Removing bias from health care AI tools

February 22 2024

Credit: Pixabay/CC0 Public Domain

Rapid advances in artificial intelligence (AI) have opened the way for the creation of a huge range of new health care tools, but to ensure that these tools do not exacerbate preexisting health inequities, researchers urge the use of more representative data in their development.

Researchers from Oxford University's Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), University College London and the Center for Ethnic Health Research, supported by Health Data Research UK, have for the first time studied the full detail of ethnicity data in the NHS. They outline the importance of using representative data in health care provision and have compiled this information into a research-ready database.

The new study, published in *Scientific Data*, is the first part of a three-phase project that aims to reduce bias in AI health prediction models which are trained on real-world patient data. The project, which addresses ethnicity disparities that were highlighted during the pandemic, is part of the UK Government's COVID-19 Data and Connectivity National Core Study led by Health Data Research UK.

The researchers used de-identified data on ethnicity and other characteristics from general practice and hospital health records, accessed safely within NHS England's Secure Data Environment (SDE) service, via the British Heart Foundation Data Science Center's CVD-COVID-UK/COVID-IMPACT Consortium.

This is the first time that patient ethnicity data has been studied at this depth and breadth for the whole population of England. The researchers were able to combine records to analyze patient self-identified ethnicity recorded through over 489 potential codes.

Researchers analyzed how more than 61 million people in England identified their ethnicity in over 250 different groups. They also looked at the characteristics of those with no record of their ethnicity, and how conflicts in patient ethnicity data can arise. The data, now available for other researchers to use, shows that 1/10 patients lack ethnicity records, and around 12% of patients had conflicting ethnicity codes in their patient records.

Sara Khalid, Associate Professor of Health Informatics and Biomedical Data Science at NDORMS, explained, "Health inequity was highlighted during the COVID19 pandemic, where individuals from ethnically diverse backgrounds were disproportionately affected, but the issue is long-standing and multi-faceted.

"Because AI-based health care technology depends on the data that is fed into it, a lack of representative data can lead to biased models that ultimately produce incorrect health assessments. Better data from real-world settings, such as the data we have collected, can lead to better technology and ultimately better health for all."

Professor Cathie Sudlow, Chief Scientist at Health Data Research UK and Director of its BHF Data Science Center said, "We are delighted to be supporting hundreds of researchers to harness the power of the UK's rich health data. This study on ethnicity recording highlights how different sources of health data from the whole English population can be accessed and analyzed in a safe and secure way, providing insights that are relevant to everyone.

"The findings will empower health professionals, patients, carers and policymakers to make better decisions that will benefit people of all ages, ethnic groups, and social backgrounds across the country."

The study assessed the available detail of ethnicity data in NHS England, including across different types of ethnicity codes. For example, NHS hospitals record patient data via 19 ethnicity codes, while GPs use the globally recognized SNOMED-CT Codes, of which there are 489. However, health researchers lose the finer detail from these recording systems as they typically collapse these groups into just five or six, potentially leading to less accurate research.

The researchers plan to demonstrate the value of these findings in the

subsequent phases of the project, which will first focus on using these detailed results on ethnicity data to better describe how different ethnicities were impacted by the COVID-19 pandemic, and then feed into more equitable artificial intelligence and machine learning tools suitable for use by diverse patient groups.

More information: Marta Pineda-Moncusí et al, Ethnicity data resource in population-wide health records: completeness, coverage and granularity of diversity, *Scientific Data* (2024). DOI: 10.1038/s41597-024-02958-1

Provided by University of Oxford