All of Us Biobank:
Scalable Biospecimen Resource

January 27, 2022

Mark S. Caulder, Biobank Program Director
All of Us Research Program

NIH National Institutes of Health

All of Us Research Program Biobank

- Concept Clearance: Continuation of the *All of Us* Research Program Biobank
- Title: *All of Us* Research Program Biobank: A Scalable Biospecimen Resource (U24)
- Objectives/Purpose: The *All of Us* Research Program Biobank will have the responsibility of monitoring and enhancing state of the art methods and technologies for biosample collection, processing, handling, management, storage, distribution and providing all support services needed for biospecimen collection.
- Funds Available and Anticipated Number of Awards: $35M (direct costs) in FY2023 for 1 meritorious award
- Award Project Period: 5 years
- Council Action: Vote for continued support of the *All of Us* Research Program Biobank

All of Us Mission

Nurture partnerships
for decades with at least a **million participants** who reflect the diversity of the US

Deliver one of the largest, richest biomedical datasets
that is broadly available and secure

Our Mission

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

Catalyze an ecosystem
of communities, researchers, and funders who make *All of Us* an indispensable part of health research

Made possible by a team that maintains a culture built around the program's core values

All of Us Research Program: Core Protocol

Enroll, Consents & Authorize EHR

- Recruiting 18+ years old initially; plan to include children later
- Online, interactive consent
- Includes authorization to share Electronic Health Record data
- Consent to get DNA Results

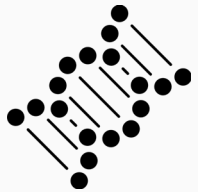
Answering Surveys

- Six initial surveys: The Basics, Overall Health, Lifestyles, Health Care Access & Utilization, Personal Medical History, Family Health History
- Additional surveys will be released on an ongoing basis.

Physical Measurements*

- Blood pressure
- Heart rate
- Height
- Weight
- BMI
- Hip circumference
- Waist circumference

**Based on diverse sampling and capacity*



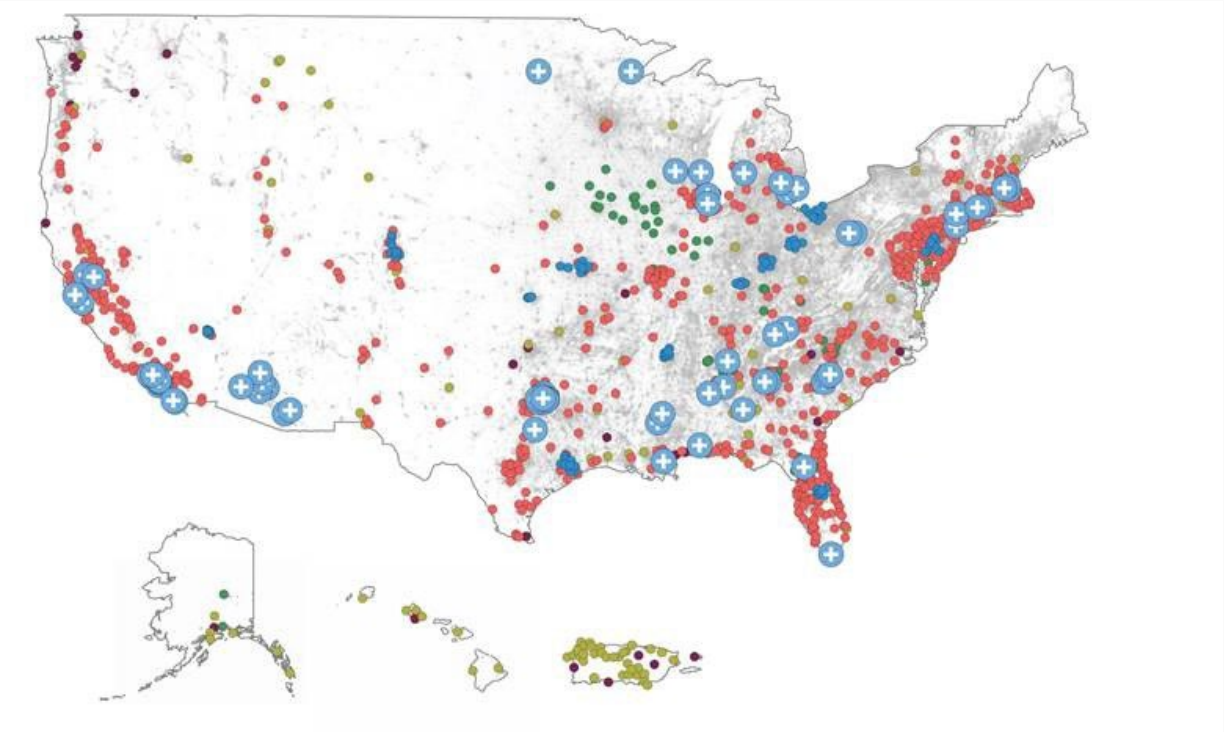
Provide Biosamples*

- Blood specimens
- Saliva specimen
- Urine specimen
- Biosamples will be stored at the program's biobank

