

The pursuit of health equity in the era of artificial intelligence

Tania Manríquez Roa^{ab}, Markus Christen^{ab}, Andreas Reis^c, Nikola Biller-Andorno^{ab}

^a Institute of Biomedical Ethics and History of Medicine (IBME), University of Zurich, Switzerland

^b Digital Society Initiative (DSI), University of Zurich, Switzerland

^c Health Ethics & Governance Unit, World Health Organisation, Switzerland

Artificial intelligence (AI) is expected to become a prominent technology in the healthcare domain. What role will it play regarding health equity? More specifically, will AI help societies improve equity, or will it exacerbate or create further inequities? To discuss health equity in the era of AI, we present definitions of key terminology and then address four questions related to health equity focusing on access to AI and on outcomes when using AI. Furthermore, we stress the importance of establishing equity monitoring and integrating sociocultural sensitivity to advance health equity when using AI technologies.

AI is an umbrella term used to refer to a collection of technologies prominently based on machine learning algorithms – rules that use statistical methods to automatically learn and infer patterns from the data [1]. In healthcare, the use of AI technologies has enabled fast and scalable analysis of complex data and has started to impact clinical decision-making. For instance, machine- and deep-learning algorithms are being used to predict clinical outcomes, aiming to optimise the allocation of hospital palliative care resources [2, 3]. AI applications that involve pattern recognition using deep neural networks are currently helping healthcare professionals to interpret medical scans and images in the fields of radiology, neurology, pathology, dermatology, ophthalmology, gastroenterology and cardiology [3]. Also, direct-to-consumer AI technologies embedded in smartwatches and smartphones are used to detect atrial fibrillation, to identify ear infections, migraine headaches and retinal diseases, and to help patients deal with common chronic conditions, such as depression, hypertension and asthma [3].

Health equity is a concept that refers to just distributions of health [4]. Importantly, while *health inequalities* are observable health differences among individuals and groups, *health inequities* are health inequalities that are unjust [5]. The World Health Organization (WHO) defines equity as the absence of avoidable, unnecessary and unfair differences among groups of people, whether these groups are established according to social, economic, demographic or geographic factors [6]. The WHO considers health equity as a core ethical principle to guide decisions around the development and use of AI technologies for health [7]. In the private sector, health equity is framed as a desirable means to optimise quality of life and to promote salutary effects on the larger economy [8].

The distinction between *equitable access* and *equitable outcomes* will help us grasp better the potential role of AI technologies for health equity. Understood in terms of *access*, health equity is commonly framed normatively. Equitable access to healthcare refers to the idea that the use of health services should reflect real needs for care [9]. Understood as an *outcome*, health equity is achieved when everyone can attain their full potential for health and well-being [6], and no one is disadvantaged in reaching this potential because of socially determined circumstances. In our view, the use of AI in healthcare raises four key questions relevant to equity:

1. Can AI technology facilitate equitable access to professional healthcare services?

Yes, but it might also encourage people who currently lack access to professional healthcare services (e.g., the uninsured) to start using AI-based services that are cheap or free but are unsupervised by healthcare professionals. On the one hand, AI can facilitate equitable access to healthcare based on patients' needs, for example, by using machine learning algorithms to accurately match patients with primary care doctors [3]. On the other hand, AI might also discourage the use of professional healthcare services. Take the case of AI apps for mental health. Some of these apps have shown promising results, facilitating access to mental healthcare, disseminating psychoeducation and helping patients deal with symptoms of depression and anxiety [10]. However, other AI apps offer automated mental health “diagnosis” and “treatments”. People in need of healthcare might wrongly believe these are standard forms of healthcare instead of getting diagnosed and treated by a mental health professional.

2. Can we have equitable access to healthcare services supported by AI?

Theoretically yes, but this is not the case yet. In fact, AI might increase existing global inequities in access to healthcare [11]. The digital divide refers to differences both in availability of infrastructure and technologies to interact with digital systems, and in having functional access to them in terms of skills [12]. But let's assume that there is no digital divide: everyone has access to digital infrastructure and sufficient digital skills to use them. Would this en-

Prof. Dr. Nikola Biller-Andorno
Institute of Biomedical Ethics and History of Medicine
WHO Collaborating Center
University of Zurich
Winterthurerstrasse 30
CH-8006 Zürich
biller-andorno[at]ibme.uzh.ch

sure equitable access to AI technologies? Not really, unless we advance in technology sharing, i.e., that AI technologies and especially those that are considered necessary for healthcare are available for everyone.

3. Can AI generate equitable outcomes for all patients when used to support medical decisions?

AI has the potential to do so, but there are cases when the opposite is true. An example of this was the deployment of a machine learning algorithm to identify cancerous skin lesions. The algorithm was trained mainly with data of white patients and therefore did not generate accurate results for black patients. Although AI might amplify and perpetuate biases, it also has potential to make them visible and reduce them. In the example of the cancerous skin lesions, we see that the biased results are not necessarily an intrinsic feature of AI. Rather, biases can originate in the fact that data collection, diagnosis and treatments have focused on the needs of white people. AI can support physicians provide diagnosis and prescribe tailored treatments – provided that the algorithms are designed, trained and tested to work well for all groups of patients. Specifically, AI technologies can help physicians calculate doses of medicines adjusted to the severity of a disease, age and sex of patients. Note that the goal of AI-tailored treatments is to provide the best possible outcome for patients according to their needs. This means that different treatments can bring equitable comparable outcomes for all patients.

