Life Studies, 2025, 1(4), 32-43
https://doi.org/10.71204/sprd8n54

Ethical Risks in Medical Artificial Intelligence and the Normative Function of Medical Humanities: A Study of AI and Emerging Medical Technologies

Liwei Liu [1, *]

[1] Shenzhen Open University, Shenzhen 518001, China

*** Correspondence:**

Liwei Liu

liwei082313@163.com

Received: 10 December 2025/Accepted: 27 December 2025/Published online: 31 December 2025

Abstract

The rapid expansion of artificial intelligence and emerging digital technologies in medicine has fundamentally reshaped clinical decision-making, healthcare governance, and biomedical knowledge production. While medical artificial intelligence promises enhanced efficiency, diagnostic accuracy, and system optimization, it simultaneously generates complex ethical risks that challenge traditional medical norms and regulatory approaches. Existing discussions of medical AI ethics often prioritize technical safeguards, algorithmic transparency, or regulatory compliance, yet they frequently underestimate the need for deeper normative reflection on responsibility, moral agency, and the meaning of care. This paper argues that medical humanities plays an indispensable normative role in identifying, interpreting, and addressing the ethical risks embedded in medical AI applications. Focusing explicitly on artificial intelligence and frontier medical technologies, the study analyzes the structural sources of ethical risk in algorithm-driven medicine and examines how medical humanities contributes to ethical norm construction beyond procedural governance. By situating medical AI within humanistic concerns such as moral responsibility, interpretive judgment, and human dignity, the paper demonstrates that medical humanities is essential for ethically robust and socially legitimate AI-enabled healthcare.

Keywords: Medical Artificial Intelligence; Ethical Risk; Medical Humanities; Normative Framework; Emerging Medical Technologies

1. Introduction

Artificial intelligence has become an increasingly influential force in contemporary medicine. Machine learning algorithms are now widely applied in medical imaging, diagnostic decision support, risk stratification, and predictive modeling, while frontier technologies such as big data analytics, algorithmic triage systems, and digital health platforms continue to reshape healthcare

delivery and clinical governance (Topol, 2019; Rajkomar, Dean, & Kohane, 2019). These developments signal not merely a technological upgrade, but a structural transformation in how medical knowledge is generated, validated, and applied.

Alongside their potential benefits, medical AI systems introduce ethical risks that exceed the scope of traditional medical ethics. Unlike earlier medical technologies, AI systems increasingly participate in epistemic and normative dimensions of practice, influencing how clinical judgments are formed and how responsibility is distributed (Floridi et al., 2018). Decisions that were once grounded in professional expertise and interpersonal interaction are now mediated by computational processes that may be opaque, probabilistic, and detached from individual clinical narratives. In response, a rapidly growing literature on AI ethics and digital health governance has emerged. Much of this literature focuses on developing ethical principles, regulatory frameworks, and technical design requirements, such as transparency, explainability, and fairness (Jobin, Ienca, & Vayena, 2019). While these approaches are necessary, they often frame ethical risk as a problem of system malfunction or regulatory deficiency. Ethical challenges are treated as issues to be managed through compliance mechanisms rather than as symptoms of deeper normative tensions within medical practice.

This paper adopts a different analytical stance. It argues that ethical risks in medical AI are not merely operational or technical problems, but manifestations of structural transformations in the moral foundations of medicine. These transformations concern how judgment is exercised, how responsibility is attributed, and how patients are recognized as moral subjects within technologically mediated systems. Addressing such risks requires normative resources that extend beyond procedural ethics and regulatory control. Medical humanities provides precisely such resources. As an interdisciplinary field drawing on philosophy, ethics, history, and interpretive social inquiry, medical humanities conceptualizes medicine as a human practice embedded in values, meanings, and moral commitments (Evans, Ahlzén, Heath, & Macnaughton, 2016). In the context of medical AI, medical humanities contributes not by offering technical solutions, but by clarifying ethical concepts, articulating normative expectations, and sustaining critical reflection on the purposes of medical practice. Importantly, this study does not examine specific algorithms, clinical cases, or educational interventions. Instead, it focuses on ethical risk at a conceptual and structural level, asking how AI and emerging medical technologies reshape the normative landscape of medicine and how medical humanities functions as a source of ethical normativity in response. By explicitly limiting its scope to AI and frontier medical technologies, the paper aims to clarify the distinctive ethical challenges posed by computational medicine and the normative role of medical humanities in addressing them.

