficient management of increasingly complex medical data. For instance, AI-enabled imaging tools have demonstrated remarkable accuracy in detecting early-stage cancers, while predictive analytics have shown potential in anticipating disease progression and tailoring interventions (Bandekar et al., 2024). Furthermore, AI has proven valuable in streamlining administrative processes, such as electronic health record (EHR) management, thereby reducing the workload of healthcare providers and improving operational efficiency. Yet, despite these advantages, the ethical debates surrounding AI adoption remain unresolved. Scholars and policymakers alike stress the importance of developing comprehensive ethical frameworks that can safeguard patient rights, promote fairness, and ensure accountability in AI-driven healthcare environments (Amini et al., 2023; Das & Khatua, 2024).

A critical issue in the ethical deployment of AI in healthcare relates to patient data privacy. AI systems typically rely on large datasets for training and validation, necessitating the use of sensitive health information. While this reliance enables more robust and accurate models, it simultaneously raises concerns regarding the security and protection of patient data (Mennella et al., 2024). Regulations such as the General Data Protection Regulation (GDPR) have been introduced to address these concerns, mandating strict standards for data collection, storage, and use (Bandekar et al., 2024). However, compliance with such frameworks remains uneven across institutions and jurisdictions, resulting in varying levels of protection for patient data. High-profile data breaches and unauthorized access to medical records have amplified public concerns and highlighted the risks of inadequate governance structures (Sargsyan et al., 2024). These challenges underscore the need for robust data management practices and heightened awareness among healthcare stakeholders about the ethical dimensions of data use.

Algorithmic bias represents another pressing ethical challenge. Bias may arise from multiple sources, including unrepresentative training datasets, flawed algorithmic design, and errors in interpreting AI outputs (Jones, 2025). Such biases can disproportionately impact marginalized groups, exacerbating existing inequities in healthcare access and outcomes. For example, models trained primarily on data from high-income countries may underperform when applied in low- and middle-income contexts, leading to misdiagnoses or inadequate treatment recommendations (Pham, 2025). Similarly, demographic imbalances in datasets can result in tools that perform poorly for women, ethnic minorities, or elderly populations, thereby reinforcing systemic disparities (Badawy et al., 2024). Addressing these challenges requires an emphasis on algorithmic transparency, inclusivity in data collection, and the active participation of diverse stakeholders in AI system development (Graaf, 2025).

In addition to bias, the lack of transparency in AI decision-making has been a source of widespread concern. Many advanced AI systems operate as “black boxes,” generating outputs without providing interpretable reasoning or justification (Karra & RamaRao, 2025). This opacity complicates the ability of healthcare professionals to understand, trust, and explain AI-driven recommendations to patients, potentially undermining confidence in clinical settings (Amini et al., 2023). Moreover, the absence of accountability mechanisms in cases of AI error or harm raises questions about liability, further complicating the integration of AI technologies into healthcare workflows. These issues point to the urgent need for explainable AI (XAI) models that can enhance interpretability while maintaining accuracy, thereby bridging the gap between technological performance and ethical responsibility.

Despite growing recognition of these challenges, significant gaps remain in the literature on AI ethics in healthcare. Much of the existing research emphasizes technical efficacy—accuracy rates, predictive performance, and efficiency—while underexploring the broader social, cultural, and ethical consequences of AI adoption (Pham, 2025). Studies have also tended to overlook the ways in which AI may reinforce existing inequities, particularly for vulnerable populations with limited access to healthcare infrastructure (Amann et al., 2020). Furthermore, there is a notable lack of consensus on the development of practical ethical frameworks or guidelines to govern the use of patient data, ensure fairness in algorithmic outcomes, and establish transparent accountability structures (Amini et al., 2023; Giansanti & Pirrera, 2025; Perez et al., 2025). These gaps highlight the necessity for research that critically examines the intersection of technology, ethics, and healthcare practice.

