



The *All of Us* Research Program

NASA Health and Air Quality Applied Sciences Team Meeting

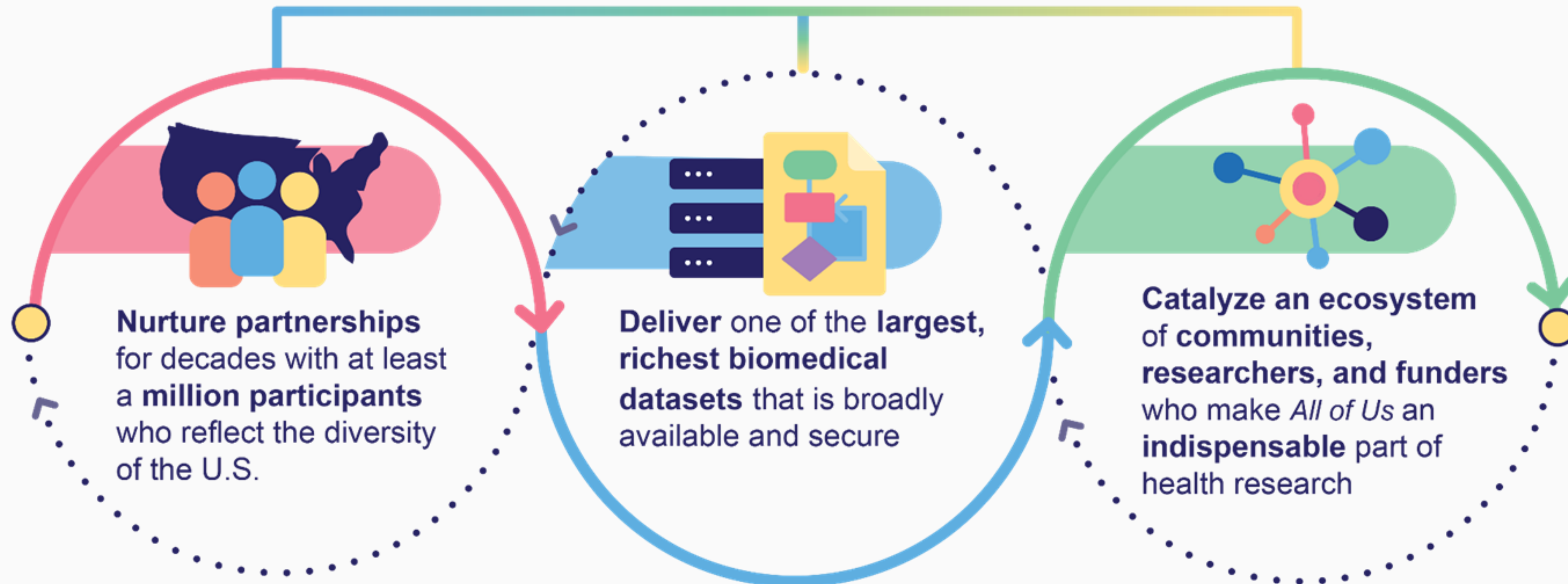
Sheri Schully, Ph.D.