4. Can AI generate equitable outcomes for all patients when used to support resource allocation decisions within the healthcare system?

Yes, but not necessarily. An AI system designed to support decision making might comply well with efficiency parameters of resource allocation in healthcare but can at the same time discriminate groups by providing inequitable outcomes (e.g., being unfair to the elderly) [7]. Interestingly, AI technologies might generate inequitable outcomes because of the inner workings of their algorithms, or the datasets deployed for training, but also because of conflicting health goals. For example, a mechanism to allocate scarce resources in the healthcare system can comply with efficiency goals to the detriment of equity goals. In cases like this, health goals come into conflict, regardless of whether AI technologies are deployed or not.

Equity monitoring

We have entered an era of rapid and market-driven introduction of AI. Pursuing health equity in this context involves ensuring that these technologies support health services to meet the needs of all people and contribute to attain everyone's full potential for health. In our view, we need to establish an *equity monitoring* when introducing AI technologies to make sure they do not inadvertently increase or create inequities. Key points for *equity monitoring* are:

- creating AI algorithms that adapt to the specific characteristics and needs of the subgroups in the populations who are going to interact with them;

- training AI algorithms with databases that represent all subgroups in the population;
- clarifying which tasks are best performed by a combination of humans and AI technologies, and which tasks are best achieved by humans only to serve health needs;
- finding mechanisms to make AI technologies available to everyone in cases in which such technologies are the best possible tool to achieve a health goal.

Finally, these points require to be reflected upon considering the local contexts in which AI is used. How to develop AI technologies that are sensitive to sociocultural contexts and that enable us to advance health equity is a question that needs further discussion.

Acknowledgements

The authors thank the [Strategy Lab](#) at the Faculty of Medicine and the Digital Society Initiative (University of Zurich) for sharing their work on digitalisation in medicine.

Tania Manriquez Roa would like to thank the Digital Society Initiative for funding her research. Andreas Reis is a staff member of the World Health Organization. The authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of the World Health Organization.

Potential competing interests

All authors have completed and submitted the International Committee of Medical Journal Editors form for disclosure of potential conflicts of interest. No potential conflict of interest was disclosed.

References

1. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ Glob Health*. 2018 Aug;3(4):e000798. <http://dx.doi.org/10.1136/bmjgh-2018-000798>.
2. Rojas JC, Fahrenbach J, Makhni S, Cook SC, Williams JS, Umscheid CA, et al. Framework for Integrating Equity Into Machine Learning Models: A Case Study. *Chest*. 2022 Jun;161(6):1621–7. <http://dx.doi.org/10.1016/j.chest.2022.02.001>.
3. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019 Jan;25(1):44–56. <http://dx.doi.org/10.1038/s41591-018-0300-7>.
4. Whitehead M. The concepts and principles of equity and health. *Int J Health Serv*. 1992;22(3):429–45. <http://dx.doi.org/10.2190/986L-LHQ6-2VTE-YRRN>.
5. MacKay D, Sreenivasan G. Justice, Inequality, and Health [Internet]. *Stanford Encyclopedia of Philosophy*; 2021. [cited 2023 Jan 11]. Available from: <https://plato.stanford.edu/entries/justice-inequality-health/>
6. World Health Organization [Internet]. *Health Equity* [cited 2023 Jan 11]. Available from: https://www.who.int/health-topics/health-equity#tab=tab_1
7. World Health Organization. *Ethics and governance of artificial intelligence for health: WHO guidance*. Geneva: World Health Organization; 2021.
8. World Economic Forum. [Internet]. White paper [cited 2023 Jan 11]. Available from: https://www3.weforum.org/docs/WEF_A_Blue-print_for_Equity_and_Inclusion_in_Artificial_Intelligence_2022.pdf
9. Aday LA, Begley CE, Lairson DR, Balkrishnan R. *Equity*. 3rd ed. Chicago (IL): Health Administration Press; 2004.
10. Manriquez Roa T, Biller-Andorno N, Trachsel M. The Ethics of Artificial Intelligence in Psychotherapy. In: Trachsel M, Gaab J, Biller-Andorno N, Tekin Ş, Sadler JZ, editors. *The Oxford Handbook of Psychotherapy Ethics*. Oxford: Oxford University Press; 2020. pp. 744–58.
11. Leslie D, Mazumder A, Peppin A, Wolters MK, Haggerty A. Does “AI” stand for augmenting inequality in the era of covid-19 healthcare? *BMJ*. 2021 Mar;372(304):n304. <http://dx.doi.org/10.1136/bmj.n304>.
12. d’Elia A, Gabbay M, Rodgers S, Kierans C, Jones E, Durrani I, et al. Artificial intelligence and health inequities in primary care: a systematic scoping review and framework. *Fam Med Community Health*. 2022 Nov;10 Suppl 1:e001670. <http://dx.doi.org/10.1136/fmch-2022-001670>.