The primary aim of this review is to explore the ethical considerations associated with the integration of AI into healthcare, with a particular focus on issues of privacy and bias. Specifically, this paper seeks to analyze how data governance, algorithmic design, and systemic inequities intersect to shape the ethical landscape of AI-driven healthcare. By synthesizing recent findings, this study intends to identify key challenges, highlight areas requiring further inquiry, and propose strategies for mitigating ethical risks while maximizing the benefits of AI technologies. In doing so, this review contributes to a growing body of literature that calls for a balanced and responsible approach to AI adoption in healthcare.

The scope of this review encompasses both high-income and low- and middle-income countries, acknowledging the diverse contexts in which AI technologies are deployed. Comparative analysis across geographic and cultural settings is critical, as ethical challenges and regulatory approaches vary widely depending on local infrastructure, governance systems, and sociocultural values (Perez et al., 2025). In resource-rich environments, concerns may center on algorithmic opacity and liability, while in resource-limited settings, issues such as inadequate datasets, infrastructural constraints, and unequal access to technology may dominate (García-Saisó et al., 2024). This review therefore considers global perspectives, with particular attention to the ethical implications of AI deployment in marginalized communities, rural areas, and regions where healthcare resources remain scarce. By situating the analysis within this broad scope, the paper aims to provide a holistic understanding of the ethical dilemmas surrounding AI in healthcare and to propose pathways toward equitable and responsible technological integration.

METHOD

The methodology employed in this study was designed to ensure a rigorous and systematic approach to the identification, selection, and analysis of literature relevant to ethical considerations in AI-driven healthcare. Given the complex and multidisciplinary nature of the topic, the review drew upon multiple databases to capture a comprehensive body of evidence. Four primary databases were utilized: Scopus, PubMed, Google Scholar, and Web of Science. Each of these databases was chosen for its unique strengths and coverage. Scopus was selected for its extensive bibliographic scope and analytical tools, which allow for the assessment of citation impact across disciplines, including healthcare and technology. PubMed was deemed essential due to its focus on biomedical and clinical research, ensuring inclusion of studies directly addressing AI applications in health contexts. Google Scholar, despite being less structured, provided broad access to a wide variety of sources, including theses, technical reports, and grey literature, which were considered valuable for a comprehensive review. Web of Science was also included as it provides extensive citation tracking and indexing of international journals, complementing the coverage of Scopus while ensuring multidisciplinary relevance.

The search strategy was constructed around carefully selected keywords that aligned with the scope of the study. Keywords such as "Artificial Intelligence in Healthcare Ethics," "Data Privacy in AI Healthcare," "Algorithmic Bias in Medical AI," "Explainable AI in Healthcare," "Healthcare Accountability and AI," "AI Applications in Health Systems," and "Ethical Considerations in AI Medicine" were used as the foundation for the searches. These terms were chosen to reflect the key thematic areas of ethical concern within AI-driven healthcare. Boolean operators were employed strategically to refine search outcomes. For example, combinations such as "AI in healthcare AND data privacy" or "algorithmic bias OR ethical considerations" were applied to expand or narrow results depending on the specific aspect of the inquiry. This ensured that the search process was both inclusive of broad discussions and precise in targeting literature directly relevant to privacy and bias issues in AI healthcare.

To maintain clarity and rigor in selection, the review established explicit inclusion and exclusion criteria. Studies published between 2020 and 2025 were prioritized to capture the most current debates and technological developments in AI-driven healthcare. Articles were required to be published in peer-reviewed journals, ensuring that only validated and high-quality research informed the analysis. The inclusion criteria further specified that studies must explicitly address AI applications in healthcare with a focus on ethical considerations such as privacy, bias, transparency, and accountability. Excluded from the review were publications that focused exclusively on technical performance metrics of AI systems without addressing ethical or social implications. Non-English articles were also excluded due to limitations in translation resources and to maintain consistency in the analysis.

The review incorporated a diverse range of study designs to capture multiple dimensions of the topic. These included randomized controlled trials, cohort studies, and case studies that explored AI applications in clinical and healthcare settings. Narrative reviews and systematic reviews were also included, given their relevance in synthesizing broader discussions of ethics and policy implications in AI-driven healthcare. Grey literature, such as government and organizational reports, was considered when it provided critical insights into emerging frameworks, guidelines, or case-specific ethical challenges that may not yet have appeared in peer-reviewed outlets.

