



EQUALITY

Mind the gap

Artificial intelligence has great potential to improve health care, but without careful implementation, only a privileged portion of society will benefit.

BY LINDA NORDLING

When data scientists in Chicago, Illinois, set out to test whether a machine-learning algorithm could predict how long people would stay in hospital, they thought that they were doing everyone a favour. Keeping people in hospital is expensive, and if managers knew which patients were most likely to be eligible for discharge, they could move them to the top of doctors' priority lists to avoid unnecessary delays. It would be a win-win situation: the hospital would save money and people could leave as soon as possible.

Starting their work at the end of 2017, the scientists trained their algorithm on patient data from the University of Chicago academic

hospital system. Taking data from the previous three years, they crunched the numbers to see what combination of factors best predicted length of stay. At first they only looked at clinical data. But when they expanded their analysis to other patient information, they discovered that one of the best predictors for length of stay was the person's postal code. This was puzzling. What did the duration of a person's stay in hospital have to do with where they lived?

As the researchers dug deeper, they became increasingly concerned. The postal codes that correlated to longer hospital stays were in poor and predominantly African American neighbourhoods. People from these areas stayed in hospitals longer than did those from more affluent, predominantly white areas. The reason for this disparity evaded the team. Perhaps

people from the poorer areas were admitted with more severe conditions. Or perhaps they were less likely to be prescribed the drugs they needed.

The finding threw up an ethical conundrum. If optimizing hospital resources was the sole aim of their programme, people's postal codes would clearly be a powerful predictor for length of hospital stay. But using them would, in practice, divert hospital resources away from poor, black people towards wealthy white people, exacerbating existing biases in the system.

"The initial goal was efficiency, which in isolation is a worthy goal," says Marshall Chin, who studies health-care ethics at University of Chicago Medicine and was one of the scientists who worked on the project. But fairness is also important, he says, and this was not explicitly



Anthropologist Kadija Ferryman studies the cultural and ethical impacts of using AI in health care.

considered in the algorithm's design.

This story from Chicago serves as a timely warning as medical researchers turn to artificial intelligence (AI) to improve health care. AI tools could bring great benefits to people who aren't currently served well by the medical system. For example, an AI tool for screening chest X-rays for signs of tuberculosis, developed by start-up Zebra Medical Vision in Shefayim, Israel, is being rolled out in hospitals in India to speed up diagnosis of people with the disease. Machine-learning algorithms could also help scientists to tease out which people are likely to respond best to which treatments, ushering in an era of tailor-made medicine that might improve outcomes.

But this revolution hinges on the data that are available for these tools to learn from, and those data mirror the unequal health system we see today. "In some health-care systems, there are very basic things that are being ignored, basic quality of care that people are not receiving," says Kadija Ferryman, an anthropologist at the New York University Tandon School of Engineering who studies the social, cultural and ethical impacts of the use of AI in health care. These inequalities are preserved in the terabytes of health data being generated around the world. And these data have primed the health-care industry for the kind of disruption that is being driven by ride-sharing platforms in the transport sector and home-rental platforms such as Airbnb in the hotel industry, Ferryman says. "Apple, Google, Amazon — they are all making inroads into the health-care space." But because AI algorithms learn from existing data, there is a risk, Ferryman says, that the tools that result from this gold rush could entrench or deepen

inequalities — such as the fact that black people in US emergency rooms are 40% less likely to receive pain medication than are white people¹.

The Chicago story is an example of bias being documented in a system before it is implemented. But not all occurrences are caught. In January, at the Conference on Fairness, Accountability and Transparency in Atlanta, Georgia, scientists from the University of California, Berkeley, and the University of Chicago presented evidence of "significant racial bias" in an algorithm that determines health-care decisions for more than 70 million people in the United States².

The algorithm in question allocates 'risk scores', which are used to enrol people at high risk of future complex health needs into specially resourced care programmes. The researchers found that black people had significantly more chronic illnesses than did white people with the same risk scores. This means that white people are more likely to be enrolled in targeted programmes than are black people with the same level of health. If the algorithm scored black and white people equally, the researchers said, black people would be enrolled into the programmes at more than twice the current rate.

RUBBISH IN, RUBBISH OUT

Impaired access to care for certain people is just one way in which AI tools could widen the health gap globally. Another problem is making sure that AI-powered tools can be applied equally to different groups of people. Information from certain population groups tends to be missing from the data with which these tools learn, meaning that the tool might work less well

for members of those communities.

White, adult men are strongly over-represented in existing medical data sets, at the expense of data from white women and children and people of all ages from other ethnic groups. This lack of diversity in the data is likely to result in biased algorithms³.

There are some efforts to plug these gaps. In 2015, the US National Institutes of Health (NIH) created the All of Us initiative with US\$130 million in funding. The research programme aims to form a database of genetic and health data from one million volunteers, expanding the data sets available for guiding the development of precision medicine to provide better quality care for everyone in the United States. It specifically targets previously under-represented communities for data collection. As of July, more than 50% of the fully enrolled participants in the programme were from minority ethnic groups.

But even such diverse data sets might not translate to AI tools that can be rolled out reliably in low-income countries, where disease profiles often differ from those in high-income nations. In sub-Saharan Africa, women are diagnosed with breast cancer younger, on average, than are their peers in developed countries, and their disease is more advanced at diagnosis⁴. Diagnostic AI tools trained on mammograms from Europe are primed to identify disease in its early stages in older women, therefore might not travel well, says Kuben Naidu, president of the Radiological Society of South Africa in Cape Town.

