



Black people were less likely than white people to be sent for personalized care, a study found.

MILLIONS AFFECTED BY RACIAL BIAS IN HEALTH-CARE ALGORITHM

Study reveals widespread racism in decision-making software used by US hospitals.

By Heidi Ledford

An algorithm widely used in US hospitals to allocate health care to patients has been systematically discriminating against black people, a sweeping analysis has found.

The study, published in *Science* on 24 October, concluded that the algorithm was less likely to refer black people than white people who were equally sick to programmes that aim to improve care for patients with complex medical needs (Z. Obermeyer *et al. Science* 366, 447–453; 2019). Hospitals and insurers use the algorithm and others like it to help to manage care for about 200 million people in the United States each year.

This type of study is rare, because researchers often cannot gain access to proprietary algorithms and the reams of sensitive health data needed to fully test them, says Milena Gianfrancesco, an epidemiologist at the University of California, San Francisco, who has studied sources of bias in electronic medical records. But smaller studies and anecdotal reports have documented unfair and biased decision-making by algorithms used in everything from criminal justice to education and health care.

“It is alarming,” says Gianfrancesco of

the latest study. “At the same time, it’s not surprising.”

Ziad Obermeyer, who studies machine learning and health-care management at the University of California, Berkeley, and his team stumbled across the problem while examining the impact of programmes that provide additional resources and closer medical supervision for people with multiple, sometimes overlapping, health problems.

When Obermeyer and his colleagues ran routine statistical checks on data they received from a large hospital, they were surprised to find that people who self-identified as black were generally assigned lower risk scores than equally sick white people. As a result, the black people were less likely to be referred to the programmes that provide more-personalized care.

The researchers found that the algorithm assigned risk scores to patients on the basis of total health-care costs accrued in one year. They say that this assumption might have seemed reasonable because higher health-care costs are generally associated with greater health needs. The average black person in the data set that the scientists used had similar health-care costs to the average white person.

But a closer look at the data revealed that the average black person was also substantially sicker than the average white person, with

a greater prevalence of conditions such as diabetes, anaemia, kidney failure and high blood pressure. Taken together, the data showed that the care provided to black people cost an average of US\$1,800 less per year than the care given to white people with the same number of chronic health problems.

The scientists speculate that this reduced access to care is due to the effects of systemic racism, ranging from distrust of the health-care system to direct racial discrimination by health-care providers.

And because the algorithm assigned people to high-risk categories on the basis of costs, those biases were passed on in its results: black people had to be sicker than white people before being referred for additional help. Only 17.7% of patients that the algorithm assigned to receive extra care were black. The researchers calculate that the proportion would have been 46.5% if the algorithm was unbiased.

When Obermeyer and his team reported their findings to the algorithm’s developers – Optum of Eden Prairie, Minnesota – the company repeated their analysis and got the same results. Obermeyer is working with the firm without salary to improve the algorithm.

He and his team collaborated with the company to find variables other than health-care costs that could be used to calculate a person’s medical needs, and repeated their analysis after tweaking the algorithm accordingly. They found that making these changes reduced bias by 84%.

“We appreciate the researchers’ work,” Optum said in a statement. But the company added that it considered the study’s conclusion to be “misleading”. “The cost model is just one of many data elements intended to be used to select patients for clinical engagement programs.”

Obermeyer says that using cost prediction to make decisions about patient engagement is a pervasive issue. “This is not a problem with one algorithm, or one company – it’s a problem with how our entire system approaches this problem,” he says.

Examining assumptions

Correcting bias in algorithms is not straightforward, Obermeyer adds. “Those solutions are easy in a software-engineering sense: you just rerun the algorithm with another variable,” he says. “But the hard part is: what is that other variable? How do you work around the bias and injustice that is inherent in that society?”

This is in part because of a lack of diversity among algorithm designers, and a lack of training about the social and historical context of their work, says Ruha Benjamin, author of *Race After Technology* (2019) and a sociologist at Princeton University in New Jersey.

“We can’t rely on the people who currently design these systems to fully anticipate or mitigate all the harms associated with

ED KASHI/VI/REDUX/EYEVINE

automation,” she says.

