Ethical framework for artificial intelligence in healthcare research: A path to integrity

Ahmad A Abujaber, Abdulqadir J Nashwan

Specialty type: Medical laboratory technology

Provenance and peer review: Invited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report’s classification
Scientific Quality: Grade B
Novelty: Grade B
Creativity or Innovation: Grade B
Scientific Significance: Grade B

P-Reviewer: Shao J, China

Received: March 10, 2024
Revised: April 18, 2024
Accepted: May 6, 2024
Published online: September 20, 2024
Processing time: 106 Days and 2 Hours

Ahmad A Abujaber, Department of Nursing, Hazm Mebaireek General Hospital (HMGH), Doha 3050, Qatar
Abdulqadir J Nashwan, Department of Nursing, Hamad Medical Corporation, Doha 3050, Qatar

Corresponding author: Abdulqadir J Nashwan, MSc, Research Scientist, Department of Nursing, Hamad Medical Corporation, Rayyan Road, Doha 3050, Qatar. anashwan@hamad.qa

Abstract

The integration of Artificial Intelligence (AI) into healthcare research promises unprecedented advancements in medical diagnostics, treatment personalization, and patient care management. However, these innovations also bring forth significant ethical challenges that must be addressed to maintain public trust, ensure patient safety, and uphold data integrity. This article sets out to introduce a detailed framework designed to steer governance and offer a systematic method for assuring that AI applications in healthcare research are developed and executed with integrity and adherence to medical research ethics.

Key Words: Artificial intelligence; Ethical framework; Healthcare research; Ethical Principles; Integrity; Patient safety

©The Author(s) 2024. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: This editorial sets out to introduce a detailed framework designed to steer governance and offer a systematic method for assuring that artificial intelligence applications in healthcare research are developed and executed with integrity and adherence to medical research ethics.

Citation: Abujaber AA, Nashwan AJ. Ethical framework for artificial intelligence in healthcare research: A path to integrity. World J Methodol 2024; 14(3): 94071
URL: https://www.wjgnet.com/2222-0682/full/v14/i3/94071.htm
DOI: https://dx.doi.org/10.5662/wjm.v14.i3.94071
INTRODUCTION

The integration of artificial intelligence (AI) into healthcare research marks a pivotal shift towards groundbreaking advancements in diagnostics, treatment, and patient care management. This evolution, however, introduces a spectrum of ethical challenges that necessitate meticulous scrutiny and governance. At the heart of these challenges are concerns over privacy and confidentiality, as AI solutions require access to extensive patient data, raising significant risks to individual privacy[1]. The issue of informed consent also becomes more complex, as the applications of AI in healthcare research may extend beyond the scope of traditional consent frameworks, necessitating updated procedures that transparently communicate the potential uses of patient data[2,3].

Moreover, the inherent risk of bias in AI algorithms presents a critical ethical dilemma, with the potential to perpetuate existing disparities in healthcare outcomes[4]. Ensuring fairness and addressing biases is paramount to uphold ethical standards in AI-driven healthcare solutions. Transparency and explainability of AI decision-making processes are essential to maintain trust and accountability, particularly in a field that is sensitive such as healthcare[5]. The question of accountability for AI-driven decisions further complicates the ethical landscape[6], alongside concerns about equitable access to AI benefits, which could inadvertently widen health disparities[7].

Addressing these ethical challenges is crucial to leveraging AI in healthcare research responsibly. It requires a collaborative effort from researchers, ethicists, policymakers, and the broader healthcare community to develop a comprehensive ethical framework. Such a framework aims not only to mitigate risks but also to ensure that AI advancements contribute positively to patient care, uphold patients’ rights, and promote equity. Although there is significant discussion in academic circles about the essential need for a robust ethical framework to oversee the integration of AI in healthcare research, current literature lacks a detailed framework that articulates foundational ethics and sets the operationalization guidelines and the implementation principles. Our research represents an innovative step toward creating a thorough and solid ethical structure. This proposed framework is designed to protect ethical integrity and ensure the highest ethical practices in AI healthcare research.

THE IMPERATIVE OF AN ETHICAL FRAMEWORK

The rapid advancement and integration of AI in healthcare research emphasizes the pressing need for an ethical framework to guide its application[8]. The absence of such a framework risks ethical lapses that could undermine public trust, compromise patient privacy, and exacerbate healthcare disparities. An ethical framework serves as a compass, guiding researchers and practitioners in navigating the complex moral terrain of AI in healthcare[9].

Adopting a robust ethical framework offers several remedies to these challenges. First, it ensures that privacy and confidentiality are paramount, safeguarding patient data against misuse. By establishing clear guidelines for data handling, the framework can mitigate risks associated with data breaches and unauthorized access. Second, it enhances informed consent processes, ensuring that participants are fully aware of how their data will be used, including potential AI applications, thus respecting their autonomy[10].

