The Human Hand Behind the Algorithm: Ethical Perspectives on AI in Clinical Practice

Tanner Buckway[1]

1. Rocky Vista University, College of Osteopathic Medicine, Ivins, UT

Keywords: Artificial Intelligence, Healthcare, Ethics

Abstract

The incorporation of artificial intelligence (AI) into clinical practice represents a pivotal shift in healthcare delivery. AI promises greater efficiency, diagnostic support, and precision, yet its adoption raises ethical concerns related to autonomy, equity, transparency, and the human dimensions of care. This paper explores these tensions by analyzing key ethical themes through both philosophical perspectives and a firsthand clinical encounter with AI-assisted documentation. The case underscores the need for critical human oversight while highlighting the potential utility of AI when effectively integrated. The goal is to inform decision-making by emphasizing foundational values in ethical clinical practice.

Introduction

Artificial intelligence is transforming healthcare from diagnostics and triage to documentation and administrative efficiency. As this transformation accelerates, the medical community must grapple with fundamental ethical questions: what role should AI play in patient care? How do we ensure its fair and responsible use? What safeguards must exist to prevent harm? These questions require a nuanced, balanced approach informed by lived experience, research, and ethical analysis. This paper examines four core areas of concern, clinical autonomy, health equity, data ethics, and the human element in care, while incorporating a personal story from medical training that illustrates the promise and limits of AI in clinical settings.

I. Clinical Autonomy and the Role of the Physician

One of the clearest ethical tensions arises in the interaction between physician judgment and AI-generated recommendations. While some worry that AI might replace human reasoning, others argue it can serve as a tool to augment clinical decision-making, particularly in high-volume, high-complexity environments.

During my third year of medical school, I worked with a preceptor who used an AI-powered recording system to document patient encounters. The software transcribed conversations and drafted clinical notes in real time. On one occasion, the system inaccurately documented the patient's gender, misstated physical exam findings, and included an incorrect differential diagnosis. However, my preceptor immediately identified these discrepancies and made swift corrections, ensuring the chart was accurate

and clinically sound. His efficiency and fluency in working with the AI highlighted how, with proper oversight, such systems could expedite documentation while preserving accuracy and safety.

This experience supports a consequentialist interpretation of AI in medicine, where ethical action is judged by its outcomes. The AI tool introduced correctable errors, but it also reduced clerical burden, increased the physician's time with the patient, and supported a more streamlined workflow. This aligns with real-world findings. A 2024 pilot study presented at the American Society of Clinical Oncology (ASCO) evaluated an ambient AI scribe (DeepScribe) across nearly 5,500 patient encounters at Texas Oncology. Providers reported significant reductions in administrative burden and documentation time, while the accuracy of billing code capture was maintained or improved. Crucially, physician satisfaction improved across multiple dimensions including emotional exhaustion, professional fulfillment, and perceptions of workplace efficiency (Doshi et al., 2024 & Shah et al., 2025). Although AI may introduce correctable errors, the net benefits support its ethical use, particularly when human oversight is maintained. Physician burnout has real consequences for patient care quality and safety. Tools that alleviate cognitive and clerical overload can be viewed as morally favorable when balanced with accountability and patient trust.

From this perspective, the net outcome was positive, and the use of AI could be considered ethically sound, so long as oversight remains in place. Consequentialism, however, also demands continual evaluation: what is the harm if errors go unnoticed? What safeguards must exist to maximize benefit and minimize risk? In this light, ethical AI implementation isn't just about what the system does, but about how it is monitored, updated, and used by responsible clinicians.

From a broader standpoint, clinicians must remain active decision-makers, not passive recipients of AI suggestions. The physician's role must not be reduced to clicking "accept" or "reject." Instead, AI should support cognitive offloading where appropriate, while preserving the clinician's ultimate authority and accountability (Bracken et al., 2025).

II. Bias, Justice, and Health Equity

Despite their data-driven appearance, AI systems can perpetuate or even exacerbate existing inequities in healthcare. Algorithms trained on biased datasets can unintentionally produce discriminatory outcomes, particularly along lines of race, socioeconomic status, or gender identity.

A well-known example is an algorithm used in population health management that underestimated the healthcare needs of African American patients by using historical healthcare spending as a proxy for need (Obermeyer et al., 2019). Since African American patients often receive less care due to systemic disparities, the algorithm mistakenly concluded they were less in need of medical intervention.

In the case of my preceptor's AI note-taking system, the initial misidentification of gender, although quickly corrected, demonstrates how even basic demographic details can be mishandled. These errors, if left uncorrected, could propagate throughout medical records, leading to inappropriate treatments or omissions of care. Beyond the immediate clinical consequences, such inaccuracies can distort the integrity of large-scale retrospective studies, quality improvement initiatives, or machine learning models that rely on electronic health record (EHR) data (Aldosari et al., 2025). When clinical documentation becomes a data source, even small inconsistencies, such as misgendering or inaccurate exam findings, can bias population-level research, misguide clinical policy, or train future AI systems on flawed assumptions. Thus, ensuring accuracy in AI-assisted documentation is not only a clinical responsibility but also a research imperative (Cestonaro et al.,2023).

From a justice-oriented framework, fairness requires that systems be designed with attention to the least advantaged. In practice, this means auditing algorithms for bias, including diverse data in training sets, and engaging marginalized communities in the design and implementation process.

III. Data Privacy and Informed Consent

AI systems often depend on large datasets that include protected health information. The ethical use of such data hinges on consent, transparency, and privacy. While some systems claim to anonymize data, the risk of re-identification remains, particularly when data are aggregated from multiple sources.

