Does AI Respect Patient Autonomy? Navigating the Ethical Tensions in Digital Health

Rasit Dinc

Rasit Dinc Digital Health & AI Research

Published: January 31, 2022 | Medical Imaging AI

DOI: 10.5281/zenodo.17998040

Abstract

The rapid integration of Artificial Intelligence AI into healthcare promises revolutionary improvements in diagnostics, treatment planning, and administrativ...

The rapid integration of Artificial Intelligence (AI) into healthcare promises revolutionary improvements in diagnostics, treatment planning, and administrative efficiency. However, this technological advancement introduces profound ethical questions, none more central than the impact of AI on **patient autonomy**. Autonomy, the foundational principle of medical ethics, asserts the patient's right and capacity to make informed decisions regarding their own health. As AI systems become increasingly involved in clinical decision-making, the core challenge is to ensure that this technology augments, rather than diminishes, the patient's ability to exercise self-determination.

The Black Box and the Crisis of Informed Consent

The primary threat AI poses to patient autonomy stems from the **"black box" problem** and the subsequent challenge to **informed consent**. For a patient's consent to be truly informed, they must have a sufficient understanding of the proposed intervention, including its risks, benefits, and alternatives. When an AI algorithm, whose decision-making process is opaque even to its developers, generates a clinical recommendation, the traditional process of informed consent breaks down [1].

This lack of transparency is compounded by the need to communicate complex AI-generated information to patients in an understandable manner. While simplifying medical reports is an ethical imperative to promote patient comprehension, excessive simplification can inadvertently compromise clinical accuracy. Academic research has highlighted this tension, noting that while a 7th-grade reading level is an ethical ideal for patient materials, AI tools that simplify reports below an 11th-grade level risk introducing inaccuracies that could alter patient management [2]. This finding reveals a critical ethical trade-off: the pursuit of readability, which supports comprehension, must be balanced against the preservation of clinical detail, which is essential for

accurate decision-making and, ultimately, autonomy.

> "The ethical imperative to balance readability and accuracy in patient-centered reporting necessitates rigorous examination and clear ethical standards for deploying LLM-based simplifications." [2]

Beyond Comprehension: Epistemic Injustice and Relational Autonomy

The ethical implications extend beyond mere comprehension. The deployment of AI in healthcare also raises concerns about **epistemic injustice**, where simplified or filtered AI-generated reports may marginalize a patient's access to nuanced clinical information, thereby limiting their understanding and decision-making capabilities [2]. If AI acts as a gatekeeper to knowledge, it risks creating a power imbalance that undermines the patient-physician relationship.

Furthermore, patient autonomy is not exercised in a vacuum; it is a **relational autonomy**, influenced by social contexts, family, and cultural values. For AI to respect this relational aspect, the information it provides must be contextually sensitive and sufficient to support decision-making within the patient's unique circumstances. The ethical deployment of AI, therefore, requires a shift from a purely technical focus to one that addresses foundational bioethical concerns: the patient-clinician relationship, information asymmetry, and the integrity of consent in the age of artificial intelligence.

Safeguarding Autonomy in the AI Era

To ensure that AI serves as a tool for empowerment rather than a barrier to self-determination, several safeguards must be implemented.

1. **Layered Communication Strategies:** Healthcare providers must adopt communication models that offer information at varying levels of complexity, allowing patients to delve deeper into the AI's rationale as needed. 2. **Model Explainability (XAI):** Continued research and regulatory pressure are necessary to develop AI systems that can provide clear, human-interpretable explanations for their outputs. 3. **Patient-Centered Design:** AI tools must be designed with patient empowerment at the forefront, promoting active participation in their care rather than positioning them as passive recipients of AI-driven recommendations.

The journey toward ethically integrating AI into digital health requires continuous vigilance and a commitment to prioritizing the patient's voice. The question is not whether AI *can* respect patient autonomy, but whether we, as developers, clinicians, and policymakers, *will* design and regulate it to do so. For more in-depth analysis on this topic, the resources at www.rasitdinc.com provide expert commentary on the intersection of digital health, ethics, and patient-centered care.

**

References

[1] Lamanna, C. (2018). Should Artificial Intelligence Augment Medical Decision Making? The Case for an Autonomy Algorithm. AMA Journal of

Ethics*, 20(9), E840-847. https://journalofethics.ama-assn.org/article/should-artificial-intelligence-augment-medical-decision-making-case-autonomy-algorithm/2018-09 [2] Lee, H. S., et al. (2025). The ethics of simplification: balancing patient autonomy, comprehension, and accuracy in AI-generated radiology reports.* BMC Medical Ethics*, 26(136). https://bmcmedethics.biomedcentral.com/articles/10.1186/s12910-025-01285-3
