Striving for health equity with machine learning

Health disparities need to be addressed so that the benefits of medical progress are not limited to selected groups. Big data and machine learning approaches are transformative tools for public and population health, but need ongoing support from insights in algorithmic fairness.

The fast progress of medical research in the past century has enabled the diagnosis and treatment of more complex diseases on an increasingly granular level, making it possible to tailor cures to individual patients. However, a wide range of decisions and factors outside of the clinical context influence our health. Public policies that promote healthy lifestyles and diets, that tax or ban unhealthy compounds in food and that efficiently allocate public funds to health resources can substantially reduce non-communicable diseases such as cancer, heart diseases, chronic respiratory diseases and diabetes, which currently kill 41 million people worldwide each year. Various social, cultural, economic and environmental factors are at play in such policy decisions and public health data increasingly reveal disparities in health outcomes between different population groups; certain communities are more at risk than others, as grimly demonstrated by the COVID-19 pandemic which has disproportionately affected ethnic and minority groups.

A Perspective in this issue by Rumi Chunara and colleagues surveys the challenges in using machine learning tools to improve health equity in public and population health. As the influences on health are complex, there is hope that increased collection of data and sophisticated analysis methods can allow us to understand these relations. Machine learning efforts could also help with better design of interventions, improved prediction of patient outcomes and better allocation of resources, paying special attention to disadvantaged communities. But a first challenge is to identify the factors that contribute to health outcomes. The use of proxy variables, such as race and ethnicity, is in places necessary to assess where bias happens, but the authors call for caution in using and interpreting these variables, as

they mask effects of the underlying, more complex relations. For example, a study that analysed blood pressure within African Americans and white people showed that racial disparities regarding hypertension may be better explained by differences in the level of education than by genetic variation.

Challenges around the identification and mitigation of racial, socioeconomic, gender or other disparities are central to the new subfield of machine learning studying 'algorithmic fairness'. Bias appears at various stages of data and algorithm use, from the initial decision to use predictive software for a specific problem and what variables to predict, to data collection and algorithmic design. Whether a method is fair, as in whether humans would agree that the decisions made are fair, often becomes apparent only after the fact when patterns of discrimination reappear.

A recent Correspondence in our June issue, by Supriya Kapur, highlighted the issue of algorithmic fairness in the context of clinical applications in which machine learning models are developed to emulate assessments of experts with automated methods, for faster and more accessible diagnoses. But it is well documented that clinical experts themselves, like most of us, are subject to biases, unconscious or otherwise. Kapur argues that regulatory policies have to change, as the focus on replicating existing clinical data currently encourages mimicking rather than questioning discriminatory mechanisms.

Chunara et al. describe in their Perspective two types of public health data: Firstly, data gathered in surveys conducted by public health organizations, which are generally well established and regulated, and secondly, a new type of person-generated data such as via smartphones and social media platforms. There is a clear potential for studying complex population health issues by analysing the latter, but

much of our digital lives are obtained and commercialized by big technology companies and there is little control each one of us has over our own data. Jathan Sadowski and colleagues argue for a more democratic approach in a recent *Nature Comment* to enable more ethical, equitable and scientifically sound use of data, including for social science applications. They call for the creation of public data trusts to gather person-generated data, with privacy protection and safeguards against use in which the participants or users did not agree. With more public scrutiny of datasets and of machine learning models trained on them, questions can be asked about assumptions and biases of algorithmic decisions. For instance, the authors mention research from 2019 that identified significant racial bias in the training data for a proprietary algorithm for medical resource allocation, which led to a reduction of extra care for black patients compared with white patients with the same level of health, as healthcare spending was erroneously assumed to equate to healthcare needs. Challenges remain but a democratization of data ownership could empower efforts for a more ethical approach and equity in health research with machine learning.

It is vital that health disparities are addressed so that the benefits of medical progress are not limited to selected groups. This requires more inclusion in research, but also a constant revision of categories that are used to control for and ensure fairness. Although increased collection of data and analysis can help in this effort, data are never objective and people who generate the data need to have a say in how their data are used in algorithm development. □

Published online: 17 August 2021
<https://doi.org/10.1038/s42256-021-00385-0>