This issue became personally relevant to me during my clinical experience with an AI-assisted documentation system. Before beginning the encounter, my preceptor took a moment to explain to the patient that we would be using an AI-powered recorder to transcribe our discussion and asked for the patient’s permission. The patient agreed without hesitation, but that brief moment of transparency stood out to me. As a student still learning how to balance technological fluency with patient-centered care, I realized how easy it would be for such tools to be used without disclosure and how important it is to maintain trust by explicitly involving the patient in decisions about their data.

Although the AI system was only used for notetaking in that encounter, it raised a broader ethical question for me: what happens to this data after the visit ends? Is it stored, shared, or used to improve the AI itself? Patients may not be aware that their data could be contributing to secondary uses such as algorithm refinement or institutional analytics. Even when data are anonymized, patients deserve to know when and how their words are being processed beyond the immediate clinical context.

By asking for the patient’s consent, my preceptor modeled an approach that respects patient autonomy and establishes an ethical baseline for technology use. This action demonstrated that transparency doesn't have to be time-consuming or complicated; it can be seamlessly integrated into the clinical workflow when clinicians are intentional. As AI becomes more embedded in healthcare, we must ensure that informed consent evolves

alongside it. Autonomy is not just about making medical decisions, but also about understanding how one's data and voice are used in the process of care.

IV. Human Connection and the Therapeutic Relationship

Medicine is inherently relational. The therapeutic alliance between clinician and patient involves more than data exchange; it involves empathy, trust, and human presence. Critics of AI in healthcare worry that technology may erode this relational aspect by mechanizing encounters or displacing face-to-face communication (Riedl et al., 2024).

In my clinical experience, the presence of AI-assisted notetaking freed my preceptor to focus more on the patient, allowing him to maintain eye contact, listen actively, and engage with the patient's narrative rather than typing notes during the visit. When used appropriately, AI can support, not hinder, the human connection by reducing clerical burden (Lin et al., 2019). But this outcome depends on thoughtful implementation.

Philosophically, this reflects a balance between utilitarian and care ethics. While efficiency matters, care ethics emphasizes responsiveness, empathy, and relational context. AI systems must be designed and deployed in ways that prioritize, not marginalize, these values.

Conclusion: Toward Ethical and Effective Integration

Artificial intelligence in healthcare holds transformative promise, but its ethical integration depends on maintaining vigilance, accountability, and core professional values. Clinicians must remain stewards of judgment, equity, and empathy. My experience working alongside a physician using AI for documentation illustrates how technology can support efficiency when paired with thoughtful oversight, but also how errors, left unmonitored, could compromise care.

To ethically guide AI integration, stakeholders must:
1. Preserve clinical authority and human judgment,
2. Design for equity and audit for bias,
3. Ensure transparency and informed consent in data use, and
4. Promote the humanistic core of medical practice.

Ethics does not require rejecting AI, it requires integrating it with wisdom and care. As medicine continues to evolve, so too must our ethical frameworks, ensuring that new tools serve timeless goals: healing, justice, and respect for the individual.

References

1. Aldosari, B., Aldosari, H., & Alanazi, A. (2025). Challenges of artificial intelligence in medicine. *Studies in Health Technology and Informatics*, *323*, 16–20. https://doi.org/10.3233/SHTI250039
2. Bracken, A., Reilly, C., Feeley, A., Sheehan, E., Merghani, K., & Feeley, I. (2025). Artificial intelligence (AI)-powered documentation systems in healthcare: A systematic review. *Journal of Medical Systems*, *49*(1), 28. https://doi.org/10.1007/s10916-025-02157-4
3. Cestonaro, C., Delicati, A., Marcante, B., Caenazzo, L., & Tozzo, P. (2023). Defining medical liability when artificial intelligence is applied on diagnostic algorithms: A systematic review. *Frontiers in Medicine*, *10*, 1305756. https://doi.org/10.3389/fmed.2023.1305756
4. Doshi, G. K., Jensen, T. L., Graziano, A., Enenmoh, C., & Lindsey, J. (2024). Use of ambient AI scribing: Impact on physician administrative burden and patient care. *JCO Oncology Practice*, *20*(10_suppl), 418. https://doi.org/10.1200/OP.2024.20.10_suppl.418
5. Lin, S. Y., Mahoney, M. R., & Sinsky, C. A. (2019). Ten ways artificial intelligence will transform primary care. *Journal of General Internal Medicine*, *34*(8), 1626–1630. https://doi.org/10.1007/s11606-019-05035-1
6. 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. https://doi.org/10.1126/science.aax2342
7. Riedl, R., Hogeterp, S. A., & Reuter, M. (2024). Do patients prefer a human doctor, artificial intelligence, or a blend, and is this preference dependent on medical discipline? Empirical evidence and implications for medical practice. *Frontiers in Psychology*, *15*, 1422177. https://doi.org/10.3389/fpsyg.2024.1422177
8. Shah, S. J., Devon-Sand, A., Ma, S. P., Jeong, Y., Crowell, T., Smith, M., Liang, A. S., Delahaie, C., Hsia, C., Shanafelt, T., Pfeffer, M. A., Sharp, C., Lin, S., & Garcia, P. (2025). Ambient artificial intelligence scribes: Physician burnout and perspectives on usability and documentation burden. *Journal of the American Medical Informatics Association*, *32*(2), 375–380. https://doi.org/10.1093/jamia/ocae295