The process of literature selection involved several steps to ensure transparency and replicability. Initially, all search results were compiled and duplicates removed. Titles and abstracts were then screened for relevance according to the inclusion and exclusion criteria. This stage eliminated studies that did not directly address ethical issues within AI healthcare. Articles that passed this stage underwent full-text review, during which methodological quality, thematic relevance, and depth of ethical analysis were evaluated. Studies deemed insufficiently rigorous or too narrowly focused on technical aspects without ethical discussion were excluded at this stage. The screening and evaluation process was conducted by multiple reviewers to minimize bias and ensure inter-rater reliability. Any discrepancies in selection decisions were resolved through discussion and consensus, further reinforcing the robustness of the methodology.

Data extraction from the selected studies followed a structured approach. Each study was analyzed for core elements including research objectives, methodological design, geographic focus, key findings related to privacy and bias, and the ethical frameworks or recommendations proposed. The extracted data were then synthesized thematically, allowing for the identification of recurrent issues, emerging challenges, and areas of consensus or contention across the literature. This thematic synthesis provided the basis for the subsequent analysis and discussion sections of the study, ensuring that the conclusions were grounded in a broad and balanced body of evidence.

Overall, the methodology adopted in this review ensured that the literature search and analysis were comprehensive, systematic, and aligned with academic standards for conducting narrative reviews in health-related fields. By combining multiple databases, applying carefully constructed search strategies, and adhering to clear inclusion and exclusion criteria, the study established a robust foundation for examining ethical considerations in AI-driven healthcare. This methodological rigor enhances the reliability of the findings and supports the relevance of the review to both academic and policy-oriented audiences.

RESULT AND DISCUSSION

Data Privacy

The findings of the literature consistently highlight that patient data privacy is among the most pressing concerns associated with the integration of AI into healthcare. Numerous studies confirm that sensitive health data is highly vulnerable to breaches, unauthorized access, and misuse when managed by AI-driven systems (Amini et al., 2023; Radanliev, 2025). The risk of privacy violations grows in proportion to the volume of data required for training AI models, with large-scale datasets increasing the likelihood of exposure to third-party access without patient consent (McGrath et al., 2025; Jaime et al., 2023). Reports of data leakage and unauthorized exploitation underscore the necessity of robust data security protocols to safeguard patient trust in AI-enabled systems. Evidence from international contexts demonstrates that patient trust is often eroded in settings where data governance practices are weak, with patients expressing reluctance to share sensitive information that could otherwise be valuable for improving diagnostic accuracy (Sargsyan et al., 2024; Mennella et al., 2024).

Differential privacy has emerged as a promising solution to this dilemma, offering ways for AI systems to learn from patient data without revealing specific individual details (Amini et al., 2023). By embedding mathematical noise into datasets, differential privacy ensures anonymity while preserving statistical utility, thus maintaining analytical validity. Despite its potential, implementation remains inconsistent across health systems. Countries with advanced data governance structures have integrated differential privacy into national strategies, embedding it within regulatory frameworks to align with policies such as the General Data Protection Regulation (GDPR) (Jaime et al., 2023; Radanliev, 2025). However, in low- and middle-income countries, the absence of robust infrastructures has slowed adoption, leaving patients more vulnerable to exploitation. These disparities highlight the global unevenness in protecting privacy, stressing the urgent need for harmonized frameworks that balance innovation with safeguarding individual rights.

Algorithmic Bias

Algorithmic bias emerged as another recurrent theme in the literature, with studies documenting how AI systems can perpetuate or exacerbate inequities in healthcare delivery. A particularly concerning example arises in oncology, where AI models developed for cancer detection show reduced accuracy for underrepresented racial and ethnic groups due to skewed training datasets (Jones, 2025). Such biases compromise diagnostic reliability, leading to misdiagnoses or inadequate treatment for marginalized populations. Similarly, in radiological imaging, algorithms often exhibit higher precision when analyzing images from populations well-represented in training datasets, while producing less accurate results for minority groups (Jones, 2025). These findings underscore the direct clinical risks posed by biased AI tools.