2. Medical Artificial Intelligence as a Source of Structural Ethical Risk

Ethical risk in medical AI does not arise solely from errors, misuse, or insufficient regulation. Rather, it is embedded in the structural characteristics of algorithmic decision-making and data-driven healthcare systems. Understanding these risks requires shifting attention from isolated

ethical dilemmas to the ways in which AI transforms the normative architecture of medical practice itself.

2.1. Algorithmic Mediation and the Transformation of Clinical Judgment

One of the most significant ethical implications of medical AI lies in its mediation of clinical judgment. AI systems increasingly influence diagnostic reasoning, prognostic assessment, and treatment selection, often by processing vast datasets and identifying statistical patterns beyond human cognitive capacity (Esteva et al., 2017). Although such systems are typically described as decision-support tools, their epistemic authority can substantially shape clinical behavior.

From an ethical standpoint, this mediation alters the character of judgment in medicine. Clinical judgment has traditionally been understood as a practice that integrates scientific knowledge, experiential insight, and moral responsibility toward individual patients (Montgomery, 2006). When algorithmic recommendations become dominant reference points, clinicians may experience a shift from active interpretive judgment to passive endorsement of system outputs.

Medical humanities highlights that judgment in medicine is not value-neutral computation, but a moral activity involving interpretation, responsibility, and responsiveness to context. Ethical risk emerges when algorithmic mediation diminishes the space for reflective judgment, thereby weakening the moral agency of clinicians even in the absence of technical error.

2.2. Opacity, Explainability, and Moral Accountability

A second structural source of ethical risk arises from the opacity of many AI systems. Machine learning models, particularly those based on deep neural networks, often lack transparent reasoning processes that can be meaningfully explained to clinicians or patients (Burrell, 2016). In medicine, where trust and accountability are central ethical norms, such opacity poses serious challenges. Accountability in medical practice presupposes the ability to provide reasons for decisions, to justify actions in moral and professional terms, and to engage in communicative relationships with patients. When clinical decisions are significantly influenced by systems that cannot offer intelligible explanations, the moral basis of accountability is undermined (London, 2019). Medical humanities contributes to ethical analysis by emphasizing that responsibility is not merely a legal or procedural concept, but a relational and narrative one. Ethical risk arises not only when harm occurs, but when the conditions for moral explanation and justification are eroded by technological opacity.

2.3. Datafication, Bias, and the Moral Recognition of Patients

AI-driven medicine relies on large-scale datafication, transforming patients into sources of data for algorithmic modeling. While this process enables predictive analytics and population-level insights, it also introduces ethical risks related to bias, exclusion, and moral recognition (Obermeyer et al., 2019).

Bias in medical AI is often discussed in statistical terms, such as unequal error rates across demographic groups. From a humanistic perspective, however, bias also represents a failure of moral recognition. When patients are primarily encountered as data points, their lived experiences, social contexts, and individual narratives may be obscured. Medical humanities underscores the

importance of narrative understanding and person-centered recognition in ethical medical practice (Charon, 2006). Ethical risk arises when data-driven abstraction conflicts with the humanistic commitment to treat patients as moral subjects rather than interchangeable instances within datasets.