Moreover, an ethical framework addresses biases in AI algorithms, promoting fairness in healthcare outcomes[11]. It mandates regular audits of AI systems for bias and requires the implementation of corrective measures when disparities are identified. Transparency and explainability become foundational, with AI systems designed to provide understandable outputs, thereby fostering trust among patients and healthcare providers[10].

Finally, an ethical framework ensures accountability and equitable access to AI-driven healthcare innovations. It delineates responsibilities among AI developers, healthcare providers, and policymakers, ensuring that those impacted by AI decisions have recourse[12]. By prioritizing equitable access, the framework also works to prevent the widening of health disparities, making the benefits of AI in healthcare research accessible to all[10].

PROPOSED ETHICAL FRAMEWORK FOR AI IN HEALTHCARE RESEARCH

The proposed framework, as illustrated in Figure 1, is centered around the four key ethical principles in medicine, further supported by actionable guidelines to operationalize these principles in practical research settings. The framework is visualized in a diagram that illustrates its multifaceted approach, focusing on the core ethical principles of respect for autonomy, beneficence, non-maleficence, and justice as its core pillars.

Core ethical principles

The core ethical principles guiding AI in healthcare research are deeply rooted in the foundational principles of medical research ethics. These principles are not only relevant but essential in ensuring that AI technologies are developed and used in ways that align with the ethical conduct of medical research. Here’s how these core principles relate to medical research ethics:

Respect for autonomy: In medical research, respecting autonomy involves acknowledging and upholding the rights of participants to make informed decisions about their involvement. This principle is crucial in AI healthcare research, especially in the context of informed consent. It emphasizes the importance of ensuring that individuals are fully informed about how their data will be used in AI applications, reflecting their autonomy in the decision-making process.
Beneficence: Beneficence in medical research ethics refers to the obligation to maximize benefits and minimize harms to research participants. In the context of AI, this principle mandates that the development and application of AI technologies should aim to improve healthcare outcomes and patient care, ensuring that the benefits of AI advancements are realized and maximized in clinical settings.

Non-maleficence: The principle of non-maleficence, or "do no harm," is vital in medical research, emphasizing the importance of avoiding harm to participants. In AI healthcare research, this translates to ensuring that AI systems do not inadvertently cause harm, such as through biases in algorithms that could lead to incorrect diagnoses or treatment recommendations. It highlights the need for rigorous testing and validation of AI models to prevent potential adverse outcomes.

Justice: Justice in medical research ethics focuses on ensuring equitable access to the benefits of research and fair distribution of risks and burdens. When applied to AI in healthcare, this principle highlights the need to address and mitigate healthcare disparities that AI solutions might exacerbate. It calls for the equitable development and deployment of AI solutions, ensuring that all patient populations can benefit from AI advancements without widening existing healthcare gaps.

Integrating these core ethical principles into the development and application of AI in healthcare research ensures that AI solutions are not only innovative but also ethically responsible. It aligns AI advancements with the longstanding commitments of medical research to respect human dignity, promote well-being, avoid harm, and distribute healthcare benefits justly among all segments of the population. This integration is crucial for maintaining trust in AI applications in healthcare and ensuring that these powerful tools serve the collective good of society.

Operational guidelines

Building upon the core ethical principles, operationalization defines the precise measures and indicators for ethical principles, translating abstract concepts into quantifiable criteria. The following operational guidelines are designed to practically guide the seamless integration of the core ethical principles into every phase of AI development and application in healthcare. However, the effective institutionalization of these guidelines hinges on the creation of precise, quantifiable, and observable metrics that capture the practical implementation of these ethical principles. This crucial task is entrusted to research institutions and Institutional Review Boards (IRBs), underlining their role in ensuring that ethical considerations are integrated consistently and effectively across the spectrum of AI healthcare initiatives.
Transparency and explainability: AI systems should be transparent in their operations, with mechanisms in place to explain decisions to both practitioners and patients.

Privacy and data protection: Strict protocols that adhere to the relevant laws and regulations such as the Health Insurance Portability and Accountability Act (HIPAA) must be established to protect patient data, respecting privacy, and confidentiality throughout the research process.

Inclusive design and bias mitigation: AI technologies should be designed with diverse populations in mind, actively working to mitigate biases in datasets and algorithms.

Stakeholder engagement: Patients, healthcare providers, ethicists, and policymakers should be involved in the development and implementation of AI applications to ensure a wide range of perspectives are considered.

Implementation guidelines
Implementing the ethical framework for AI in healthcare research is a comprehensive approach that embeds fundamental ethical principles throughout every phase of AI development and deployment. This phase transitions from theoretical planning to tangible action, executing the strategies, plans, or policies established during the operationalization phase. Implementation entails the application of operationalized components to realize specific goals, encompassing all practical aspects of enacting a plan. This includes allocating resources, organizing schedules, and conducting the activities necessary to bring theoretical concepts and strategies to fruition.