Deputy Chief Medical and Scientific Officer



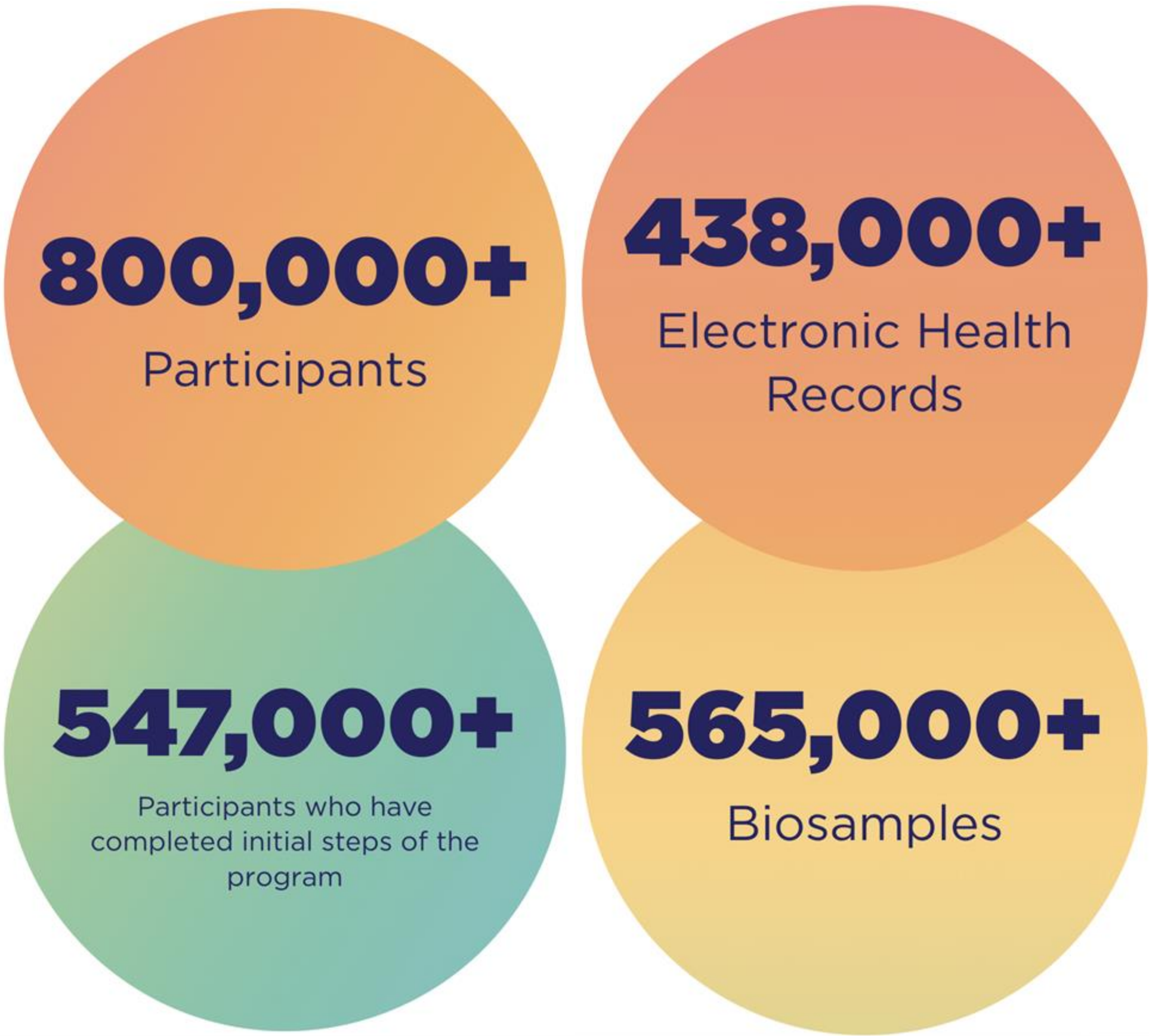
The *All of Us* Research Program Mission

Accelerate health research and medical breakthroughs,
enabling individualized prevention, treatment, and care for all of us

