DOI: 10.1111/bioe.13000

EDITORIAL

bioethics WILEY

# Who's next? Shifting balances between medical AI, physicians and patients in shaping the future of medicine

The hopes and even, to some extent, the hype that is commonly associated with artificial intelligence (AI) has long since found its way into healthcare. Starting with the first approaches of knowledge modeling for computer-aided medical decision-making as early as the 1980s, AI technologies in medicine have become more powerful and more diverse. Since technology as such is neither a blessing nor a curse, one might ask: Why does the increasing implementation of AI in medical care seem even more controversial than in other terrains? Technology's initial ethical neutrality holds, in principle, both for the most mundane mechanical machinery and for the most advanced AI. However, the application of emerging medical AI challenges us in quite specific ways, as it may have a potentially game-changing impact on some of the most delicate, existential experiences of humankind. For instance, the advent of artificial moral agency in medicine might introduce an additional agent into the patient–physician relationship. What is more, this is an artificial agent that, by generating and processing potentially decisive information for medical decision-making based on epistemological assumptions[1] and models that remain largely opaque to *both* patient and physician, might crucially alter the dynamics between patient and physician. This alteration, in turn, may affect trust in the medical profession, health literacy, and informed consent.

These and related issues appear even more pressing vis-à-vis the predominantly normativist conception in bioethics (endorsed by the World Health Organization), according to which health is a sensitive, personal endeavor, centered around what most people value highly: the duration and quality of their lives. What this means in detail, however, not only varies greatly between different people but also is very much dependent on social and cultural contexts and meanings. Accordingly, medical decisions—including those generated by medical AI—are culturally embedded and involve personal, social and cultural values and preferences, not just clinical information. Providing and receiving patient-centered healthcare, then, requires both medical expertise and a culturally sensitive, trusting relationship between patient and physician that fosters treatment shaped and informed by patients' values.

Now, human care is inherently relational—an insight that is acknowledged by the recent focus on relational autonomy in medical ethics.[2] Medical expertise, in both theory and clinical practice, thus requires more than textbook knowledge and prudential reasoning. It requires genuinely human capabilities such as empathy, compassion, and intuition, all of which seem to be realized by interacting biological beings. It is widely known, for example, that hormone release induced by human touch is conducive to healing;[3] touch starvation or "skin hunger," on the other hand, can have devastating effects on both physiological health parameters and subjective well-being. This implies that medical AI can be no surrogate for human care, since the relationship between patients and human caregivers is different in kind (not just in degree) from the relationship between patients and (embodied) medical AI such as care robots. It seems as though we would be well advised, then, to keep tabs on the ontological and moral differences between interacting people in healthcare and people interacting with medical AI.

It stands to reason that a great many potential areas of application of medical AI (including prevention, monitoring, diagnosis, and treatment recommendations) are paralleled by an unprecedented increase in (moral) responsibility regarding medical AI's performance and healthcare professionals' duty of care. Nonetheless, when it comes to improving both the quality and efficiency of healthcare, the emergence of medical AI holds great promise, and therefore calls for methodologies through which we are able to tackle medical AI's profound ethical and conceptual challenges.

Despite the rapid advance of medical AI and its progressively pervasive clinical application, there is, as yet, surprisingly little systematic bioethical work on fundamental ethical and conceptual issues—most current work focuses narrowly on the ethical implications of specific applications. The present special issue, *Promises and Challenges of Medical AI*, makes a first attempt at overcoming some of these shortcomings in the current bioethical treatment of medical AI.

This collection comprises 11 articles by scholars from across the globe, covering a wide range of cutting-edge topics. By and large, the articles focus on either conceptual, epistemological issues of medical AI or on the various ethical and normative challenges that come with medical AI's clinical application and medical decision-making.

[1]Paul, N. W. (1998). Incurable suffering from the "Hiatus Theoreticus"? Some epistemological problems in modern medicine and the clinical relevance of philosophy of medicine. *Theoretical Medicine and Bioethics*, *19*(3), 229–251.

[2]Walensi, M., Inthorn, J., & Paul, N. W. (2016). Willensfreiheit, Determinismus und die Abwägung eines Vorab Erklärten Autonomen Willens im Falle einer Natürlichen Willensäußerung. *Internationale Zeitschrift für Philosophie und Psychosomatik*, *1*, 1–16; Böttcher, B., & Paul, N. W. (2013). Personale Autonomie: Diskussion eines Zentralen Ethischen Konzepts am Beispiel von Fertilitätsprotektiven Maßnahmen bei Krebspatientinnen. *Ethik in der Medizin*, *25*(1), 47–59.

[3]Field, T. M. (2010). Touch for socioemotional and physical well-being: A review. *Developmental Review*, *30*, 367–383; Tabatabaee, A., Tafreshi, M. Z., Rassouli, M., Aledavood, S. A., AlaviMajd, H., & Farahmand, S. K. (2016). Effect of therapeutic touch in patients with cancer: A literature review. *Medical Archives (Sarajevo, Bosnia and Herzegovina)*, *70*(2), 142–147.