Cross-country comparisons further reveal that the severity and impact of algorithmic bias vary depending on the healthcare infrastructure and available datasets. Research in countries with limited healthcare resources indicates that AI models trained primarily in data-rich environments may fail to perform adequately when applied in resource-constrained settings, especially for chronic diseases requiring nuanced, context-specific data (Shah, 2024). These gaps limit the ability of AI to deliver equitable healthcare across diverse populations, reinforcing existing disparities in access and outcomes. Globally, the evidence calls for the design of inclusive and representative datasets, emphasizing the importance of collaborative, cross-cultural approaches to AI model development (Jones, 2025). Addressing these disparities is critical for ensuring fairness and accuracy in AI-driven healthcare interventions.

Human Oversight and Autonomy

The role of healthcare professionals in maintaining clinical autonomy remains central in the discourse on AI integration. Literature suggests that while AI-based clinical decision support systems provide valuable assistance, ultimate responsibility must rest with physicians and other medical practitioners. Amann et al. (2020) argue that explainable AI tools, which allow clinicians to interpret and understand algorithmic reasoning, are crucial in preserving trust and professional autonomy. By demystifying the decision-making process, such systems ensure that clinicians remain active participants in patient care rather than passive recipients of machine outputs.

Despite these benefits, concerns about over-reliance on AI are pervasive. Rony et al. (2025) caution that excessive dependence on AI risks "deskilling" medical professionals, undermining critical thinking skills that are essential for effective practice. The literature indicates that such reliance could diminish practitioners' ability to question AI outputs, leading to blind acceptance of potentially flawed recommendations. This is particularly evident in complex clinical domains such as surgery and mental health. In surgical contexts, AI-driven decision support has shown potential to enhance outcomes, yet limitations arise when algorithms fail to account for patient-specific conditions. Inaccuracies or oversights in algorithmic recommendations could jeopardize patient safety, underscoring the necessity of human oversight in final decision-making (Rad et al., 2025). In mental health, reliance on AI raises ethical concerns regarding the sensitivity of patient interactions. Systems lacking human empathy may inadvertently trivialize or misinterpret trauma, illustrating the limitations of purely technological solutions (Holohan & Fiske, 2021).

The global literature reveals a consensus that human judgment is irreplaceable in safeguarding the integrity of healthcare delivery. Active engagement of medical professionals in evaluating AI outputs is imperative to ensure that patient care remains both effective and ethically grounded. As such, the coexistence of AI and human expertise must be carefully managed to maximize benefits while avoiding the erosion of professional autonomy.

Governance and Accountability

The issue of governance and accountability emerges as a fundamental consideration in the ethical deployment of AI in healthcare. Literature emphasizes that the lack of clear accountability mechanisms creates ambiguity regarding liability when AI systems cause harm or produce errors. Pham (2025) stresses the necessity of integrating robust ethical principles into AI governance frameworks, advocating for regulatory approaches that explicitly delineate responsibility for patient outcomes. Without such clarity, patients may face challenges in seeking redress, and healthcare professionals may be reluctant to fully integrate AI tools into practice.

Evidence also highlights the need for transparency and fairness in governance frameworks. Graaf (2025) argues that multi-stakeholder engagement, involving healthcare providers, ethicists, regulators, and technologists, is critical for developing policies that reflect diverse perspectives and societal values. Collaborative governance structures are essential in ensuring that AI systems align with both clinical needs and ethical standards. This approach recognizes that no single group possesses the expertise to fully address the complex ethical dilemmas posed by AI. By fostering dialogue across stakeholders, governance frameworks can be more responsive and adaptive to the evolving landscape of AI technologies.

Global perspectives further underscore disparities in governance approaches. High-income countries have begun to implement comprehensive strategies to regulate AI in healthcare, incorporating strict privacy standards, accountability mechanisms, and guidelines for equitable deployment. In contrast, resource-limited nations often lack the institutional capacity to enforce such measures, leaving gaps in accountability and oversight (Pham, 2025). This imbalance risks exacerbating global health inequities, as populations in underregulated regions remain more vulnerable to unethical AI practices. Literature thus calls for international cooperation to establish harmonized principles, ensuring that AI governance transcends national boundaries and reflects universal commitments to ethical healthcare delivery.