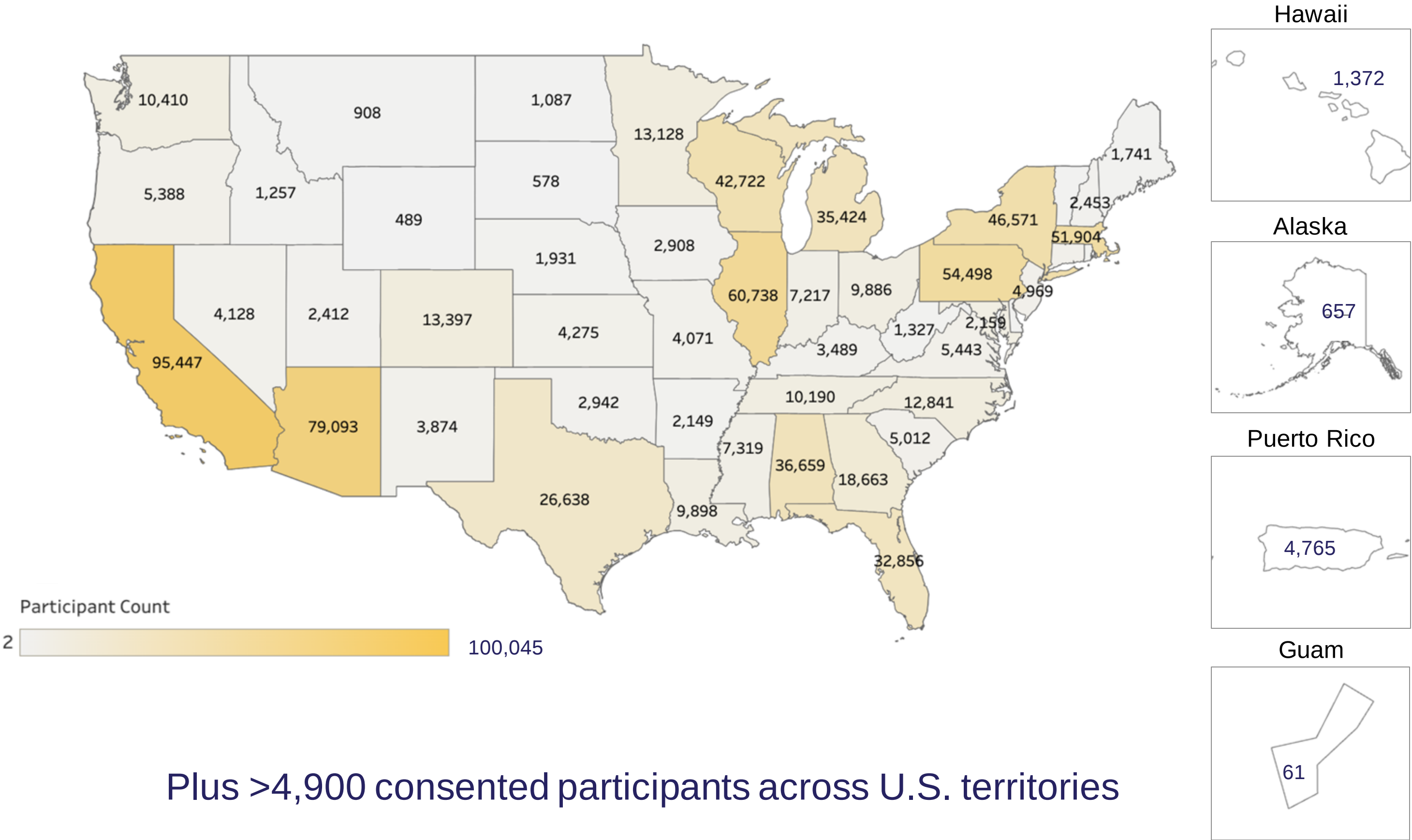


Enrolled 800K+ Participants With Continued Growth

Participant Enrollment



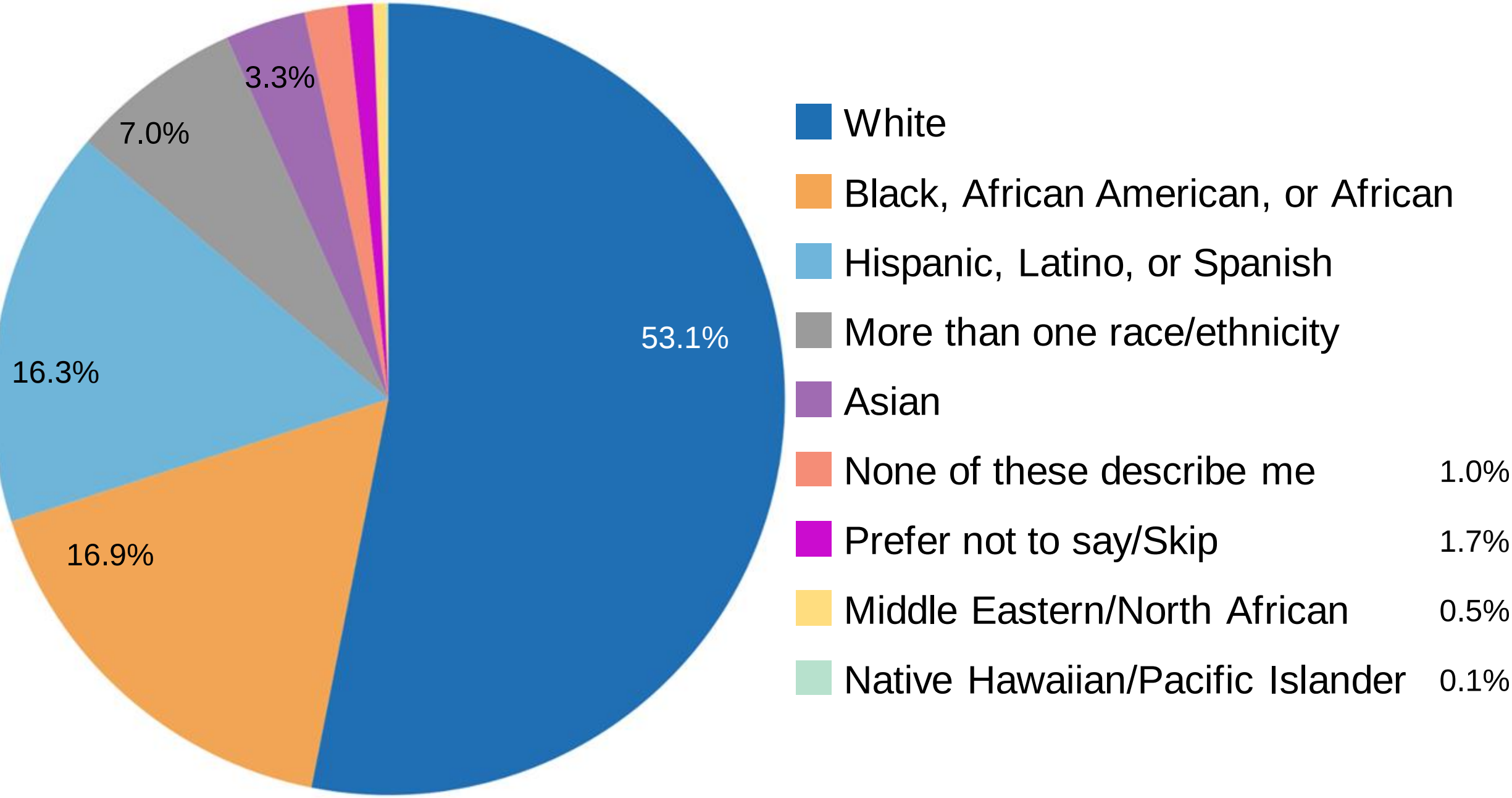
Map of Consented Participants



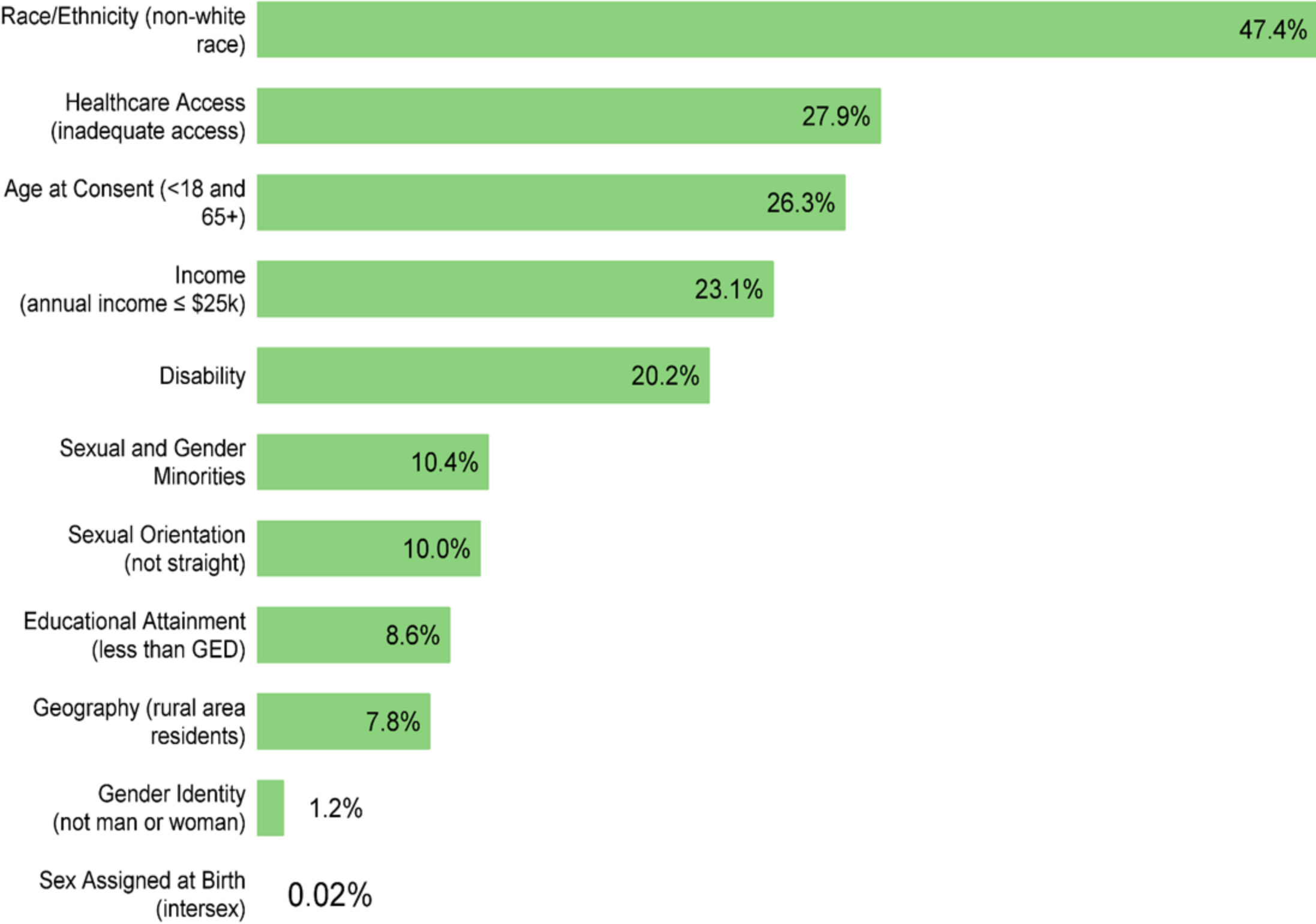
Numbers as of May 10, 2024

Participant Diversity

Race & Ethnicity of Participants



UBR Category



Over 85% of *All of Us* participants are underrepresented in biomedical research

Numbers as of May 10, 2024

Program Progress: Nearly 250,000 Whole Genome Sequences Available

The *All of Us* Research Workbench contains the one of the largest sets of whole genome sequences widely available for research.



413,350+
Survey
Responses



337,500+
Physical
Measurements



312,900+
Genotyping
Arrays



287,000+
Electronic Health
Records



245,350+
Whole Genome
Sequences



11,350+
Structural Variants

NEW! In 2023



1,000+
Long-Read Sequences

NEW! In 2023



15,600+
Fitbit
Records

NEW! Sleep Data



The whole genome sequence dataset includes variation at more than **1 billion** locations, which is nearly **one-third** of the entire human genome



Includes 275 million genetic variants never seen before, including 4 million with protein coding consequences

Environmental Data in *All of Us*

Data Linkages and Surveys

Location

Integrate locational information into the *All of Us* workbench through a decentralized address geocoding tool.

