

## Autonomy and AI-Based Healthcare

### Ádám Kovács\*

Semmelweis University, Üllői út 26, 1085 Budapest, Hungary

\*: All correspondence should be sent to: Mr. Ádám Kovács.

Author's Contact: Ádám Kovács, M.P.H., M.A., E-mail: [adam.kovacs@hotmail.com](mailto:adam.kovacs@hotmail.com)

DOI: <https://doi.org/10.15354/si.26.op155>

Funding: No funding source declared.

COI: The author declares no competing interest.

AI Declaration: The author affirms that artificial intelligence did not contribute to the process of preparing the work.

The integration of artificial intelligence (AI) into healthcare has the potential to revolutionize diagnostics, treatment planning, and patient care, but it also raises critical questions about autonomy. AI-based systems can analyze vast datasets, identify patterns beyond human perception, and provide predictive insights, offering unprecedented support to clinicians and patients. However, as AI assumes greater decision-making roles, concerns emerge regarding patient autonomy, clinician authority, and the ethical delegation of responsibility. This article argues that while AI can enhance healthcare efficiency and precision, it should complement rather than replace human judgment. Maintaining autonomy requires transparent algorithms, explainable AI, informed patient consent, and clear boundaries for AI intervention. Balancing the benefits of AI with ethical and regulatory oversight is essential to preserve trust, accountability, and human-centered care. By critically examining autonomy in AI-driven healthcare, stakeholders can harness innovation responsibly while safeguarding patient rights and clinical integrity.

**Keywords:** AI in Healthcare; Autonomy; Ethics; Patient-Centered Care; Clinical Decision-Making

Science Insights, May 31, 2026; Vol. 48, No. 5, pp.2207-2210.

© 2026 Insights Publisher. All rights reserved.



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the [Creative Commons Attribution-NonCommercial 4.0 License](https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed by the Insights Publisher.

**A**RTIFICIAL intelligence is transforming healthcare at a pace that challenges conventional frameworks of decision-making, ethics, and clinical responsibility. Machine learning algorithms can now process vast amounts of medical data, detect subtle patterns in imaging, predict patient outcomes, and even propose treatment strategies (Topol, 2019; Rajkomar et al., 2019). AI applications span diagnostic radiology,

pathology, predictive analytics, telemedicine, and personalized medicine. While these technologies promise efficiency, accuracy, and scalability, they also raise profound questions about autonomy—both for patients and clinicians (Jiang et al., 2017). Autonomy, understood as the capacity to make informed and voluntary decisions, is a cornerstone of ethical medical practice (Beauchamp & Childress, 2019). The increasing reli-

ance on AI to guide or make clinical decisions threatens to redefine traditional roles and responsibilities in healthcare. Addressing these tensions is critical to ensure that AI serves as an enhancement rather than a replacement of human judgment.

Patient autonomy is central to contemporary healthcare, emphasizing informed consent, shared decision-making, and the right to choose or refuse treatment. AI-based recommendations, however, can complicate these principles. When algorithms provide treatment suggestions based on complex, opaque analyses, patients may be inclined to defer to “machine intelligence,” assuming it is inherently more accurate or objective than human judgment (Longoni et al., 2019). This deference can unintentionally undermine patient autonomy, transforming consent into passive compliance. To counteract this, AI systems must be designed to provide explainable outputs, translating complex predictive models into comprehensible recommendations (Doshi-Velez & Kim, 2017). Transparency is key: patients need accessible explanations of how AI arrives at conclusions, the certainty of predictions, and the risks and benefits of proposed interventions. Without such transparency, autonomy becomes symbolic rather than substantive.

