Preliminary mapping study. *Journal of Medical Internet Research* 24 (1):e33081. doi:10.2196/33081.
Möllmann, N. R., M. Mirbabaie, and S. Stieglitz. 2021. Is it alright to use artificial intelligence in digital health? A systematic literature review on ethical considerations. *Health Informatics Journal* 27 (4). doi:10.1177/14604582211052391.

THE AMERICAN JOURNAL OF BIOETHICS
2023, VOL. 23, NO. 9, 84–86
https://doi.org/10.1080/15265161.2023.2237437

Taylor & Francis
Taylor & Francis Group

OPEN PEER COMMENTARIES

AI-Based Medical Solutions Can Threaten Physicians' Ethical Obligations Only If Allowed to Do So

Benjamin Gregg[a]

University of Texas at Austin

Mildred Cho and Nicole Martinez-Martin (2023) distinguish between two of the ways in which humans can be represented in medical contexts. One is technical: a digital model of aspects of a person's health. The other concerns the moral worth of actual persons. Moral value is vulnerable to neglect, denial, or violation by the technical approach. The authors concede that the dangers posed by relevant technologies follow not from human intention but from AI-based technologies inadvertently reproducing, even exacerbating, various biases and inequalities already present in society.

Yet Cho and Martinez-Martin also portray AI as a tool likely to escape the control of the toolmakers. This strand of their argument over-estimates the capacity of AI-technologies to make themselves independent of the humankind that seeks their benefits. Correspondingly, their argument does not speak adequately to the ways in which physicians can regulate their use of digitalized medical solutions toward recognizing and protecting the moral worth of patients. They might redress their over-estimation of AI's dangers by revising their notion of representation along three dimensions.

First, the authors claim that AI-based approaches "shift moral obligations in health research": "away from traditional biomedical scientific methods and the logic of clinical reasoning" and "away from the people on whom the simulations are based." They contend that AI-based approaches easily lead us to what they call a "data-first approach" that favors "digital simulations." The notion of such a distinction——between morally sensitive "clinical reasoning" and amoral digital simulations——cannot be sustained with respect to issues of the digital representation of patients. A model of a person's health condition does not of itself entail the analyst's lack of moral concern for the person modeled. The technical representation of a health condition treats the body as an object. But a competent physician can both treat the body as an object of technical analysis and at the same time grasp that it is always also a morally relevant human subject (He et al. 2019).

Second, the authors assert that an institutional "focus on technical solutions and individual actors ignores the systems issues, practices and discourse that contribute to bias" and so "lead[s] to misrepresentation in digital simulacra." But a competent physician can deploy AI in analyzing an individual while cognizant of that person's social embeddedness——for example, cognizant of how the individual may have been victimized by various social inequalities and prejudices some of which may be relevant to medical diagnosis. The authors tend to dismiss this capacity by asserting that "Human biases such as those introduced by underlying inequalities in health care and societal prejudices" are both "difficult to identify" and can be "amplified by AI." Identifying often subtle social determinants of an individual's health is a technical task that requires interdisciplinary analysis

CONTACT Benjamin Gregg bgregg@austin.utexas.edu Government, University of Texas at Austin, Austin, TX, USA.

combining, say, sociological scholarship with clinical insights. Such identification will be more difficult in some cases than others. But the authors cannot show that such identification is always difficult as such, or that it is necessarily difficult in ways that follow from the technical properties of AI-based medical analysis (Ferryman 2020).

The authors also claim that "features of digital simulacra have the potential to increase bias" and to "obscure[e] values and inequities that are embedded in the decisions made throughout the design process." We know from experience that scientists and engineers may unknowingly embed social prejudices in developing AI-systems. But we also know that the goal of bias-free AI need not be defeated by this enduring danger. Researchers are well able to identify the biases where they manifest themselves and to make corrections in design or programming (Ahmad, Eckert, and Teredesai 2018).

Third, the vulnerabilities in Cho and Martinez-Martin's analysis have one source. The authors reify AI-based virtual representation of bodies as the "epistemic culture of data science." This they contrast with a clinical notion of the body subject to etiological causes. That is, they treat an abstraction——the general cognitive interests and perspectives typical of a particular technical endeavor, in this case data science——as if it somehow constrained persons entertaining it to think in very particular ways. They assert that it brings data scientists to "shift moral attention from actual patients and principles, such as equity, to simulated patients and patient data." But a medical professional who deploys a digital model of a patient need not confuse the model with reality. She need not displace bioethical principles of clinical practice with patient data (Mittelstadt 2019).

In short, the authors assert that one kind of representation——digital models of physical beings, models that stress accuracy with respect to data——necessarily precludes another kind of representation: epidemiological models oriented on justice with respect to some of the social determinants of health within the relevant population. These determinants range from socioeconomic to ethnic or racial factors, and from lifestyle to environmental features. But the cognitive act of representing does not require the analyst to exclude alternative acts of representation, nor does it constrain her to represent in only certain ways (Benedek and Fink 2019).

The project of Cho and Martinez-Martin might yield greater analytic insights, and perhaps even practical guidance for medical professionals, if re-formulated so as not to reify the notion of representation. The authors could do so by conceding that digital medical solutions pose no inherent threat to patients because social bias is not primarily a computational or algorithmic phenomenon (Rasheed et al. 2022; Dey et al. 2022). It is a product of institutions, cultural inheritances, poverty, and other environments that produce and perpetuate social inequities as well as some health disparities. Hence AI-based medical practices can threaten physicians' ethical obligations only if allowed to do so.