Developers should routinely run tests such as those performed by Obermeyer’s group before they deploy an algorithm that affects human lives, says Rayid Ghani, a computer scientist at Carnegie Mellon University in Pittsburgh, Pennsylvania. That kind of auditing is more common now, he says, since reports of biased algorithms have increased.

He thinks that the results of these audits should always be compared to human

decision-making. Unpublished analyses by Ghani’s team have compared algorithms used in public health, criminal justice and education to human decision-making, and found that the machine-learning systems were biased – but less so than the people.

“We are still using these algorithms called humans that are really biased,” says Ghani. “We’ve tested them and known that they’re horrible, but we still use them to make really important decisions every day.”

gradient of same-sex attraction. The app cited the *Science* study and warned users that it did not predict same-sex attraction.

The researchers behind the *Science* study say that Bellenson’s app misrepresents their work. The test “is not grounded in science. It is not predictive. It won’t tell you anything”, says Benjamin Neale, a geneticist at the Broad Institute and an author of the *Science* analysis. He and his colleagues examined the DNA of around 475,000 people and found 5 genetic variations loosely correlated with people who said they’d had sex with someone of the same sex at least once. But none of the variants was so prevalent that the researchers could use them to predict a person’s sexual identity.

Neale sent a letter to GenePlaza on 14 October asking that it take down the app – or remove references to his study. The next week, Bellenson renamed the app ‘122 Shades of Gray’ and added a note explaining that the authors of the *Science* study weren’t affiliated with the project. He says that because the app has always warned users that it is not predictive, it does not misrepresent the study.

But the chorus of angry scientists on Twitter grew louder. Some echoed Vitti’s concern that the app could be abused. In his petition, Vitti noted that Bellenson lives in Uganda, where gay sex is punishable by life in prison. Vitti worried that, regardless of the app’s scientific flaws, Ugandan authorities could get hold of a person’s results and use them as evidence of sexual preferences.

Bellenson says that there are much simpler ways of discovering a person’s sexual preference, such as looking at their social-media accounts. “The idea that a government would need a DNA test to figure out if someone is gay is ridiculous,” he says.

‘GAY GENE’ APP PROVOKES FEARS OF A GENETIC WILD WEST

Debate highlights broader concerns about apps that use the results of direct-to-consumer genetic testing.

By Amy Maxmen

Joseph Vitti’s stomach turned when he opened a link an acquaintance had sent him. It took him to an app called ‘How Gay Are You?’ that purported to gauge a person’s level of attraction to others of the same sex, according to their genes.

The app’s creator, Joel Bellenson, a US entrepreneur living in Kampala, Uganda, based the test on the findings of a massive study on the genetics of same-sex sexual behaviour – even though the analysis, published in *Science* in August, concluded that a person’s genes cannot predict their sexuality (A. Ganna *et al. Science* 365, eaat7693; 2019).

Vitti, a computational geneticist at the Broad Institute in Cambridge, Massachusetts, thinks the app was misleading – even dangerous. “There are vulnerable queer people all over the world,” says Vitti, “and this app stands to hurt them.” On 11 October, he started an online petition to remove the test. Within two weeks, more than 1,660 people had signed it.

Bellenson says that the idea his test could endanger people is an “absurd scenario” and notes that the test also included a warning that it could not predict same-sex attraction.

But the furore over his app highlights a growing problem in the field of genetics. Researchers conduct statistically sophisticated analyses of hundreds of thousands of genomes, searching for associations between genetic variations and diseases, behaviours or other characteristics. Anyone can take the variations identified by such studies, strip them of caveats and nuance, and market a simple genetic-interpretation tool online.

Scientists and genetic counsellors say that

these unregulated tools can harm individuals and society, causing anxiety, unnecessary medical expenses, stigmatization and worse. “It’s the Wild West of genetics,” says Erin Demo, a genetic counsellor at Sibley Heart Center Cardiology in Atlanta, Georgia. “This is just going to get harder and harder.”

Bellenson posted his app on GenePlaza, an online marketplace for DNA-interpretation tools, in early October. For US\$5.50, a person could upload their genetic data – as supplied by consumer DNA sequencing companies such as 23andMe of Mountain View, California – and the app would place them along a



Millions of people have had their DNA sequenced by consumer genetic-testing companies.