Clinician autonomy is also affected by the rise of AI. Medical professionals have traditionally exercised expertise, judgment, and ethical deliberation in patient care. AI has the potential to augment this expertise by synthesizing data, highlighting overlooked patterns, or predicting complications before they occur (Topol, 2019). Yet as reliance on AI grows, clinicians may feel pressured to follow algorithmic recommendations, even when their judgment suggests alternatives. This phenomenon, sometimes referred to as “automation bias,” can subtly erode professional autonomy and clinical confidence (Goddard et al., 2012). Maintaining clinician autonomy requires integrating AI as a supportive tool, not a prescriptive authority. Healthcare systems must establish clear guidelines delineating when and how AI recommendations should influence decision-making, ensuring that clinicians retain ultimate responsibility and discretion.

The ethical dimension of autonomy intersects with accountability. If AI suggests a treatment that results in harm, who bears responsibility—the clinician, the AI developer, or the healthcare institution? Traditional medico-legal frameworks assume human actors as decision-makers. As AI becomes more sophisticated, attributing responsibility becomes complex (Mittelstadt et al., 2016). Preserving autonomy therefore demands not only transparency and interpretability but also regulatory structures that clarify liability, enforce ethical standards, and promote accountability. Only by defining the limits and responsibilities of AI can healthcare systems prevent the erosion of ethical autonomy in clinical practice.

AI-based healthcare offers particular promise in enhancing patient autonomy when implemented thoughtfully. Personalized medicine, driven by AI, can provide patients with highly tailored treatment options based on genetic, environmental, and lifestyle factors (Obermeyer & Emanuel, 2016). Predictive analytics can empower individuals to take preventive measures or make informed choices about disease management. Telemedicine platforms powered by AI chatbots can expand access to information and healthcare guidance, particularly in underserved regions

(Jiang et al., 2017). These applications demonstrate that AI can complement autonomy, enabling patients to make better-informed choices, provided the technology is designed to be transparent, interpretable, and supportive of human judgment rather than coercive or opaque.

Despite its potential, AI in healthcare is not inherently neutral. Biases in training data, algorithmic opacity, and unequal access to technology can compromise both patient and clinician autonomy. For example, AI models trained on datasets lacking diversity may misdiagnose or inadequately treat underrepresented populations, undermining equitable decision-making (Obermeyer et al., 2019). Similarly, socioeconomic disparities in access to AI-driven tools can exacerbate health inequities, limiting the autonomy of patients in resource-limited settings. Addressing these challenges requires proactive design, rigorous validation, continuous monitoring, and inclusion of diverse populations in training data. Ensuring fairness and transparency is essential to prevent AI from inadvertently constraining autonomy.

Autonomy in AI-based healthcare also intersects with cultural and social contexts. Different societies have varying expectations regarding medical authority, consent, and the role of technology in healthcare (WHO, 2021). Implementing AI globally necessitates sensitivity to these differences, ensuring that tools respect cultural norms and patient expectations. In some contexts, AI may enhance autonomy by providing previously inaccessible information; in others, it may conflict with established patient-clinician relationships. Policymakers and developers must consider these nuances to preserve both ethical and practical autonomy across diverse healthcare settings.

Training and education are critical to sustaining autonomy in AI-driven healthcare. Clinicians must understand AI’s capabilities, limitations, and interpretive nuances to make informed decisions and effectively communicate recommendations to patients (Topol, 2019). Similarly, patients require accessible explanations of AI tools and their implications for care. Education fosters informed consent, reduces over-reliance on algorithms, and enhances trust. By cultivating digital literacy among both clinicians and patients, healthcare systems can integrate AI without diminishing autonomy, ensuring that human judgment remains central to care.

The concept of shared decision-making provides a framework for reconciling AI integration with autonomy. In this model, AI serves as an advisor, providing evidence-based insights that inform discussions between patients and clinicians (Elwyn et al., 2012). The final decisions, however, remain grounded in human judgment, values, and preferences. By positioning AI as a facilitator rather than a replacement, healthcare systems can leverage technological advances while preserving ethical principles. Shared decision-making underscores the complementary potential of AI: it can enhance knowledge, reduce uncertainty, and optimize treatment without overriding the autonomy of stakeholders.