Center for Linkage and Acquisition of Data (CLAD)

Environmental Justice Index

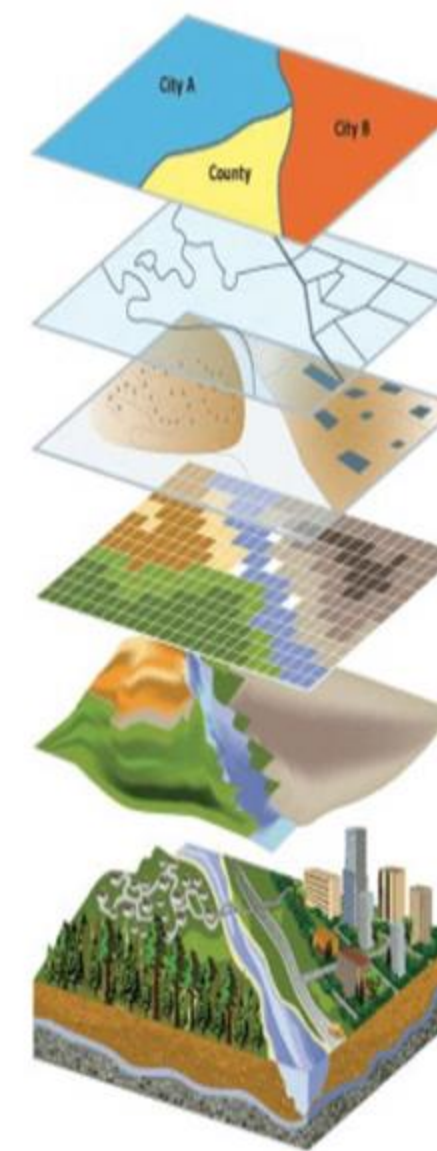
Environmental Exposure and Occupational Health (expected 2025)

Geospatial Data

Geographic Information System (GIS) Data Linkage

Growing Number of Data Layers

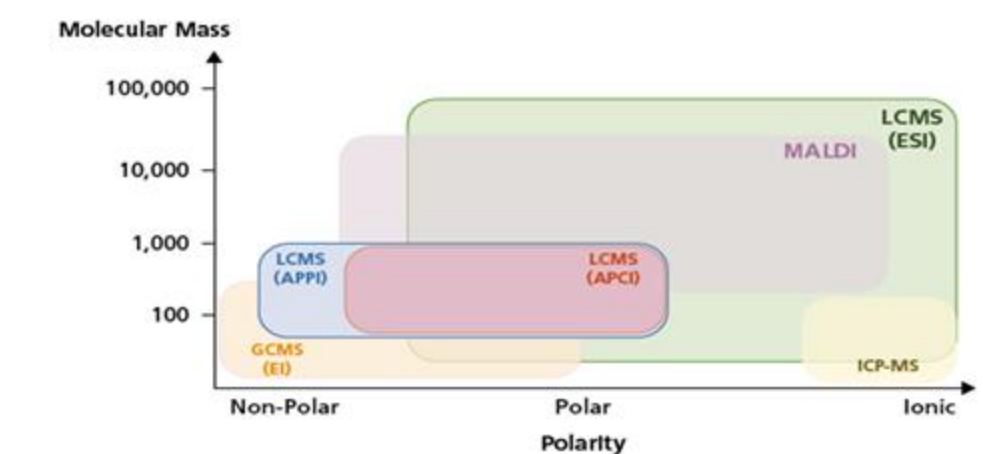
- Airports
- CAFOs
- Cellular towers
- Drinking water
- Dry cleaners
- Hazardous waste
- Highways
- Nuclear sites
- Wastewater
- Population info
- Power lines
- PR landfills
- Railroads
- Spills
- Sanitary landfills
- Superfund sites
- Toxic release sites,
- Etc.



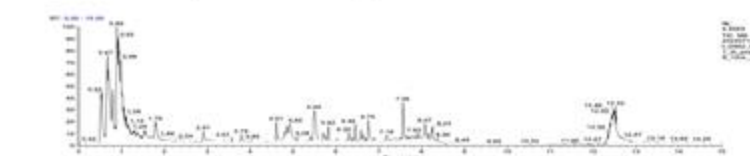
Address at the time of survey completion and longest lived childhood address

Exposomics and Health

LC- and GC-based high-resolution mass spectrometry



Complementary LC and GC HRMS



Untargeted HRMS (exposomics) through the **NIEHS** Human Health Exposure Analysis Resource.

