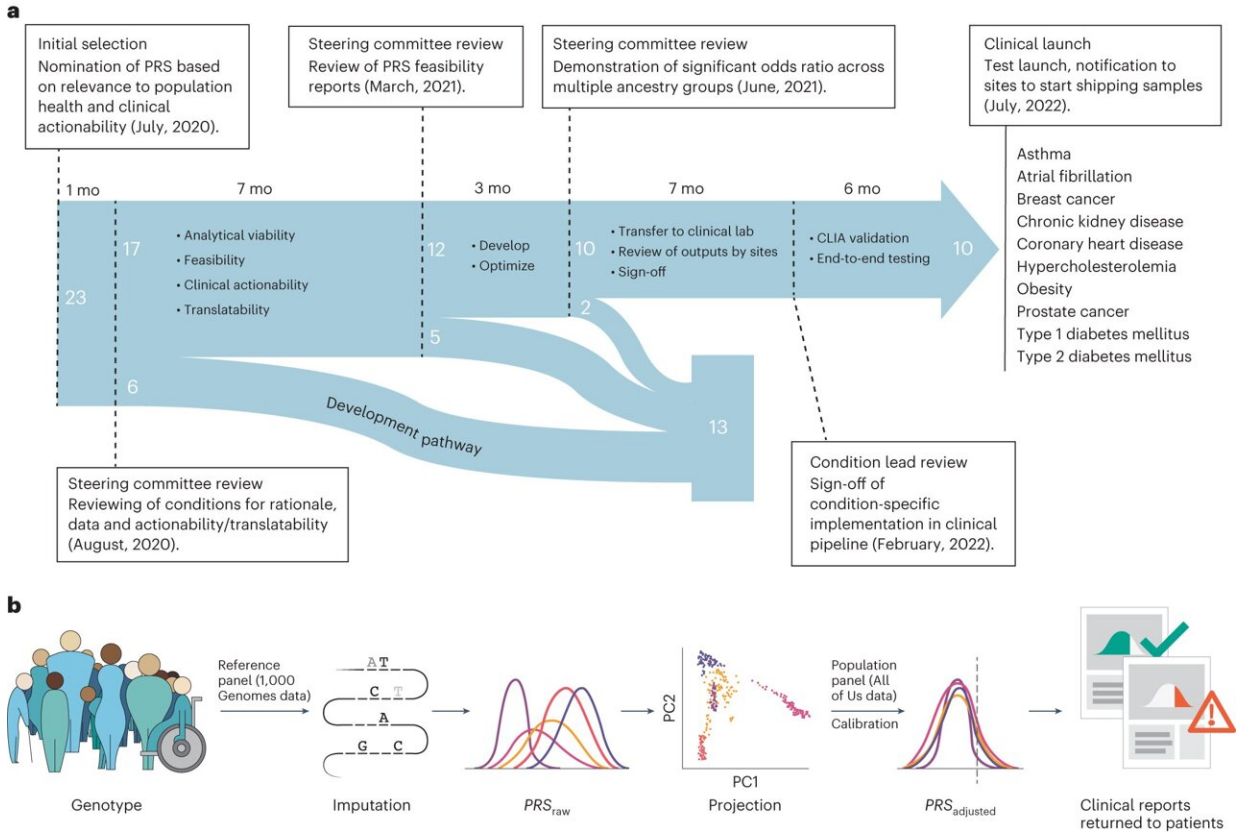
Analysis identifies more than 275 million previously unreported genetic variations

February 19 2024, by Bill Snyder



Timeline and process overview. Credit: *Nature Medicine* (2024). DOI: 10.1038/s41591-024-02796-z

An analysis of genomic data from nearly 250,000 participants in the National Institutes of Health's All of Us Research Program has identified

more than 275 million previously unreported genetic variations, nearly 4 million of which have potential health consequences.

The data, reported Feb. 19 in the journal *Nature*, constitutes a research resource that is unprecedented in its scale and diversity, as 77% of the participants historically have been underrepresented in biomedical research, and 46% are from underrepresented racial and ethnic minorities.

"Collectively we anticipate that this work will advance the promise of precision medicine for all Americans. It is a significant step to addressing the health of the diverse population of the United States," said the paper's corresponding author, Alexander Bick, MD, Ph.D., of Vanderbilt University Medical Center.

All of Us is a historic effort to gather genomic and health data from 1 million or more people of widely diverse backgrounds. VUMC, which has pioneered studies of genetic contributions to disease and why people vary in response to medications, leads the All of Us Data and Research Center.

Historically biomedical and genomic research has represented people who are primarily of European genetic ancestry. The exclusion of large groups of people from these studies has made it difficult to obtain a comprehensive understanding of human health.

As a result, estimates of the cumulative effect of multiple genetic variants, called polygenic risk scores, may not accurately reflect the true risk for developing certain diseases in underrepresented groups.

In a companion paper published Feb. 19 in *Nature Medicine*, researchers from the Electronic Medical Records and Genomics (eMERGE) Network describe how they used the All of Us Researcher Workbench to

calibrate polygenic risk scores for 10 common conditions, including diabetes, heart disease and prostate cancer, in 25,000 individuals of diverse ancestry.

This robust evaluation of risk scores across multiple genetic ancestries would not have been possible without access to the richly diverse All of Us dataset, the researchers noted.

The eMERGE network was launched in 2007 by the National Human Genome Research Institute of the NIH to connect DNA biorepositories with electronic health record (EHR) systems for large scale, high-throughput genetic research. VUMC has played a major role in eMERGE as a network site and as the national coordinating center since the network's inception.

Recruitment of All of Us participants began in May 2018. To date, more than 500,000 participants have agreed to share their EHRs, provided physical measurements and other health-related information, and donated at least one biospecimen, such as a blood sample, for storage in one of the program's biobanks.

Working with Verily, the life science subsidiary of Google's parent company, Alphabet Inc., and the Broad Institute of MIT and Harvard, VUMC developed processes for cleaning, de-identifying and standardizing data collected from participants, and built tools and cloud-computing capacity to ensure the data were accessible and secure.

In 2020 the program launched the beta version of its cloud-based research platform, called Researcher Workbench. Access is open to registered researchers affiliated with institutions that have signed a data use and registration agreement with All of Us.

As of February 2024, 680 institutions had agreements in place, allowing

more than 8.900 registered researchers to work on more than 8,400 projects. In 2022, the program also began to share health-related genetic research results with participants who choose to receive them. These results include a Hereditary Disease Report and a Medicine and Your DNA report

As of January 2024, approximately 200,000 participants had been sent invitations to review their results, and about half of them have accepted. Participants who choose to receive their health-related DNA results can schedule a meeting with a genetic counselor to discuss the results.

In their conclusion, Bick and his colleagues predicted their partnership with All of Us participants will enable scientists to "move beyond large-scale genomic discovery to understanding the consequences of implementing genomic medicine at scale"—at the level of the individual.

More information: Genomic data in the All of Us Research Program, *Nature* (2024). DOI: 10.1038/s41586-023-06957-x

Niall J. Lennon et al, Selection, optimization and validation of ten chronic disease polygenic risk scores for clinical implementation in diverse US populations, *Nature Medicine* (2024). DOI: 10.1038/s41591-024-02796-z

Provided by Vanderbilt University Medical Center