The use of Big Data (large, aggregated datasets) to inform the provision of health care leaves out context and details. Shutterstock

Small Data approaches provide nuance and context to health datasets

September 11, 2019 4.21pm EDT

The trend of tracking health and well-being using digital technologies has permeated mainstream culture. The real-time monitoring capabilities, interactive decision-support algorithms and diagnostic testing features of digital health devices have drawn the interest of users everywhere, including the Global South.

Applying tools such as predictive analytics and prescriptive analytics has benefited businesses such as insurance companies and health-care providers. It has also led to unequal treatment and discrimination of individuals as consumers and recipients of health-care services, and to ill-advised decision making by clinicians and policy makers.

Our research, which was undertaken with participants largely from North America, has investigated attitudes towards the sharing of personal health data with various stakeholders within the wider health sector. It also explores alternative data approaches which could mitigate marginalization and exclusion.

Authors

Debora Irene Christine
Researcher, United Nations University

Mamello Thinyane
Principal Research Fellow, United Nations University institute on Computing and Society, United Nations University
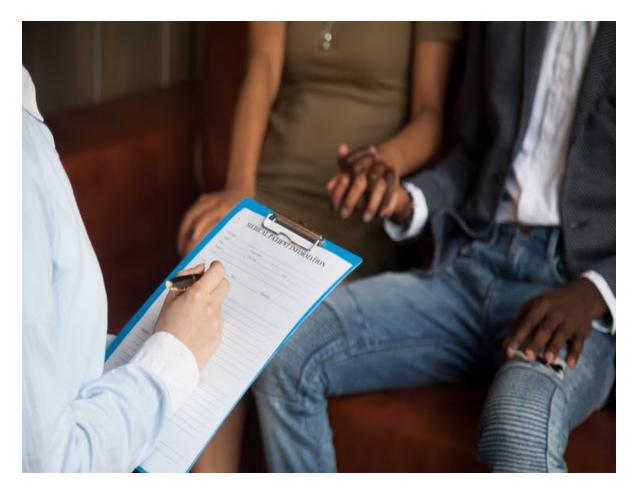
Big data and discrimination

Personal health data collected through digital devices like glucose monitors and personal health trackers (such as Fitbit) and mobile devices are naturally small data. When this data become aggregated into big data sets, they also become highly valuable and profitable as knowledge resources.

Discrimination based on health data collected from self-tracking apps is rampant. Already, health information is being combined with other types of personal data to make population-wide inferences and correlations that carry value on the market of patient health data.

How does this work? Imagine yours — and other people's — personal health data pooled together by data brokers and processed using machine learning algorithms to identify patterns and conditions that affect the overall health across groups. These patterns are then used to predict health risks and health-care costs. Combined with the data of purchasing habits, personal health data are used by insurance companies to offer customers personalized, data-driven, dynamic prices. Processed using predictive analytics, the same data could be used as the basis for health prediction discrimination.

These unfair practices further discriminate against people who are not viewed as profitable. Data-driven profiling on the basis of discriminatory attributes related to group membership such as race, gender or sexual orientation has the potential to perpetuate historical data marginalization and existing disparities by excluding vulnerable groups from health care.

Sharing personal health data

Our research explored individuals' attitudes towards sharing of their personal health and well-being data with stakeholders within the health ecosystem. Three distinct groups of stakeholders were identified. Participants were most willing to share their personal health and well-being data with their doctors who directly provide them with pertinent services. This was followed by their families and friends, who shared high social proximity with them. The latter also reflects the motivation of sharing information for social sense-making and support.

The participants show least willingness to share data with entities within the wider health sector, such as pharmaceutical companies, the national statistics offices and multilateral health organizations such as the World Health Organization. Sharing personal health data could inform the monitoring of specific health indicators and contribute to reporting on national health and well-being. Subsequently, it could lead to the development of health interventions and policies.

We found that concerns about personal data privacy, data ownership and personal benefit affected participants' willingness to share their personal data.

Human-centered approaches

The potential harms of the increased application of big data analytics tools and to personal health data need to be addressed. On the other hand, there is an increasing need for greater participation of individuals as data producers and users to better understand social phenomena related to health and disease. The use of data collected from individuals, however, needs to be grounded in the understanding of individuals' preferences, human rights principles and ethical standards.

Preserving individual privacy and providing protection from potential discrimination based on sensitive health data requires putting fair, accountable and transparent algorithms in place. It also requires regulations which limit data use that might cause harm to certain individuals or groups (e.g. the use of health data to increase premiums in insurance or to deny access from specific health services).

Apps that track an individual's personal health data need to be transparent in the ways they gather data, what they do with it and who they provide it to. Guaranteeing privacy preservation will help to foster trust among users. Users should be informed of when, how and why their data is being used, as well as of the risks associated with the external use of the data, such as data breaches. Users should also retain control over their data, with the option to opt out and to request for their data to be deleted.

---

*Read more: **My Health Record: Deleting personal information from databases is harder than it sounds***

Small versus big data

Alternative data approaches, including the use of Small Data could mitigate the limitations of data approaches which rely heavily on Big Data analytics. As an approach to data processing, Small Data centers the individual in collecting, analyzing and applying personal data.

Using Small Data approaches means the sociocultural context from which data is collected is considered, enabling a detailed understanding of causal relations of health and well-being problems.

Personal health data can be used to address health inequalities or disparities in quality of life. For example, personal well-being data, when pooled, can demonstrate how physiological stress is tied to

societal norms and pressures rather than to individual weaknesses. A Small Data approach provides ways for the collection, analysis and application of personal health data which works towards giving individuals more meaning and insights of their data "through looking closely at others who are like us."

Small data policies

Governments can encourage people's greater contribution and participation in informing the reporting on national and global health and well-being. Developing small data tools enables people to collect their own health data and have full control of their participation in the wider health data ecosystem.

[*Expertise in your inbox. Sign up for The Conversation's newsletter and get a digest of academic takes on today's news, every day.*]

Privacy   Health policy   Health care   Big data   Health data

Before you go...

Journalists and researchers have one thing in common: we seek the truth. I work with academics and scientists every day to communicate knowledge, discovery, and facts to readers like you who also care about evidence over opinion. If you think this is important – and I'm confident you do – please help The Conversation grow with a gift in whatever amount you can afford. Thank you.

Donate now

Martin LaMonica
Deputy Editor