3. Typologies of Ethical Risk in Medical Artificial Intelligence and Frontier Medical Technologies

Ethical risks in medical artificial intelligence are neither accidental nor merely derivative of technical malfunction. Rather, they emerge from the ways in which AI systems restructure epistemic authority, redistribute responsibility, and redefine the moral relationships at the heart of medical practice. To clarify the normative challenges posed by AI and frontier medical technologies, it is analytically useful to distinguish several interrelated types of ethical risk. These typologies do not represent isolated problems; instead, they illuminate overlapping dimensions of ethical vulnerability that require humanistic interpretation and normative judgment.

3.1. Responsibility Gaps and the Fragmentation of Moral Agency

One of the most widely discussed ethical risks associated with medical AI is the emergence of responsibility gaps. In traditional medical practice, responsibility for clinical decisions is largely attributable to identifiable agents—physicians, healthcare teams, or institutions—whose judgments can be evaluated and justified. AI systems complicate this structure by introducing distributed decision-making processes involving developers, data curators, clinicians, institutions, and automated systems (Matthias, 2004; Floridi et al., 2018). In AI-mediated medicine, outcomes often result from interactions among multiple actors and algorithmic processes, making it difficult to identify a single morally responsible agent. When adverse outcomes occur, responsibility may be diffused across technical design choices, data quality, institutional deployment decisions, and clinical use. This diffusion creates ethical uncertainty rather than simple liability failure. From the perspective of medical humanities, this risk is not merely procedural but normative. Moral agency in medicine has historically been grounded in the capacity of practitioners to deliberate, decide, and answer for their actions. Responsibility gaps threaten this moral architecture by weakening the link between action and accountability. Medical humanities emphasizes that responsibility is not exhausted by causal attribution; it involves moral authorship, narrative explanation, and ethical self-understanding (Pellegrino & Thomasma, 1993). Ethical risk thus arises when AI systems undermine the conditions under which clinicians can meaningfully assume responsibility, even if formal accountability mechanisms remain in place. Addressing such risk requires normative clarification of responsibility that cannot be achieved through technical design alone.

3.2. Algorithmic Injustice and the Reproduction of Structural Inequality

A second major category of ethical risk concerns fairness and justice. AI systems in medicine are typically trained on large datasets that reflect existing social, economic, and healthcare inequalities. As a result, algorithmic outputs may systematically disadvantage certain populations, even when systems perform well according to aggregate accuracy metrics (Obermeyer et al., 2019; Benjamin, 2019). In medical contexts, algorithmic injustice may manifest in biased risk predictions, unequal access to advanced diagnostics, or differential treatment recommendations

across demographic groups. While technical approaches to fairness seek to adjust models to reduce bias, such efforts often address symptoms rather than underlying normative issues.

Medical humanities contributes a broader ethical lens by situating algorithmic injustice within historical and social contexts. Inequality in AI systems is not merely a data problem, but a reflection of deeper patterns of exclusion and marginalization in healthcare and society. Humanistic inquiry emphasizes that justice in medicine involves moral recognition, respect for persons, and responsiveness to lived experience—not solely statistical parity (Daniels, 2008). Ethical risk arises when algorithmic systems normalize injustice under the guise of objectivity. Medical humanities challenges this normalization by foregrounding ethical questions about whose lives are valued, whose suffering is rendered visible, and whose interests are prioritized in technological design and deployment.

3.3. Value Misalignment and the Reconfiguration of Medical Ends

Beyond responsibility and justice, AI introduces ethical risk through value misalignment. AI systems are typically optimized for specific objectives, such as efficiency, accuracy, or cost reduction. While these goals may align with certain institutional priorities, they do not necessarily correspond to the normative ends of medicine, which include care, compassion, and the promotion of human flourishing (Beauchamp & Childress, 2019). In AI-driven healthcare systems, there is a risk that technical optimization subtly reshapes the goals of medical practice. For example, prioritizing predictive accuracy may marginalize contextual judgment, while emphasizing efficiency may reduce time for patient engagement. These shifts do not require explicit ethical transgression; they emerge gradually through system design and performance metrics.