Synthesis of Findings

Across these themes, the literature converges on the conclusion that ethical considerations are integral to the sustainable and equitable deployment of AI in healthcare. Data privacy concerns reflect the tension between the need for large datasets to enhance AI capabilities and the imperative to protect patient rights. Algorithmic bias reveals how unrepresentative data and flawed design perpetuate inequities in care delivery, underscoring the urgency of inclusive model development. The role of human oversight highlights the irreplaceable value of professional judgment in maintaining autonomy, critical thinking, and patient-centered care. Finally, governance and accountability emphasize the importance of robust, transparent frameworks that can ensure fairness, liability, and ethical alignment across diverse contexts.

Taken together, these findings demonstrate that while AI holds immense potential to transform healthcare delivery, realizing its benefits requires ongoing attention to ethical risks. Addressing privacy, bias, autonomy, and governance issues is not ancillary but central to the responsible integration of AI in healthcare systems. By situating these challenges within both local and global perspectives, the literature underscores the need for collaborative, multidisciplinary solutions that prioritize patient rights and equity alongside technological innovation.

The ethical challenges associated with AI-driven healthcare cannot be disentangled from broader systemic factors that shape how technology is deployed and perceived in medical contexts. The literature reveals that regulatory frameworks, data access, and digital infrastructure are central in influencing both the opportunities and limitations of AI adoption in healthcare. Weak or fragmented regulatory regimes create significant uncertainty that hampers innovation while simultaneously failing to safeguard patient rights. Amann et al. (2020) emphasize that robust legal frameworks are necessary to ensure safe and ethical integration of AI into healthcare systems. Without such clarity, issues such as data privacy violations and algorithmic bias may become exacerbated rather than mitigated. This is particularly evident in countries where regulation has not kept pace with technological developments, leaving patients vulnerable to data misuse and inequitable treatment outcomes.

Data access inequities also play a crucial role in perpetuating ethical challenges. Studies demonstrate that limited access to representative datasets results in the development of AI models that do not adequately capture the diversity of patient populations. As Giansanti and Pirrera note, unequal access to technology, particularly in rural or underserved regions, restricts the ability of AI to provide inclusive healthcare solutions (Perez et al., 2025). These disparities contribute to models that perform well for majority groups while failing to deliver accurate or fair outcomes for minorities and marginalized populations. Furthermore, the lack of high-quality, standardized datasets in many contexts results in models that are difficult to validate across different clinical environments, raising concerns about both efficacy and fairness.

Infrastructural deficiencies further compound these challenges. Many developing countries or remote regions lack the digital capacity to implement and sustain AI systems, thereby widening the gap between resource-rich and resource-poor environments. Outdated digital infrastructures not only limit the technical feasibility of AI deployment but also increase the risks of cybersecurity breaches, creating additional vulnerabilities for patient data protection. The global digital divide thus emerges as both a technical and ethical issue, reinforcing inequities in healthcare delivery and undermining the promise of AI as a tool for universal health improvement.

The literature also points to a variety of potential solutions to mitigate the ethical risks of privacy violations and algorithmic bias. One widely cited approach is the implementation of differential privacy, which allows AI systems to analyze large datasets without compromising individual identities (Amini et al., 2023). This technique balances the dual imperatives of protecting patient confidentiality while enabling meaningful insights from data. Countries with advanced data governance structures have already begun integrating differential privacy into national health strategies, yet implementation remains inconsistent globally. Ensuring wider adoption of such methods requires not only technological investment but also regulatory commitment to enforcing privacy standards.

Beyond technical solutions, strengthening governance frameworks is essential. Mennella et al. argue that stakeholder engagement is critical for building accountable AI systems, with policymakers, clinicians, ethicists, and patients all playing a role in shaping ethical guidelines (Badawy et al., 2024). Multi-stakeholder governance ensures that AI systems are designed and evaluated within a framework that prioritizes transparency, accountability, and inclusivity. Practical approaches to governance include embedding risk-benefit analyses into the AI development pipeline and mandating regular audits of algorithmic outcomes to detect and correct bias. In addition, fostering international cooperation to harmonize ethical standards across borders is particularly important in the context of global data exchange, where inconsistent policies may otherwise undermine ethical safeguards.