The obvious way to solve this problem is to give AI developers access to data from low-income countries. But doing so raises concerns related to data protection for vulnerable populations, says Naidu. Medical data is highly sensitive — information such as HIV status could be used to discriminate against certain populations, for instance. Naidu recalls feeling troubled by the eagerness he was met with when visiting a gathering of radiologists in the United States a few years ago. AI companies among the exhibitors "were very excited to hear that I was from Africa, and asked how they could get access to our data", he says.

Companies offer to pay for such data, he says, which might tempt cash-strapped national health systems or individual researchers to part with patient data, perhaps without thinking hard about the rights of the those whose data they are sharing. A number of developing countries are introducing data-protection laws, but in those countries where law enforcement is lax, such regulations could be circumvented.

Of course, privacy isn't just a worry in developing countries. Even in nations with strong data-protection legislation, such as the United States, keeping personal data private might be harder than expected. The University of Chicago is currently facing a class-action lawsuit for sharing patient records with Google. The project stripped out identifiers such as social security numbers and names from the data, in

accordance with the country's privacy laws. But the plaintiffs in the lawsuit argue that the dates of patient visits, which were not excised from the data, could be combined with other information held by Google, such as smartphone locations, to match people to their health records.

A related concern is that data companies could tempt people to give up their privacy in return for medical care or financial reward. Such practices could create a privacy divide between rich and poor along the same lines as the digital divide that already separates different socio-economic groups⁵.

Ferryman, who sits on the institutional review board of the All of Us programme, admits that she struggles with the tension between the push — no matter how benevolent — to gather data from historically marginalized and maligned populations, and the need to protect those very populations from being exploited. “On the one hand, we want to help these people by gathering more information about them. But on the other hand, what’s to say that data will not be used in ways that discriminate against them?”

PROMOTING FAIRNESS THROUGH AI

One way to ensure that AI tools don't worsen health inequalities is to incorporate equity into the design of AI tools. The University of Chicago Medicine data team did this after discovering the issues with its proposed hospital-discharge optimization algorithm. Its data science unit now partners closely with the university's diversity, inclusion and equity department. This means addressing equity in AI is not an afterthought “but rather, a core of how we implement AI in our health system,” says John Fahrenbach, a data scientist with the university's Center for Healthcare Delivery Science and Innovation.

Fahrenbach worries that not enough attention is being paid to equity in the design of most machine-learning models. “There are so many machine-learning models in health care being developed, deployed and pitched, and I rarely hear them even mention these concerns. This really has to change and formalized regulation is likely the best way for this to happen,” he says.

There is some way to go in this respect. The UK's National Health Service has received criticism⁶ for not giving enough attention to the potential for AI to widen health gaps in its updated Code of Conduct for Data-driven Health and Care Technologies, released in February. Similarly, the US Food and Drug Administration (FDA), which regulates and approves new medical technologies, has been urged by the American Medical Association to highlight bias as a significant risk of machine-learning in its approval process for medical software. A modification to the process, proposed in April, would allow AI tools that continually improve their performance by learning from new data to do so without having to undergo another review by the FDA.

Some research funders are tackling the issue



A member of the US National Institutes of Health All of Us team collects samples from a participant.

head on, by launching research programmes to study how the introduction of AI tools affects access to care and its quality. Wellcome, a London-based biomedical charity, launched such a programme in June this year. The £75-million (US\$90-million), five-year programme will look at ways to make sure that innovations in the use of health data will benefit everyone — not just in the United Kingdom, but also in other parts of the world, such as East and Southern Africa and India, where Wellcome has a strong presence.

Determining whether patchy or biased data could be resulting in unequal health care will play a part in the programme, says Nicola Perrin, head of data for science and health at Wellcome, but it won't be the primary focus. The initiative will drill down into how

the unique make-up of individual hospitals — such as the availability of doctors, medicine or equipment, and the hospital's relationship with the communities that depend on it for care — affects whether the AI tools deliver.

“That's the bit that's always neglected, the unglamorous, unsexy part,” she says. In developing countries, especially, it's about making sure that tools are actually meeting demand on the ground, and building trust and buy-in from the communities they are intended to help, she says. “We need to understand people's expectations, and where the boundaries should be.”

This point echoes that of Naidu. Rolling out health-care tools in developing countries is

never easy, he says, and it requires an intimate understanding of the existing bottlenecks in the health system. For example, the AI that can identify people with tuberculosis from chest X-rays, primed for use in India, could also save time, money and lives in South Africa — especially in rural areas where there aren't specialists to examine such images, he says. But to obtain images in the first place, communities will need X-ray machines and people to operate them. Failure to provide those resources will mean that AI tools will simply serve those already living near better-resourced clinics.

Ferryman thinks it is right to be cautious about new medical technologies. “There is no absolute guarantee that the tools will have benefits that outweigh the potential harm they can do,” she says. But she also thinks that most people working in health care in the United States want a more equitable system. Health systems are built around highly trained specialists whose primary motivation is caring for people, and many doctors are hungry for innovations that make the system fairer, she says. “That gives me hope.” ■

Linda Nordling is a science journalist in Cape Town, South Africa.

1. Lee, P. et al. *Am. J. Emer. Med.* <https://doi.org/10.1016/j.ajem.2019.06.014> (2019).
2. Obermeyer, Z. & Mullainathan, S. in *Proc. Conf. Fairness, Accountability, and Transparency* 89 (ACM, 2019).
3. Blasimme, A. & Vayena, E. in *Oxford Handbook of Ethics of Artificial Intelligence* Preprint at <https://doi.org/10.2139/ssrn.3368756> (2019).
4. Blac, E. & Richmond, R. *Glob. Health* **15**, 3 (2019).
5. Montgomery, K., Chester, J. & Kopp, K. J. *Info. Policy* **8**, 34–77 (2018).
6. Smallman, M. *Nature* **567**, 7 (2019).