This phase requires active collaboration among various stakeholders, including researchers, clinicians, patients, ethicists, and policymakers[8]. Here are key steps to effectively implement the ethical framework:

Multi-disciplinary collaboration: The multidisciplinary collaboration entails: (1) Establishing interdisciplinary teams that include ethicists, data scientists, healthcare professionals, and preferably a patient representative to guide the ethical development and deployment of AI solutions[13]; and (2) Facilitating regular discussions and workshops to address ethical concerns and integrate diverse perspectives into AI research and development processes.

Education and training: Education and training include: (1) Developing educational programs and resources for AI researchers and healthcare professionals that focus on the ethical implications of AI in healthcare. Abujaber et al[1] proposed that educational institutions, particularly those specializing in health-related fields, should begin integrating AI into their curricula during the collegiate years. Such early exposure to AI is recommended to ease future acceptance and adoption among students entering healthcare professions[1]; and (2) Including modules on ethical decision-making, bias recognition, and mitigation strategies in AI development curricula.

Policy development and regulatory compliance: This can be achieved by two steps: (1) Working with regulatory bodies to ensure that policies and guidelines for AI in healthcare research reflect the core ethical principles[13]; and (2) Encouraging the adoption of standards and best practices that promote transparency, accountability, and equity in AI applications.

Ethical review and oversight: This requires: (1) Implementing ethical review processes specifically tailored to AI projects in healthcare research, akin to traditional human subjects’ research oversight; and (2) Establishing ethics committees or boards with expertise to evaluate AI projects, focusing on the potential risks, benefits, and ethical considerations.

Public and stakeholder engagement: The advocacy for integrating AI into medical research is driven by the objective to fully harness its potential in improving patient care, providing support to patients’ families, and benefiting the wider community. Consequently, the successful implementation of AI depends on the active involvement of patients, their families, and other key stakeholders. Engagement should encompass: (1) Collaborating with patients, the public, and stakeholders via consultative and participatory design methods to solicit feedback on the development of AI and its potential impacts[14]; and (2) Promoting transparency by publicly sharing information about AI projects, including objectives, methodologies, and ethical considerations[15].

Continuous monitoring and evaluation: This involves: (1) Monitoring the outcomes of AI applications in healthcare research continuously to identify unforeseen ethical issues or adverse effects; and (2) Implementing mechanisms for ongoing evaluation and adaptation of AI technologies, ensuring they remain aligned with ethical principles and societal values[16].

Feedback loop: This requires: (1) Creation of a feedback loop that allows for the continuous integration of lessons learned from the implementation of AI technologies back into the ethical framework; and (2) Refinement and updating the ethical framework regularly based on new insights, technological advancements, and evolving societal norms.

Implementing this ethical framework is an ongoing process that requires commitment, transparency, and adaptability. By taking these steps, stakeholders can ensure that AI in healthcare research is conducted with the highest ethical standards, ultimately benefiting society while safeguarding individual rights and promoting equity.
ETHICAL ISSUES AND PROPOSED SOLUTIONS

Addressing how AI can potentially compromise patient confidentiality and the measures that can be implemented to safeguard sensitive health information.

Discussion on the risks of bias in AI algorithms, its impact on treatment and diagnosis outcomes across different demographics, and strategies for mitigation.

Exploring the complexities of informed consent when using AI, including the use of patients' data for training AI systems.

Clarifying who is held accountable when AI systems make errors or cause harm, and how liability is managed in the deployment of AI technologies in healthcare.

Proposing specific guidelines for data management that comply with existing regulations and ethical standards, like HIPAA, to ensure data privacy and security.

Providing clear guidelines for maintaining transparency in the operations of AI systems and the logic behind AI decision-making processes.

Recommending strategies for designing AI systems that are inclusive and equitable, ensuring fair representation and treatment of all patient groups.

Outlining the role of IRBs and other ethics committees in ongoing oversight, including regular ethical reviews and the monitoring of AI systems post-deployment to quickly identify and address new ethical issues.

CONCLUSION

As AI technologies become increasingly integrated into healthcare research, establishing ethical foundations is imperative to ensure these innovations serve the public good. While the proposed ethical framework serves as a foundational step towards guiding AI applications in healthcare research, it is evident that further refinement is essential to align it seamlessly with the existing medical research governance systems. Achieving higher ethical standards in AI healthcare research relies not only on the institutional adoption of this framework but also on its continuous enhancement to improve practical implementation. Success in this domain is contingent upon both the robust integration of the framework within institutional practices and the ongoing efforts to increase its operational effectiveness.

FOOTNOTES

Author contributions: Abujaber AA contributed to the manuscript's writing, editing, and literature review; Nashwan AJ contributed to the manuscript's writing, editing, and literature review.

Conflict-of-interest statement: All the authors report no relevant conflicts of interest for this article.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: https://creativecommons.org/Licenses/by-nc/4.0/

Country of origin: Qatar

ORCID number: Ahmad A Abujaber 0000-0002-8704-4991; Abdulqadir J Nashwan 0000-0003-4845-4119.

S-Editor: Li L.
L-Editor: A.
P-Editor: Zhang L.

REFERENCES

Abujaber AA et al. Ethical framework for AI in healthcare research


