



Credit: Office of Intramural Training & Education
National Institutes of Health

Become an *All of Us* Researcher

The *All of Us* Research Program is working to enroll at least one million people who reflect the diversity of the United States. By sharing their health information with us, *All of Us* participants are helping us build one of the largest, richest, and most diverse biomedical datasets of its kind. Researchers can use this dataset to conduct thousands of studies, accelerating health research and enabling individualized prevention, treatment, and care for all of us.

**One of the largest, richest,
and most diverse biomedical
datasets ever built**

413,450+ Participants

75%

are from communities
underrepresented
in biomedical research*



45%

are from racial
and ethnic
minority groups

Combines biological factors, genomics, and social determinants
on a large, inclusive scale, capturing a wide variety of data types.

DATA AVAILABLE IN THE RESEARCHER WORKBENCH



413,350+
Survey
Responses



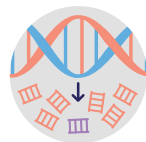
337,500+
Physical
Measurements



312,900+
Genotyping
Arrays



287,000+
Electronic Health
Records



245,350+
Whole Genome
Sequences**



15,600+
Fitbit
Records



1,000+
Long-Read Sequences

*This calculation was made using the *All of Us* Curated Data Repository (CDR) v7, available via the *All of Us* Researcher Workbench. To read more about how the *All of Us* Research Program defines communities underrepresented in biomedical research within the Researcher Workbench, read the CDR v7 data characterization report: support.researchallofus.org/hc/en-us/articles/14558858196628.

**On this graphic, whole genome sequences (WGS) refers to participants with short-read WGS data in the Researcher Workbench.

Access and tools

The *All of Us* dataset is stored on the Researcher Workbench, a secure, cloud-based platform. The program offers tiered access to the data. Researchers' institutions must first have agreements in place with *All of Us* before they can register to use the Researcher Workbench's **Registered and Controlled Tiers**.



The **Public Tier** dataset is available at **ResearchAllOfUs.org**. Visitors can explore aggregated overviews and interactive data previews—with participant identifiers removed—through the **Data Browser**. Registration is not required.



The **Registered Tier** dataset includes individual-level data from EHRs, wearables, and surveys, as well as physical measurements taken at the time of participant enrollment. Registration and a Responsible Conduct of Research training are required for access.



The **Controlled Tier** dataset includes data available on the Registered Tier, as well as genomic data and expanded demographic, survey, and EHR data. Genomic data includes whole genome sequences, long-read sequences, and genotyping arrays. Researchers must have amended institutional agreements and complete an additional training to access this tier.

The Researcher Workbench is currently supported by R, Python, Hail, and PLINK, widely-used open-source platforms for statistical analysis and data science. Tools and resources for researchers include:



Workspaces



Notebooks



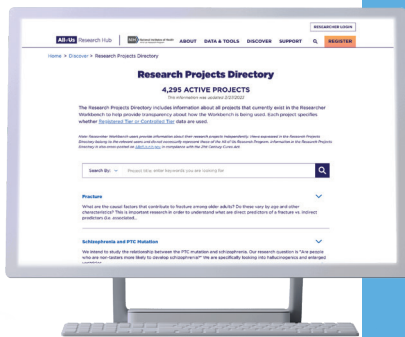
Dataset Builder



Cohort Builder



User Support Hub



Transformative research potential

- Conduct **prospective, retrospective, and cross-sectional** analyses
- Develop improved **risk assessment and prevention strategies** to preempt disease
- **Investigate health disparities** and find new approaches to **improve** health equity
- Provide **earlier and more accurate diagnoses** to reduce illness burden
- Increase **wellness and resilience** and promote healthy living
- Create a line of sight to enable new precision **treatments and interventions**

Learn how other researchers are using the data by browsing the **Research Projects Directory** and **Spotlights**.

Ready to get started? Register for access.

Visit **ResearchAllOfUs.org** to:

- Explore the data and projects
- Become a registered researcher

Connect with *All of Us*    