Medical humanities plays a critical role in identifying such value shifts. By interrogating the purposes of medicine and the meanings of care, humanistic analysis reveals ethical risks that are invisible to purely instrumental frameworks. Value misalignment becomes ethically significant not when harm is immediate, but when the orientation of practice drifts away from its humanistic foundations (Verghese, Shah, & Harrington, 2018).

This type of ethical risk underscores the importance of normative reflection prior to and alongside technological deployment. Medical humanities does not oppose optimization per se, but insists that technological goals be evaluated against broader ethical commitments that define medicine as a moral practice.

3.4. Dehumanization and the Erosion of Person-Centered Care

A further ethical risk associated with medical AI is the potential dehumanization of care. As AI systems process patients primarily as data profiles, risk scores, or diagnostic categories, there is a danger that individual persons are reduced to algorithmically legible attributes. Such reduction does not require malicious intent; it can arise from the routine operation of data-driven systems.

Dehumanization in this context refers not to overt mistreatment, but to the gradual erosion of relational and narrative dimensions of care. When clinical encounters are structured around algorithmic outputs, patient stories, emotions, and subjective meanings may be sidelined. Medical

humanities has long emphasized that illness is not merely a biological event, but a lived experience that requires interpretive engagement (Kleinman, 1988).

Ethical risk arises when technological mediation diminishes the moral salience of personhood. From a humanistic standpoint, ethical medicine requires recognition of patients as agents with values, histories, and social identities. AI systems that privilege abstraction over interpretation threaten this recognition, even when clinical outcomes improve.

3.5. Moral Deskilling and the Atrophy of Ethical Judgment

Medical AI introduces ethical risk through the phenomenon of moral deskilling. As clinicians increasingly rely on algorithmic recommendations, opportunities for exercising ethical judgment may diminish. Over time, this reliance can lead to reduced confidence or capacity for independent moral reasoning, particularly in complex or ambiguous situations (Coeckelbergh, 2020). Moral deskilling differs from technical dependency. It concerns the erosion of practical wisdom—the ability to navigate uncertainty, weigh competing values, and respond sensitively to unique circumstances. Medical humanities conceptualizes such wisdom as central to ethical medical practice, cultivated through reflection, experience, and engagement with human narratives. Ethical risk emerges when AI systems displace rather than support moral deliberation. Even if systems perform accurately, their dominance may weaken the moral agency of practitioners, reducing ethics to compliance with system outputs. Addressing this risk requires normative frameworks that reaffirm the role of human judgment in ethically charged decisions.

4. The Normative Functions of Medical Humanities in Governing Ethical Risk

While Part II analyzed the typologies of ethical risk generated by medical artificial intelligence, ethical risk identification alone is insufficient for normative governance. Risk analysis must be accompanied by normative resources capable of interpreting moral meaning, articulating value commitments, and guiding responsible action in contexts of uncertainty. This section argues that medical humanities performs four interrelated normative functions in the governance of ethical risk in medical AI: conceptual clarification, value articulation and prioritization, reconstruction of moral responsibility, and preservation of human dignity and meaning. Together, these functions enable ethical governance to move beyond procedural compliance toward substantive moral orientation.

4.1. Conceptual Clarification: Interpreting Ethical Risk beyond Technical Vocabulary

One of the primary normative contributions of medical humanities lies in conceptual clarification. Ethical debates surrounding medical AI are often conducted using technical or regulatory language, such as “accuracy,” “robustness,” “bias,” or “explainability.” While these terms are indispensable for system design and oversight, they do not exhaust the moral significance of AI-mediated medical practice.Medical humanities introduces interpretive concepts—such as judgment, care, responsibility, and personhood—that allow ethical risks to be understood as moral phenomena rather than technical anomalies. For example, algorithmic opacity is frequently discussed as a problem of explainability. From a humanistic perspective, however, opacity also implicates the ethical requirement of reason-giving in medicine, which underpins trust, consent, and accountability (London, 2019). By clarifying such concepts, medical

humanities reframes ethical risk as a challenge to the moral grammar of medical practice. This reframing is essential for avoiding category errors in ethical governance, where moral concerns are mistakenly treated as engineering problems. Conceptual clarification thus serves as the first step in normative governance, enabling stakeholders to recognize what is ethically at stake when AI systems are introduced into medical contexts.

