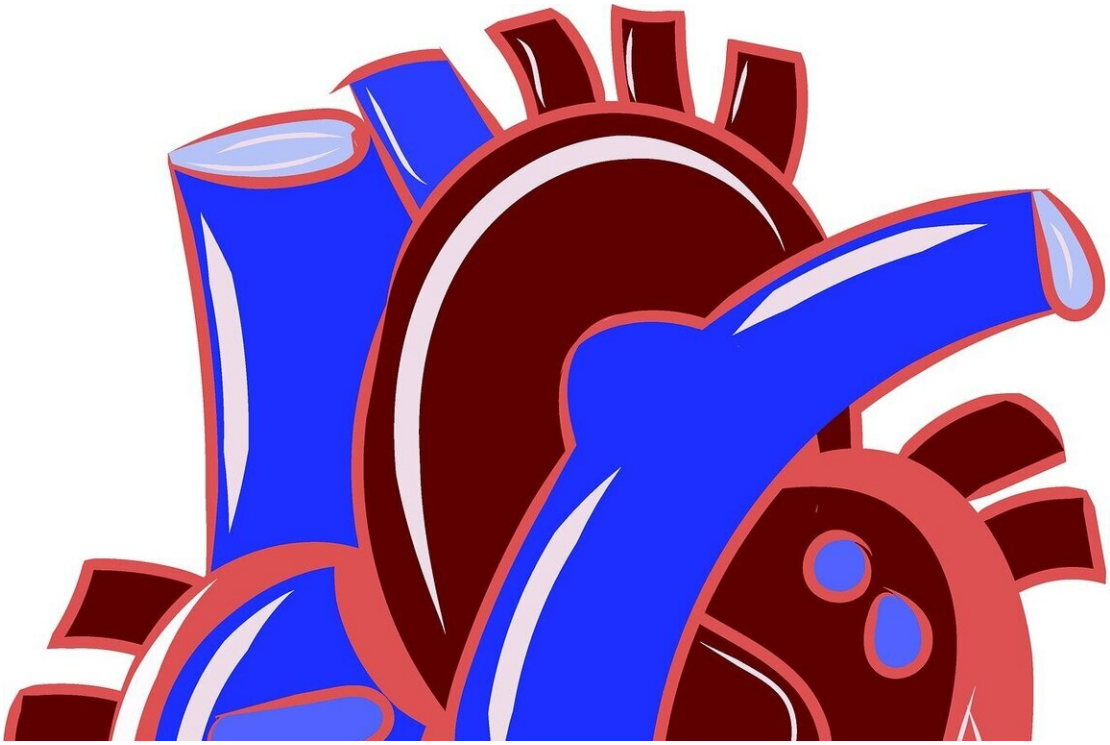
Machine learning for cardiovascular disease improves when social, environmental factors are included

July 27 2021



Credit: CC0 Public Domain

Machine learning can accurately predict cardiovascular disease and guide treatment—but models that incorporate social determinants of health better capture risk and outcomes for diverse groups, finds a new

study by researchers at New York University's School of Global Public Health and Tandon School of Engineering. The article, published in the *American Journal of Preventive Medicine*, also points to opportunities to improve how social and environmental variables are factored into machine learning algorithms.

Cardiovascular disease is responsible for nearly a third of all deaths worldwide and disproportionately affects lower socioeconomic groups. Increases in cardiovascular disease and deaths are attributed, in part, to social and environmental conditions—also known as social determinants of health—that influence diet and exercise.

"Cardiovascular disease is increasing, particularly in low- and middle-income countries and among communities of color in places like the United States," said Rumi Chunara, associate professor of biostatistics at NYU School of Global Public Health and of computer science and engineering at NYU Tandon School of Engineering, as well as the study's senior author. "Because these changes are happening over such a short period of time, it is well known that our changing social and environmental factors, such as increased processed foods, are driving this change, as opposed to genetic factors which would change over much longer time scales."

Machine learning—a type of artificial intelligence used to detect patterns in data—is being rapidly developed in cardiovascular research and care to predict disease risk, incidence, and outcomes. Already, statistical methods are central in assessing cardiovascular disease risk and U.S. prevention guidelines. Developing predictive models gives health professionals actionable information by quantifying a patient's risk and guiding the prescription of drugs or other preventive measures.

Cardiovascular disease risk is typically computed using clinical information, such as blood pressure and cholesterol levels, but rarely

take social determinants, such as neighborhood-level factors, into account. Chunara and her colleagues sought to better understand how social and environmental factors are beginning to be integrated into machine learning algorithms for cardiovascular disease—what factors are considered, how they are being analyzed, and what methods improve these models.

"Social and environmental factors have complex, non-linear interactions with cardiovascular disease," said Chunara. "Machine learning can be particularly useful in capturing these intricate relationships."

The researchers analyzed existing research on machine learning and cardiovascular disease risk, screening more than 1,600 articles and ultimately focusing on 48 peer-reviewed studies published in journals between 1995 and 2020.

They found that including social determinants of health in machine learning models improved the ability to predict cardiovascular outcomes like rehospitalization, heart failure, and stroke. However, these models did not typically include the full list of community-level or environmental variables that are important in cardiovascular disease risk. Some studies did include additional factors such as income, marital status, social isolation, pollution, and health insurance, but only five studies considered environmental factors such as the walkability of a community and the availability of resources like grocery stores.

The researchers also noted the lack of geographic diversity in the studies, as the majority used data from the United States, countries in Europe, and China, neglecting many parts of the world experiencing increases in cardiovascular disease.

"If you only do research in places like the United States or Europe, you'll miss how social determinants and other environmental factors related to

cardiovascular risk interact in different settings and the knowledge generated will be limited," said Chunara.

"Our study shows that there is room to more systematically and comprehensively incorporate social determinants of health into cardiovascular disease statistical risk prediction models," said Stephanie Cook, assistant professor of biostatistics at NYU School of Global Public Health and a study author. "In recent years, there has been a growing emphasis on capturing data on social determinants of health—such as employment, education, food, and social support—in electronic health records, which creates an opportunity to use these variables in machine learning studies and further improve the performance of risk prediction, particularly for vulnerable groups."

"Including social determinants of health in machine learning models can help us to disentangle where disparities are rooted and bring attention to where in the risk structure we should intervene," added Chunara. "For example, it can improve clinical practice by helping health professionals identify patients in need of referral to community resources like housing services and broadly reinforces the intricate synergy between the health of individuals and our environmental resources."

More information: Yuan Zhao et al, Social Determinants in Machine Learning Cardiovascular Disease Prediction Models: A Systematic Review, *American Journal of Preventive Medicine* (2021). DOI: 10.1016/j.amepre.2021.04.016

Provided by New York University

https://medicalxpress.com/news/2021-07-machine-cardiovascular-disease-social-environmental.html