Education and continuous training for healthcare professionals also emerge as indispensable solutions. Clinicians must be equipped not only to operate AI tools but also to critically assess their ethical implications. Training programs that emphasize explainable AI, algorithmic bias, and data governance can empower healthcare workers to maintain clinical autonomy while effectively integrating AI into practice. This aligns with the concerns expressed by Rony et al. (2025) regarding the risks of over-reliance on AI leading to deskilling. By promoting critical engagement with AI systems, healthcare professionals can ensure that human judgment remains central to medical decision-making.

Despite these promising avenues, the literature acknowledges several limitations in current research. Much of the scholarship remains focused on the technical efficiency of AI systems, such as accuracy rates and predictive capabilities, with insufficient attention to the broader ethical, legal, and social implications. Badawy et al. (2024) emphasize the need for deeper exploration of how AI impacts clinical contexts beyond performance metrics, particularly in terms of patient experiences and practitioner challenges. This gap suggests a research agenda that extends beyond the laboratory to the lived realities of healthcare delivery. Additionally, McGrath et al. (2025) highlight the importance of longitudinal studies that assess the long-term effects of AI implementation on health outcomes, ethical practices, and system-wide equity. Without such studies, understanding the enduring consequences of AI adoption remains limited.

There is also a scarcity of interdisciplinary research that integrates perspectives from ethics, law, and policy alongside technical analyses. Amini et al. (2023) and Pham (2025) underscore the necessity of developing comprehensive ethical frameworks that not only guide technological design but also address regulatory enforcement and patient rights. Expanding the scope of research to include legal and policy dimensions is crucial for creating governance systems capable of responding to the complex ethical dilemmas posed by AI in healthcare. Similarly, comparative cross-country studies are needed to shed light on how cultural and institutional contexts shape the ethical deployment of AI, as shown by Perez et al. (2025), who found significant variations in outcomes depending on local regulation and infrastructure.

The exchange of health data across borders introduces another underexplored area that requires urgent research attention. International collaboration in AI development and deployment necessitates robust frameworks for cross-border data sharing that can reconcile differences in legal and ethical standards. García-Saisó et al. (2024) argue that resource-limited settings face unique ethical challenges, and without adequate global cooperation, these contexts risk further marginalization. Developing ethical principles for international data governance, therefore, represents both a challenge and an opportunity for expanding the equitable benefits of AI in healthcare.

Overall, the discussion highlights that ethical challenges in AI-driven healthcare are deeply intertwined with systemic structures, ranging from regulation to infrastructure. Addressing these issues requires integrated solutions that combine technological innovation with governance, education, and cross-border collaboration. The limitations of existing research, particularly its narrow technical focus, signal the need for more holistic and interdisciplinary inquiry that can better inform policies and practices aimed at ensuring the ethical use of AI in healthcare.

CONCLUSION

This narrative review has examined the ethical considerations surrounding the integration of artificial intelligence into healthcare, with a particular emphasis on issues of data privacy, algorithmic bias, human oversight, and governance. The findings underscore that while AI holds transformative potential to enhance diagnosis, treatment, and healthcare delivery, these benefits are inseparably linked to complex ethical challenges. Patient data privacy remains vulnerable due to the reliance of AI on large datasets, highlighting the urgent need for robust data governance frameworks and the wider adoption of protective measures such as differential privacy. Algorithmic bias continues to pose a substantial threat, especially for marginalized populations, and requires more inclusive datasets, transparent algorithmic design, and cross-cultural collaboration to ensure fairness. Human oversight and clinical autonomy must remain central, as over-reliance on AI risks diminishing professional judgment and critical thinking in healthcare practice. Governance and accountability frameworks, strengthened by multi-stakeholder engagement, are essential for clarifying liability and ensuring equitable AI deployment.

The urgency of these issues calls for immediate interventions at both policy and practice levels. Stronger regulations, enhanced clinician training, and international cooperation on ethical standards are necessary to mitigate risks while maximizing benefits. Future research should move beyond technical performance to explore interdisciplinary dimensions, including long-term impacts, cross-border data governance, and patient experiences. By adopting integrated strategies that prioritize transparency, accountability, and inclusivity, healthcare systems can navigate the ethical complexities of AI, ensuring that innovation serves all populations equitably.