4.2. Value Articulation and Prioritization in AI-Mediated Medicine

A second normative function of medical humanities is the articulation and prioritization of values. Medical AI systems are inherently value-laden, as they embed assumptions about what outcomes matter, whose interests are prioritized, and which trade-offs are acceptable. However, these values are often implicit, embedded in optimization targets, performance metrics, or institutional incentives. Medical humanities provides a framework for making such values explicit and subject to ethical deliberation. Drawing on ethical theory and philosophical reflection, it interrogates how values such as efficiency, accuracy, equity, compassion, and respect for autonomy interact—and sometimes conflict—in AI-mediated medicine (Beauchamp & Childress, 2019). Crucially, medical humanities does not merely list values; it contributes to value prioritization. In situations where AI systems optimize for system-level efficiency at the expense of individual patient experience, humanistic inquiry raises questions about the proper ends of medicine. It emphasizes that technological success does not automatically translate into ethical legitimacy. Value articulation grounded in medical humanities thus provides normative orientation for evaluating AI systems not only by what they achieve, but by what they are for.

4.3. Reconstruction of Moral Responsibility under Algorithmic Mediation

A third normative function concerns the reconstruction of moral responsibility. As discussed in Part II, medical AI generates responsibility gaps by fragmenting agency across human and non-human actors. Technical approaches to responsibility often focus on assigning liability or defining accountability chains. While necessary, these approaches risk reducing responsibility to legal compliance. Medical humanities contributes a richer understanding of responsibility as a moral practice rather than a procedural allocation. Responsibility in medicine involves attentiveness to patient vulnerability, willingness to answer for one's decisions, and engagement in moral reasoning under uncertainty (Pellegrino & Thomasma, 1993). When AI systems mediate decision-making, these dimensions of responsibility are not eliminated, but transformed. Through ethical reflection and narrative analysis, medical humanities helps reconstruct responsibility in AI-mediated contexts by reaffirming the role of human agents as moral interpreters of technological outputs. Rather than treating AI recommendations as authoritative commands, clinicians are understood as ethically responsible for contextualizing, interpreting, and, when necessary, resisting algorithmic guidance. This reconstruction preserves moral agency without denying the epistemic contributions of AI.

4.4. Preserving Human Dignity and Meaning in Data-Driven Care

A fourth and foundational normative function of medical humanities is the preservation of human dignity and meaning in data-driven medicine. AI systems tend to abstract patients into datasets, risk profiles, and predictive scores. While abstraction is a necessary feature of large-

scale analytics, it carries the ethical risk of dehumanization. Medical humanities counters this risk by emphasizing the narrative and experiential dimensions of illness. Illness is not merely a biological deviation, but a disruption of lived meaning that affects identity, relationships, and moral self-understanding (Kleinman, 1988). Ethical medicine therefore requires engagement with patients as persons whose experiences cannot be fully captured by algorithmic representation. By foregrounding dignity and meaning, medical humanities provides a normative counterweight to the instrumental rationality of AI systems. It insists that ethical governance must ensure that technological mediation does not erode the relational foundations of care. This function is particularly critical in frontier medical technologies, where automation risks distancing practitioners from the human realities of illness.

