Ethics of Medical AI
Giovanni Rubeis
Springer, 2024 264 pages

Ethics of Care in the Era of Algorithms: Giovanni Rubeis' Ethical Blueprint for Medical AI

Chelsea Ma, Stanford University

In *Ethics of Medical AI*, Giovanni Rubeis, bioethicist and professor at Greifswald Medical School in Germany, explores the intersection of artificial intelligence and medical ethics, offering a nuanced analysis of how AI is reshaping healthcare. Examining the complex implications of Medical Artificial Intelligence (MAI), Rubeis focuses on its impact on clinical practices, physician–patient relationships, and the environments—personal, professional, and urban—in which healthcare operates. The book provides practical strategies for navigating AI-related challenges for a diverse audience, including healthcare professionals, policymakers, engineers, researchers, and students of medicine, social sciences, and philosophy.

Divided into two parts, *Ethics of Medical AI* first lays the theoretical groundwork necessary for the ethical analyses presented in the second half. Rubeis explores the history and definitions of AI and MAI, as well as the rise of evidence-based medicine (EBM)—a framework that MAI can enhance, but also one that may create tension when technological approaches conflict with the humanistic aspects of care. This foundational section equips readers with essential context to engage with the ethical arguments that follow. Rubeis skillfully connects technical and philosophical concepts while grounding them in the realities of clinical practice. His discussion of EBM highlights both the potential of MAI to improve clinical decision-making and the risks of reductionism and depersonalization in patient care.

While MAI is designed to enhance precision and efficiency in healthcare, its integration poses significant risks, including algorithmic bias, shifts in responsibility within physician–patient relationships, and the expansion of medicalization through constant surveillance. Rubeis introduces the concept of "smart data practices" to describe how AI organizes and aggregates data to improve predictive analytics and decision-making tools. However, these same practices can raise ethical concerns, especially regarding bias in data collection and algorithm design. He warns that "the datafication of patients risks reinforcing existing disparities unless deliberate efforts are made to address bias in data collection and algorithm design" (p. 117). This critique is central to his call for transparency and equity in the development of MAI systems. As a result, Rubeis underscores the need for frameworks that prioritize fairness and justice, especially for populations already marginalized within healthcare systems and vulnerable to inequities.

Equally compelling is Rubeis' exploration of how MAI affects relationships in healthcare. Trust and empathy, he argues, are the foundation of effective medical practice, yet these qualities are at risk when care becomes mediated by AI. He asks, "How can trust be cultivated when the therapeutic relationship is shared with an artificial agent?" (p. 62). By raising this question, Rubeis challenges readers to consider the limitations of MAI in replicating the relational aspects of care that are essential to patient well-being. While MAI may streamline workflows and improve diagnostic accuracy, Rubeis cautions against allowing these technological gains to overshadow compassion, shared decision-making, and person-centered care—core tenets of medical ethics. He advocates for models of MAI that "enhance, rather than replace, the human elements of care" (p. 172), ensuring that technology supplements, rather than undermines, the values at the heart of medicine.

The ethical implications of MAI extend beyond individual interactions, as Rubeis illustrates in his discussion of healthcare environments. He examines the rise of wearable health technologies and smart home devices, noting how these innovations blur the boundaries between clinical and private spaces and raise critical questions about privacy and agency. Rubeis writes, "When the clinical gaze is not limited to medical institutions, but penetrates the private realm, this border becomes permeable, thus transforming the home itself. The extension of the clinical gaze may change the landscape of the home and with

it its moral geography, meaning the social practices and relationships that constitute it" (p. 232). These reflections emphasize the importance of considering how MAI reshapes not only healthcare delivery but also the lived experiences of patients whose personal spaces and routines may increasingly be monitored by third parties—often without meaningful consent.

Rubeis stresses that while MAI may appear autonomous, it is ultimately designed by humans and reflects the choices, biases, and priorities of its creators. This insight reframes discussions of accountability. Ethical issues in MAI, he argues, stem from systemic flaws in data collection, algorithm development, and healthcare infrastructure. "The delegation of tasks to MAI does not absolve humans of their ethical responsibilities; rather it demands even greater scrutiny of the decisions embedded in the technology" (p. 161). By situating MAI within broader societal and institutional contexts, Rubeis ensures that evaluating AI ethics extends beyond technical analysis to include the social determinants shaping healthcare outcomes.

A central contribution of the book is Rubeis' critique of solutionism—the belief that technological innovation alone can solve systemic problems. He argues that effective ethical frameworks must integrate social, economic, and institutional dimensions to address the root causes of health disparities. This critique, coupled with his analysis of technological reductionism—the tendency to reduce patients to data points—reinforces his call for a more holistic approach to MAI. Rubeis advocates incorporating narrative

medicine and humanistic values into MAI design and implementation, ensuring that technology supports, rather than replaces, essential human practices.

The book concludes with a practical guide for stakeholders, offering seven "lessons" for the responsible implementation of MAI. His use of the term "lessons" is intentional—it suggests adaptability and reflection, rather than rigid directives. These lessons emphasize transparency, accountability, and the inclusion of social determinants in healthcare decision-making. This practical framework highlights the book's relevance and applicability across sectors, bridging theoretical ethics and real-world challenges. The result is a valuable resource not only for academics and ethicists but also for engineers, policymakers, and healthcare providers shaping the future of medicine.

Ethics of Medical AI is a thoughtful, timely, and comprehensive work that engages deeply with the ethical challenges posed by the integration of AI into healthcare. Professor Rubeis' interdisciplinary approach and balanced analysis ensure that the book resonates with a broad audience, making it an indispensable guide for those invested in the future of ethical, human-centered healthcare.