Entire *All of Us* cohort

Subset of cohort

Environmental Data in *All of Us*

Data Linkages and Surveys

Location

Integrate locational information into the *All of Us* workbench through a decentralized address geocoding tool.

Center for Linkage and
Acquisition of Data (CLAD)

Environmental Justice Index

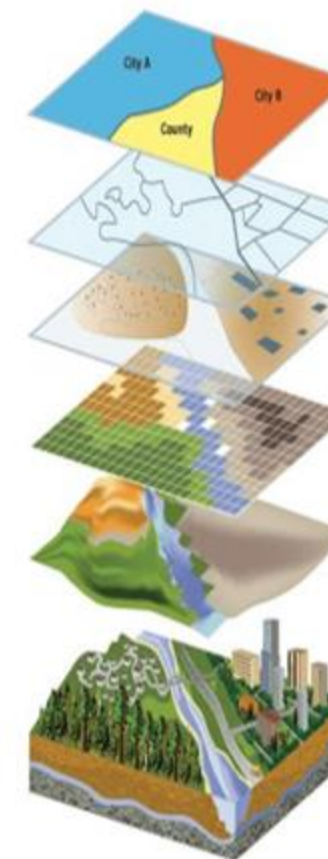
Environmental Exposure and
Occupational Health (expected 2025)

Geospatial Data

Geographic Information System (GIS) Data Linkage

Growing Number of Data Layers

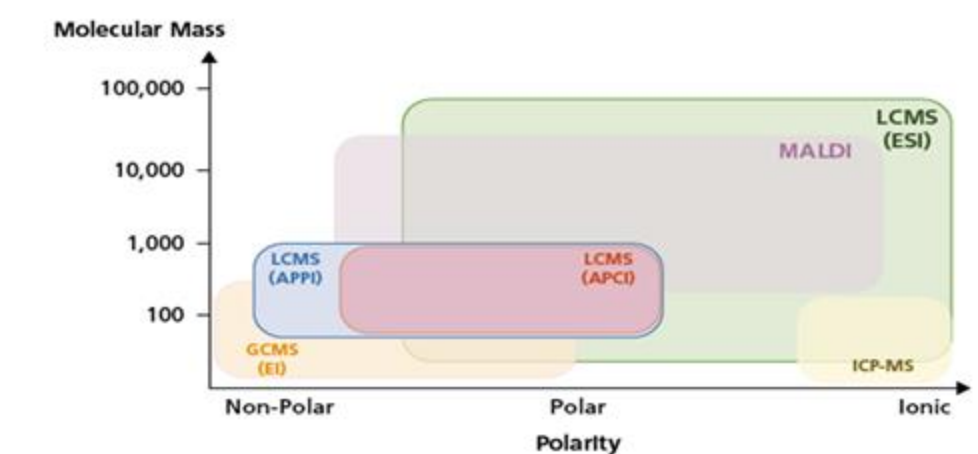
- Airports
- CAFOs
- Cellular towers
- Drinking water
- Dry cleaners
- Hazardous waste
- Highways
- Nuclear sites
- Wastewater
- Population info
- Power lines
- PR landfills
- Railroads
- Spills
- Sanitary landfills
- Superfund sites
- Toxic release sites, Etc.



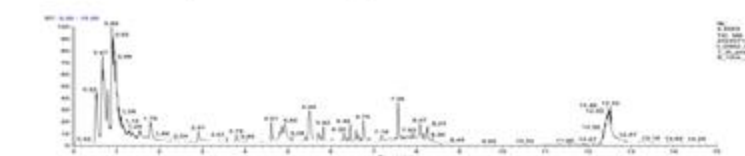
Entire *All of Us* cohort

Exposomics and Health

LC- and GC-based high-resolution mass spectrometry



Complementary LC and GC HRMS



Untargeted HRMS (exposomics)
through the **NIEHS** Human Health
Exposure Analysis Resource.

Subset of cohort

Sheri.Schully@nih.gov

Data Access →



**All
of Us**
RESEARCH PROGRAM



National Institutes
of Health

AllofUs.nih.gov



@AllofUsResearch
#JoinAllofUs