4.5. Medical Humanities as a Normative Mediator between Technology and Ethics

Taken together, these normative functions position medical humanities as a mediator between technological rationality and ethical responsibility. Rather than opposing AI or rejecting technological innovation, medical humanities interprets and evaluates AI within a broader moral horizon. It translates technical developments into ethical questions and transforms abstract values into concrete normative guidance. This mediating role is especially important in contexts where ethical governance risks becoming proceduralized. Checklists, principles, and compliance frameworks are necessary but insufficient for addressing the moral complexity of AI-driven medicine. Medical humanities sustains ethical reflection by maintaining attention to meaning, judgment, and human vulnerability. In this sense, medical humanities does not function as an external constraint on technological progress. It operates as an internal normative resource that shapes how medicine understands itself in the age of artificial intelligence. By enabling conceptual clarity, value articulation, responsibility reconstruction, and dignity preservation, medical humanities contributes to an ethical framework capable of governing medical AI in a manner consistent with the moral foundations of medicine.

5. Toward a Humanistically Grounded Ethical Framework for Medical AI Governance

(1) From Risk Identification to Normative Governance

The preceding sections have demonstrated that ethical risks in medical artificial intelligence are not incidental side effects of technological innovation, but structural features of algorithm-driven medicine. Responsibility gaps, algorithmic injustice, value misalignment, dehumanization, and moral deskilling arise from the ways in which AI systems reorganize epistemic authority, moral agency, and professional practice. Addressing these risks requires more than reactive regulation or technical safeguards. Normative governance differs fundamentally from risk management. Whereas risk management seeks to minimize harm through control mechanisms, normative governance aims to orient technological development toward ethically legitimate ends. In the context of medical AI, such orientation must be grounded in an understanding of medicine as a moral practice rather than a purely technical enterprise. This is precisely where medical humanities plays a central role. Medical humanities enables a transition from fragmented ethical responses to an integrated normative framework. By clarifying ethical concepts, articulating

values, reconstructing responsibility, and preserving human dignity, it provides the intellectual infrastructure necessary for sustained ethical governance in AI-mediated medicine.

(2) Core Normative Principles Informed by Medical Humanities

A humanistically grounded ethical framework for medical AI governance does not replace existing principles such as safety, transparency, or accountability. Instead, it deepens and contextualizes them by embedding these principles within a broader moral horizon. Several core normative orientations emerge from medical humanities scholarship. First, interpretive responsibility must be recognized as a foundational norm. AI systems do not absolve clinicians of moral responsibility; rather, they transform the conditions under which responsibility is exercised. Clinicians remain ethically accountable for interpreting algorithmic outputs in light of patient context, values, and vulnerability. This principle resists the moral displacement that can accompany automation. Second, person-centered moral recognition must guide AI deployment. Ethical governance requires that patients be treated not merely as data sources or risk profiles, but as persons with narratives, identities, and moral claims. Medical humanities emphasizes that ethical medicine involves responsiveness to lived experience, which cannot be fully captured by predictive models. Third, value reflexivity must be institutionalized. AI systems embed implicit value hierarchies through optimization goals and performance metrics. A humanistically informed framework requires continuous reflection on whether these embedded values align with the moral purposes of medicine. Efficiency and accuracy, while important, must be evaluated against commitments to care, equity, and human flourishing. Fourth, ethical humility should guide technological ambition. Medical humanities reminds us that uncertainty, ambiguity, and moral complexity are intrinsic to medicine. Ethical governance must therefore resist overconfidence in technological solutions and preserve space for doubt, dialogue, and moral deliberation.

(3) Integrating Medical Humanities into AI Ethical Governance Structures

The normative contributions of medical humanities cannot remain abstract. For ethical governance to be effective, these contributions must inform governance structures at multiple levels, including policy formulation, institutional oversight, and professional self-regulation. At the policy level, medical humanities provides conceptual resources for framing ethical guidelines that go beyond procedural checklists. Instead of treating ethics as an external constraint, policy frameworks can acknowledge the moral purposes of medicine and the interpretive responsibilities of practitioners.

At the institutional level, ethical governance bodies overseeing AI deployment benefit from humanistic expertise capable of interpreting ethical risk in context-sensitive ways. Medical humanities scholars contribute to ethical deliberation not by offering definitive answers, but by sustaining critical reflection on meaning, value, and responsibility.