REFERENCE

Amann, J., Blasimme, A., Vayena, E., Frey, D., & Madai, V. (2020). Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Medical Informatics and Decision Making, 20*(1). https://doi.org/10.1186/s12911-020-01332-6

Amini, M., Jesus, M., Sheikholeslami, D., Alves, P., Benam, A., & Hariri, F. (2023). Artificial intelligence ethics and challenges in healthcare applications: a comprehensive review in the context of the european gdpr mandate. *Machine Learning and Knowledge Extraction, 5*(3), 1023-1035. https://doi.org/10.3390/make5030053

Badawy, W., Zinhom, H., & Shaban, M. (2024). Navigating ethical considerations in the use of artificial intelligence for patient care: a systematic review. *International Nursing Review, 72*(3). https://doi.org/10.1111/inr.13059

Bandekar, N., Chaudhari, R., Yadav, Y., Figueiredo, D., & Chunkhare, M. (2024). The role of ai in emr (electronic medical record) and patient privacy enhancement., 301-320. https://doi.org/10.4018/979-8-3693-1243-8.ch016

Das, B., & Khatua, D. (2024). Ethical considerations for implementing ai-based solutions in digital health and medical analytics., 241-266. https://doi.org/10.4018/979-8-3693-6294-5.ch010

Endo, P. (2025). Artificial intelligence for women and child healthcare: is ai able to change the beginning of a new story? a perspective. *Health Science Reports, 8*(5). https://doi.org/10.1002/hsr2.70779

García-Saisó, S., Martí, M., Pesce, K., Luciani, S., Mújica, Ó., Hennis, A., … & D'Agostino, M. (2024). Artificial intelligence as a potential catalyst to a more equitable cancer care. *JMIR Cancer, 10*, e57276. https://doi.org/10.2196/57276

Giansanti, D., & Pirrera, A. (2025). Integrating ai and assistive technologies in healthcare: insights from a narrative review of reviews. *Healthcare, 13*(5), 556. https://doi.org/10.3390/healthcare13050556

Graaf, Y. (2025). Societal factors influencing the implementation of ai-driven technologies in (smart) hospitals. *PLOS ONE, 20*(6), e0325718. https://doi.org/10.1371/journal.pone.0325718

Graña-Castro, O., Izquierdo, E., Mesa, A., Menasalvas, E., & Chivato-Pérez, T. (2024). Assessing the impact of new technologies on managing chronic respiratory diseases. *Journal of Clinical Medicine, 13*(22), 6913. https://doi.org/10.3390/jcm13226913

Holohan, M., & Fiske, A. (2021). "Like I'm talking to a real person": exploring the meaning of transference for the use and design of ai-based applications in psychotherapy. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.720476

Jaime, F., Muñoz, A., Rodríguez-Gómez, F., & Jeréz-Calero, A. (2023). Strengthening privacy and data security in biomedical microelectromechanical systems by iot communication security and protection in smart healthcare. *Sensors, 23*(21), 8944. https://doi.org/10.3390/s23218944

Jones, R. (2025). Algorithmic bias and fairness in biomedical and health research., 287-324. https://doi.org/10.4018/979-8-3373-4252-8.ch008

Karra, M., & RamaRao, S. (2025). Big data and ai in sexual and reproductive health: a comment. *Studies in Family Planning, 56*(2), 317-331. https://doi.org/10.1111/sifp.70005

Lock, C., Tan, N., Long, I., & Keong, N. (2024). Neuroimaging data repositories and ai-driven healthcare—global aspirations vs. ethical considerations in machine learning models of neurological disease. *Frontiers in Artificial Intelligence, 6*. https://doi.org/10.3389/frai.2023.1286266

McGrath, C., Chau, C., & Molina, G. (2025). Monitoring oral health remotely: ethical considerations when using ai among vulnerable populations. *Frontiers in Oral Health, 6*. https://doi.org/10.3389/froh.2025.1587630

