All Of Us Research Program Protocol Summary

What is *All of Us*?

The *All of Us* Research Program is a historic effort to gather data about a wide variety of people living in the United States and to improve research and health. The program will examine the role that individual differences in genes, environment, and lifestyle play in disease and health, leading to the delivery of precision medicine.

What makes the *All of Us* Research Program unique?

- **Size:** More than one million participants will enroll.
- **Diversity:** The program will actively recruit from minority populations that have been historically underrepresented in biomedical research.
- **Scope:** Collection of a wide variety of molecular, environmental, physiologic, and behavioral information. The program expects to include detailed longitudinal health and exposure information from participants and retains the flexibility to enhance its scope as funding allows.
- **Duration:** 10 years or more
- **Accessibility:** Program resources are available in three tiers: public data that anyone can browse; registered data for authorized users; and controlled data, which may be accessed only by researchers who have completed all data access requirements.

For detailed information, download the full protocol.

Protocol Overview

PARTICIPANTS

Primary Objective: To build a robust research resource that can facilitate the exploration of biological, clinical, social, and environmental determinants of health and disease. The program will collect and curate health-related data and biospecimens from one million or more individuals who reflect the diversity of the United States; these data and biospecimens will be made broadly available for research uses.

Participants: Program Timeline and Enrollment

Protocol Component	Summary
Program Timeline	• The program will last 10 or more years, with active enrollment occurring in the first 5 years. • Participation is expected to last for the entire duration of the program, with regular data contribution and follow-up. • Continuous updates through repeated participant engagement, including sharing of participant electronic health records (EHRs), participant-provided information (PPI) module survey responses, and sensor/wearable technologies.

Protocol Component	Summary
Participants	● Participants not affiliated with health care provider organizations (HPOs) can enroll remotely on their own using digital interfaces or through participating HPOs.
Getting to One Million	● Active enrollment is expected through the first 5 years of the program.
Inclusion/Exclusion Criteria	● Included: Adults age 18 years and older who have the capacity to consent and currently reside in the U.S. or a U.S. territory. ● Children's participation is anticipated to begin in 2023. ● Excluded: People who are incarcerated at the time of enrollment and adults without the decisional capacity to consent. ● Any considerations to include a broader population and enroll vulnerable participants will be addressed in future protocol amendments.
Special Considerations	● The program is committed to working with Tribal Nations to respectfully address issues related to research involving American Indians and Alaska Natives (AI/AN). Before any action is taken that will significantly affect Indian Tribes, consultation will occur. ● Precautions will be utilized to ensure that there is no undue influence, such as compensation or reimbursement, to pressure potential participants join the *All of Us* Research Program or on the onsite staff responsible for recruiting interested parties.

Participants: Informed Consent	
Protocol Component	**Summary**
Methods of Consent	● Consent is conducted on-site at HPOs or remotely through the eConsent platform on the *All of Us* website and mobile app.
eConsent Modules	● eConsent is modularized to include primary consent, HIPAA Authorization for Research EHRs, and Consent for Return of Genomic Results. ● The modules include audiovisual content, screens, videos, electronic signature, and evaluations.

Participant Data Collection	
Protocol Component	**Summary**
Participant-Provided Information (PPI)	● PPI is provided online at participants' own convenience, at an HPO site, or via computer-assisted telephone interviewing (CATI). ● PPI collection involves questionnaires on topics including personal and family medical history, socioeconomic factors, health care access and utilization, and social determinants of health.
Electronic Health Records (EHRs)	● At enrollment, participants authorize sharing their EHRs with the *All of Us* Data and Research Center (DRC). ● EHR data may be sent directly to the DRC by the participant's health care providers or by the participant through Patient Portal API-based technologies and/or other vendors. ● Initial datatypes collected include demographics, visits, diagnoses, procedures, medications, laboratory visits, and vital signs. ● In the future, data collection will be expanded to all parts of the EHR, including health care provider notes, radiology, messaging, reports, and other testing if applicable.
Physical Measurements	● Physical measurements will be obtained by being collected in a clinical setting, via EHR extrapolation, and/or via self-reporting from home. ● Participants will have a standardized set of physical measurements collected and recorded in HealthPro. The same core set of baseline measurements will be carried out irrespective of enrollment through an HPO or as a DV. ● Physical measurements include blood pressure, heart rate, height, weight, and waist and hip circumference. Body mass index (BMI) will be calculated automatically from measured height and weight. ● The program will have the ability to do home visits when needed.
Biospecimens	● Participants may be invited to give blood, urine, and saliva samples at their HPO enrollment centers or designated DV biospecimen collection facilities. ● The program will have the ability to do home visits when needed.
Digital Health Data	● The program may eventually collect data from a subset of participants via their mobile phones, wellness and fitness devices, other sensors, and/or mobile apps.

Protocol Component	Summary
Participant Access to Data	● Participants will have Web and mobile access to their physical measurements, data on biospecimens, genomic information, and PPI responses, as well as additional information types as they are collected via their Participant Portal account.

DATA

Data Collection, Accessibility, and Storage	
Protocol Component	**Summary**
Electronic Core Dataset	● The core dataset will include PPI, physical measurements, digital health technology readings, genomic and baseline biospecimen assay data, and EHR-derived information from most participants. ● The program will remove explicit personal identifiers from the datasets made available for research purposes.
Accessing the Core Research Dataset	● Different tiers of access will give qualified researchers access to public, registered, or controlled tiers of data. ● The Research Hub opened to the public in May 2019 (see https://www.researchallofus.org/). ● The Registered Tier opened in May 2020, and the Controlled Tier is expected to open in spring 2022. Qualified researchers can access the data within a secure system and run their analyses using the cloud infrastructure. Researchers cannot download data. ● Data will be available via a researcher-based, rather than a project-based, mechanism. ● A diverse group of qualified researchers—including academic scientists, researchers at commercial organizations, and interested community scientists—will be able to request access to the resource.

Biospecimen Collection and Storage All biospecimen samples are first processed at the collection site. The program transports biospecimens at 4°C and stores aliquots at -80°C and ships them to the Mayo Medical Laboratories Centralized Biobank for more processing and storage.	
Protocol Component	**Summary**
Blood	● About 50 mL is collected.
Urine	● About 50 mL is collected.

Protocol Component	Summary
Saliva	● In cases where it may not be possible to obtain a blood sample, a saliva sample may be collected instead for the purposes of isolating DNA.
Data from Biospecimens	● The program is generating genomic data from participants. ● Decisions about other datatypes generated from biospecimens will be made in the future.
Accessing Biospecimens	● Biospecimens may be made available to qualified researchers with some explicit identifiers. Biospecimens are finite, exhaustible resources requiring coordinated and controlled access.
Data Security and Protection	● Data elements will be transferred through encrypted channels to the *All of Us* DRC for storage and creating a dataset accessible to researchers. The data will be stored in a secure cloud-computing environment that follows rigorous standards to protect individual privacy and data confidentiality.