At the professional level, ethical governance requires reinforcing the moral agency of clinicians. AI systems should be positioned as epistemic aids rather than normative authorities. Medical humanities supports this positioning by articulating the ethical dimensions of judgment and care that remain irreducibly human.

(4) Avoiding Instrumentalization of Ethics

A significant danger in AI ethics discourse is the instrumentalization of ethics itself. Ethical principles risk becoming tools for legitimizing technological deployment rather than frameworks for critical evaluation. When ethics is reduced to compliance, its normative force is weakened. Medical humanities resists this instrumentalization by insisting that ethics is an ongoing interpretive practice rather than a set of static rules. Ethical governance, from this perspective, is not a one-time certification process but a continuous engagement with evolving moral questions. This orientation is especially important in frontier medical technologies, where rapid innovation outpaces formal regulation. By maintaining ethical reflection as a living practice, medical humanities helps ensure that AI governance remains responsive to human values rather than subordinated to technological momentum.

6. Conclusion

This paper has argued that ethical risks in medical artificial intelligence are structural challenges that cannot be adequately addressed through technical design or regulatory compliance alone. AI and emerging medical technologies transform the normative foundations of medicine by reshaping judgment, responsibility, and human recognition. In this context, medical humanities plays an indispensable normative role. Through conceptual clarification, value articulation, responsibility reconstruction, and the preservation of human dignity and meaning, medical humanities provides the ethical orientation necessary for governing medical AI in a morally legitimate manner. Rather than opposing technological innovation, it mediates between technological rationality and human values, ensuring that medicine remains a humane practice in an age of artificial intelligence. As AI continues to expand its influence across medical domains, the relevance of medical humanities will only increase. Ethical governance grounded in humanistic reflection offers not a constraint on innovation, but a condition for its moral sustainability. Recognizing this normative function is essential for ensuring that the future of medical AI serves human well-being rather than undermining the ethical foundations of medicine itself.

Author Contributions:

All authors have read and agreed to the published version of the manuscript.

Funding:

This research received no external funding.

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Not applicable.

Data Availability Statement:

Not applicable.

Conflict of Interest:

The authors declare no conflict of interest.

References

Beauchamp, T. L., & Childress, J. F. (2019). Principles of biomedical ethics (8th ed.). Oxford University Press.

Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim Code. Polity Press.

Burrell, J. (2016). How the machine "thinks": Understanding opacity in machine learning algorithms. Big Data & Society, 3(1), 1–12.

Charon, R. (2006). Narrative medicine: Honoring the stories of illness. Oxford University Press.

Coeckelbergh, M. (2020). AI ethics. MIT Press.

Daniels, N. (2008). Just health: Meeting health needs fairly. Cambridge University Press.

Esteva, A., Kuprel, B., Novoa, R. A., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118.

Evans, M., Ahlzén, R., Heath, I., & Macnaughton, J. (2016). Medical humanities. In E. J. Cassell & T. J. Buchanan (Eds.), The Oxford handbook of medical ethics and law (pp. 463–482). Oxford University Press.

Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People—An ethical framework for a good AI society. Minds and Machines, 28(4), 689–707.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389–399.

Kleinman, A. (1988). The illness narratives: Suffering, healing, and the human condition. Basic Books.

London, A. J. (2019). Artificial intelligence and black-box medical decisions: Accuracy versus explainability. Hastings Center Report, 49(1), 15–21.

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. Ethics and Information Technology, 6(3), 175–183.

Montgomery, K. (2006). How doctors think: Clinical judgment and the practice of medicine. Oxford University Press.

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.

Pellegrino, E. D., & Thomasma, D. C. (1993). The virtues in medical practice. Oxford University Press.

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14), 1347–1358.

Topol, E. (2019). Deep medicine: How artificial intelligence can make healthcare human again. Basic Books.

Verghese, A., Shah, N. H., & Harrington, R. A. (2018). What this computer needs is a physician. JAMA, 319(1), 19–20.