Mennella, C., Maniscalco, U., Pietro, G., & Esposito, M. (2024). Ethical and regulatory challenges of ai technologies in healthcare: a narrative review. *Heliyon, 10*(4), e26297. https://doi.org/10.1016/j.heliyon.2024.e26297

Mooghal, M., Nasir, S., Arif, A., Khan, W., Rashid, Y., & Vohra, L. (2024). Innovations in artificial intelligence-driven breast cancer survival prediction: a narrative review. *Cancer Informatics, 23*. https://doi.org/10.1177/11769351241272389

Mushtaq, S., & Hameeda, Q. (2025). Empowering public health: ai-powered security solutions for ai-driven challenges. *Applied AI Letters, 6*(2). https://doi.org/10.1002/ail2.119

Onetiu, F., Bratu, M., Folescu, R., Bratosin, F., & Bratu, T. (2025). Assessing medical students' perceptions of ai-integrated telemedicine: a cross-sectional study in Romania. *Healthcare, 13*(9), 990. https://doi.org/10.3390/healthcare13090990

Papageorgiou, P. (2025). Artificial intelligence in primary malignant bone tumor imaging: a narrative review. *Diagnostics, 15*(13), 1714. https://doi.org/10.3390/diagnostics15131714

Pavunraj, D., Kumar, A., & Anbumaheshwari, K. (2025). AI in personalized treatment planning., 117-142. https://doi.org/10.4018/979-8-3373-1275-0.ch006

Pawelczyk, J., Kraus, M., Voigtlaender, S., Siebenlist, S., & Rupp, M. (2025). Advancing musculoskeletal care using ai and digital health applications: a review of commercial solutions. *HSS Journal®, 21*(3), 331-341. https://doi.org/10.1177/15563316251341321

Perez, K., Wisniewski, D., Ari, A., Lee, K., Lieneck, C., & Ramamonjiarivelo, Z. (2025). Investigation into application of ai and telemedicine in rural communities: a systematic literature review. *Healthcare, 13*(3), 324. https://doi.org/10.3390/healthcare13030324

Pham, T. (2025). Ethical and legal considerations in healthcare ai: innovation and policy for safe and fair use. *Royal Society Open Science, 12*(5). https://doi.org/10.1098/rsos.241873

Rad, A., Vardanyan, R., Athanasiou, T., Maessen, J., & Nia, P. (2025). The ethical considerations of integrating artificial intelligence into surgery: a review. *Interdisciplinary Cardiovascular and Thoracic Surgery, 40*(3). https://doi.org/10.1093/icvts/ivae192

Radanliev, P. (2025). Privacy, ethics, transparency, and accountability in ai systems for wearable devices. *Frontiers in Digital Health, 7*. https://doi.org/10.3389/fdgth.2025.1431246

Rony, M., Ahmad, S., Das, D., Tanha, S., Deb, T., Akter, M., … & Akter, F. (2025). Nursing students' perspectives on integrating artificial intelligence into clinical practice and training: a qualitative descriptive study. *Health Science Reports, 8*(4). https://doi.org/10.1002/hsr2.70728

Shah, S. (2024). The impact of artificial intelligence on health outcomes: a review. *PJS, 1*. https://doi.org/10.70389/pjs.100006

Sargsyan, A., Hovsepyan, S., & Muradyan, A. (2024). Ubiquitous and powerful artificial intelligence (AI)., 255-271. https://doi.org/10.1007/978-3-031-62332-5_26

Varnosfaderani, S., & Forouzanfar, M. (2024). The role of ai in hospitals and clinics: transforming healthcare in the 21st century. *Bioengineering, 11*(4), 337. https://doi.org/10.3390/bioengineering11040337

Vozna, A., & Costantini, S. (2025). Ethical, legal, and societal dimensions of ai-driven social robots in elderly healthcare. *Intelligenza Artificiale*. https://doi.org/10.1177/17248035241310192

Weerasinghe, K. (2025). The impact of ai-driven chatbot assistance on protocol development and clinical research engagement: an implementation report. *Journal of Personalized Medicine, 15*(7), 269. https://doi.org/10.3390/jpm15070269