Note three corollaries to this conclusion: (a) While AI may generate unwanted, unintended consequences, the potential moral and legal challenges that AI poses derive from inadequate precautionary measures by humans, not from features of AI as such. (b) Responsibility for failures of AI to meet normative standards for the treatment of human beings resides with human beings. After all, the moral capacity of human intelligence, necessary for responsibility-taking, depends on a conscious mind in the sense of a relationship among brain, body, and environment. AI has no such capacity, not only because it lacks this relationship. (c) The moral capacity of human cognition is the capacity for a mutual attribution of responsibility among members of political community. Outsourcing, to AI, moral and legal responsibility for social conditions that affect citizens adversely would undermine the politics of mutual responsibility (Gregg 2022a). That politics is a core feature of liberal democratic community. Such community is the venue for identifying and addressing the biases and inequalities that AI-based technologies can inadvertently reproduce (Gregg 2022b).

The project of identifying AI-based medical solutions that threaten physicians' ethical obligations would then ask: How are real bodies to be digitally represented such that all members of the population benefit from these rapidly developing technologies equitably? This question is not about the nature of AI-based representation. It is about the just distribution, within a political community, of the health benefits that medical digital solutions may offer. Clinical and research institutions can pursue distributive justice even while monitoring those representations for unintended elements of social bias.

FUNDING

The author(s) reported there is no funding associated with the work featured in this article.

ORCID

Benjamin Gregg http://orcid.org/0000-0001-9510-6147

REFERENCES

Ahmad, M. A., C. Eckert, and A. Teredesai. 2018. Interpretable machine learning in healthcare. In *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 559–560. doi:10.1145/3233547.3233667.

Benedek, M., and A. Fink. 2019. Toward a neurocognitive framework of creative cognition: The role of memory, attention, and cognitive control. *Current Opinion in Behavioral Sciences* 27 (June):116–22. doi:10.1016/j.cobeha.2018.11.002.

Cho, M., and N. Martinez-Martin. 2023. Epistemic rights and responsibilities of digital simulacra for biomedicine. *The American Journal of Bioethics* 23 (9):43–54. doi:10.1080/15265161.2022.2146785.

Dey, S., P. Chakraborty, B. C. Kwon, A. Dhurandhar, M. Ghalwash, F. J. Suarez Saiz, K. Ng, D. Sow, K. R. Varshney, and P. Meyer. 2022. Human-centered explainability for life sciences, healthcare, and medical informatics. *Patterns* 3 (5):100493. doi:10.1016/j.patter.2022.100493.

Ferryman, K. 2020. Addressing health disparities in the Food and Drug Administration's artificial intelligence and machine learning regulatory framework. *Journal of the American Medical Informatics Association* 27 (12): 2016–9. doi:10.1093/jamia/ocaa133.

Gregg, B. 2022a. Political bioethics. *The Journal of Medicine and Philosophy* 47 (4):516–29. doi:10.1093/jmp/jhac008.

Gregg, B. 2022b. *Creating human nature: The political challenge of genetic engineering*. New York: Cambridge University Press. Chapter 6.

He, J., S. L. Baxter, J. Xu, J. Xu, X. Zhou, and K. Zhang. 2019. The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine* 25 (1): 30–6. doi:10.1038/s41591-018-0307-0.

Mittelstadt, B. 2019. Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence* 1 (11):501–7. doi:10.1038/s42256-019-0114-4.

Rasheed, K., A. Qayyum, M. Ghaly, A. Al-Fuqaha, A. Razi, and J. Qadir. 2022. Explainable, trustworthy, and ethical machine learning for healthcare: A survey. *Computers in Biology and Medicine* 149:106043. doi:10.1016/j.compbiomed.2022.106043.

THE AMERICAN JOURNAL OF BIOETHICS
2023, VOL. 23, NO. 9, 86–88
https://doi.org/10.1080/15265161.2023.2237464

Taylor & Francis
Taylor & Francis Group

OPEN PEER COMMENTARIES

Check for updates

Should We be More Worried about Digital Simulacra in Healthcare Being Our "Caricatures", Rather than Our "Replicas"?

Brad Partridge[a]

The University of Queensland

The construction of digital simulacra in healthcare and medical research purportedly strives to virtually recreate some aspect of reality, whether that be a piece of human tissue, an entire organ, a bodily system, a diseased state, or even a whole individual. The terms "digital twin" and "digital replica" have been adopted to imply a data model copy of a patient based on a large number of biomedical, diagnostic, genomic, or behavioral variables. The promise of developing digital simulacra in medicine and research is the ability to conduct "*real-time optimization and testing*" of a person's bodily systems (Fagherazzi 2020, 2), of the sort done with models of engineering and industrial constructions. By focusing on a particular disease state, the goal is to use digital simulacra for improved diagnostic judgements, tailored treatments, and better predictions about the future health of patients (Braun 2021, 394). For advocates, the use of Big Data and AI/machine learning tools is what will allow the realization of true personalized medicine where disease progression can be focused on, and the effects of clinical interventions viewed in the simulated form first and then applied to the individual.

CONTACT Brad Partridge b.partridge@uq.edu.au School of Business, The University of Queensland, Saint Lucia, Queensland, Australia.