AI also has the potential to redefine autonomy in ways that extend beyond traditional clinical interactions. Predictive models can alert patients to risk factors before disease onset, enabling proactive lifestyle modifications. Wearable sensors and real-time monitoring systems, powered by AI, can empower

patients to manage chronic conditions independently while remaining connected to clinical oversight (Rajkomar et al., 2019). These innovations can shift autonomy from reactive decision-making within healthcare facilities to continuous, informed self-management. In this sense, AI has the potential to expand rather than diminish autonomy, provided it is implemented ethically and inclusively.

Ethical frameworks for AI-based healthcare must emphasize principles that protect autonomy. Transparency, explainability, accountability, and patient-centered design are essential (Floridi et al., 2018). AI systems should provide interpretable outputs, allow clinician override, and ensure equitable access. Regulatory bodies must define standards for validation, reporting, and liability. By embedding autonomy as a core ethical principle, AI can become a partner in care rather than an external authority, ensuring that technological advancement does not come at the cost of human rights or clinical judgment.

Looking ahead, the integration of AI into healthcare represents both an opportunity and a responsibility. AI can improve diagnostic accuracy, optimize treatment, and empower patients, but these benefits must be balanced against risks to autonomy,

ethics, and trust (Topol, 2019). Successful implementation requires thoughtful design, robust regulatory frameworks, continuous monitoring, and active involvement of patients and clinicians in decision-making. By foregrounding autonomy, stakeholders can ensure that AI enhances healthcare rather than undermining the foundational principles of medical ethics.

In conclusion, AI-based healthcare has transformative potential, but its integration raises complex questions about autonomy. Light-speed computation, pattern recognition, and predictive modeling can augment decision-making, improve outcomes, and empower patients. Yet the opacity of algorithms, potential biases, and pressure to defer to machine recommendations pose ethical and practical challenges. Maintaining autonomy requires transparency, explainability, clinician oversight, and informed patient engagement. AI should complement human judgment, not replace it, preserving the ethical and professional principles that underlie trust in medicine. By balancing innovation with accountability, healthcare systems can harness AI to enhance care, empower patients, and sustain clinician autonomy, ensuring that technological advancement aligns with human-centered values. ■

---

Received: November 28, 2025 | Revised: February 20, 2026 | Accepted: March 02, 2026

---

## References

- Beauchamp, T. L., & Childress, J. F. (2019). *Principles of biomedical ethics* (8th ed.). Oxford University Press.
- Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.
- Elwyn, G., Frosch, D., Thomson, R., Joseph-Williams, N., Lloyd, A., Kinnersley, P., ... Barry, M. (2012). Shared decision making: A model for clinical practice. *Journal of General Internal Medicine*, 27(10), 1361–1367. DOI: <https://doi.org/10.1007/s11606-012-2077-6>
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... Vayena, E. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines*, 28(4), 689–707. DOI: <https://doi.org/10.1007/s11023-018-9482-5>
- Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators, and mitigators. *Journal of the American Medical Informatics Association*, 19(1), 121–127. DOI: <https://doi.org/10.1136/amiainl-2011-000089>
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243. DOI: <https://doi.org/10.1136/svn-2017-000101>
- Longoni, C., Bonezzi, A., & Morewedge, C. K. (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research*, 46(4), 629–650. DOI: <https://doi.org/10.1093/jcr/ucz013>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1–21. DOI: <https://doi.org/10.1177/2053951716679679>
- Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future—Big data, machine learning, and clinical medicine. *The New England Journal of Medicine*, 375(13), 1216–1219. DOI: <https://doi.org/10.1056/NEJMp1606181>
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. DOI: <https://doi.org/10.1126/science.aax2342>
- Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *The New England Journal of Medicine*, 380(14), 1347–1358. DOI: <https://doi.org/10.1056/NEJMra1814259>
- Topol, E. (2019). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books.
- World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*. WHO Press.