  

Whereas current bioethical work has tended to shy away from discussing controversial conceptions of potentially different meanings of medical AI in specific national and cultural contexts, Weissglass[4] highlights the contextual bias that comes with medical AI in low- and middle-income countries. Grote and Berens[5] shed light on how the potentially competitive relationship between medical AI and clinicians can be reshaped toward a more collaborative approach. Alvarado,[6] Sterckx et al.,[7] as well as Ursin et al.[8] tackle different epistemological challenges that arise in the tension between medical AI's opacity in generating decisions on the one hand and the urge to make such decisions intelligible to both patients and physicians on the other. Trust is generally thought of as a uniquely human capacity that requires agents to take intentional stances toward each other, which makes Starke et al.'s[9] defense of trust in medical AI all the more intriguing. On the more clinical side, Holm[10] brings confidence and fairness as central normative concepts to the forefront of assessing medical AI. Sand et al.[11] argue that physicians' requirements for a responsible implementation of medical AI go beyond what can be achieved through ethics by design. Kühler[12] makes the case for ethically sound mobile health applications, reinvigorating traditional bioethical controversies about paternalism and respect for autonomy. Apostolova and Lanoix[13] draw attention to the importance of care-worker-centered robotic aids to alleviate suffering in long-term care. Schwan and Palmer[14] venture that employing medical AI might help to mitigate shame-induced barriers to healthcare.

In sum, the articles of this collection demonstrate, for one thing, that algorithmic and machine-learning-based problem-solving are implemented by medical AI systems with an increasing level of success in performing increasingly complex tasks (e.g., haptic feedback during surgical interventions). The results thus yielded by medical AI are often superior to human medical care in terms of process quality; but they are not necessarily superior in terms of overall treatment outcome. This is a discrepancy that once again points to the danger of banking on overly technocratic AI solutions to healthcare. A thorough discussion of medical AI appears indispensable when it comes to the future of medicine as such. This future will most possibly have to rely on AI. However, AI cannot and should not be regarded as a general problem-solving tool; but rather as a toolbox to further improve prevention, medical intervention, and rehabilitation. It will not be a responsible practice in the field of prevention if we do not use AI to analyse the vast amount of real-world data to improve our ability to manage public health and public health crises, such as the ongoing pandemic. It will be difficult to explain why complex pattern recognition as performed in pathology or radiology should not be supported by a never tired and always more precise AI. However, the human factor, the intuition that is needed to differentiate novelties and rarities from technological artefacts of diagnostic tools, and, last but not least, the healing potential of the non-technological dimension of medical practice need to be further cultivated if we do not want to lose probably the most important dimension of medicine: care. We need to discuss the potentials of AI and our imagination of a future medicine, now. The present collection of articles attempts to take a first step in this direction.

Nils-Frederic Wagner[1]
Mita Banerjee[2]
Norbert W. Paul[1]

[1]*Institute for History, Theory and Ethics of Medicine, University of Mainz Medical Center, Mainz, Germany*
[2]*Department of English and Linguistics, University of Mainz, Mainz, Germany*

**Correspondence**
Nils-Frederic Wagner, Institute for History, Theory and Ethics of Medicine, University of Mainz Medical Center, Mainz, Germany.
Email: n.wagner@uni-mainz.de

[4]Weissglass, D. (2021). Contextual bias, the democratization of expertise, and medical AI in low and middle income countries. *Bioethics*.
[5]Grote, T., & Berens, P. (2021). How competitors become collaborators—Bridging the gap between machine learning algorithms and clinicians. *Bioethics*.
[6]Alvarado, R. (2021). Pigeons, neural networks, epistemic opacity and error in medical AI. *Bioethics*.
[7]Sterckx, S., Pierce, K. R., & Van Biesen, W. (2021). A riddle, wrapped in a mystery, inside an enigma: How semantic black boxes and opaque artificial intelligence confuse medical decision-making. *Bioethics*.
[8]Ursin, F., Timmerman, C., & Steger, F. (2021). Explicability of artificial intelligence in radiology: An additional value to the four bioethical principles? *Bioethics*.
[9]Starke, G., van den Brule, R., Elger, B., & Haselager, P. (2021). Intentional machines: A defense of trust in medical AI. *Bioethics*.
[10]Holm, S. (2021). Assessing medical AI for clinical use: Confidence and fairness. *Bioethics*.
[11]Sand, M., Durán, J., & Jongsma, K. (2021). Responsibility beyond design: Physicians' requirements for ethical medical AI. *Bioethics*.
[12]Kühler, M. (2021). Exploring the phenomenon and ethical issues of AI paternalism in health. *Bioethics*.
[13]Apostolova, I., & Lanoix, M. (2021). The importance of developing care-worker-centered robotic aides in long-term care. *Bioethics*.
[14]Schwan, D., & Palmer, A. (2021). Beneficent dehumanization: Employing artificial intelligence and carebots to mitigate shame-induced barriers to medical care. *Bioethics*.