**Based on diverse sampling and capacity*

All of Us by the Numbers

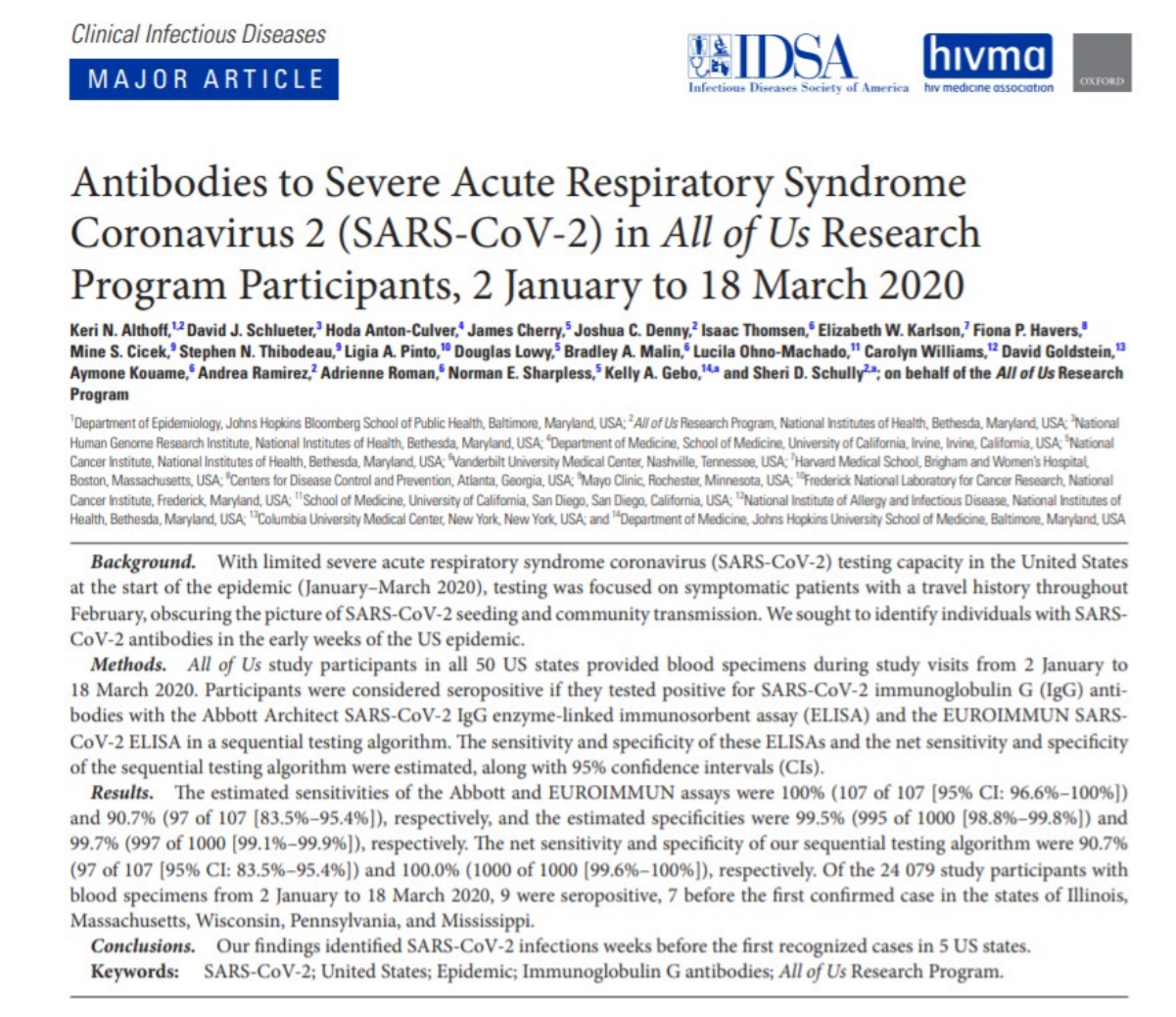
- *All of Us* Research Program
  - 447k+ participants
  - 334k+ who have donated biospecimens
    - ***2.3M primary tubes processed***
    - ***9.0M secondary aliquots stored***
  - 314k+ who have completed the core protocol
    - >80% are underrepresented in biomedical research
    - >50% are racial/ethnic minorities
- >245 clinics enrolling participants
- Published COVID serology study paper
- Launched genomic return of results pipeline



- *Participants from all 50 states*
- *Bilingual enrollment journey*

All of Us COVID-19 Serology Study

- Sought to identify individuals with SARS-CoV-2 antibodies in the early weeks of the US epidemic
- Study included 24,079 *All of Us* participants who had not withdrawn from the program and had an acceptable biospecimen for testing, which was collected during the study period (January 2, 2020 – March 18, 2020)
- Of the 24,079 study participants, 9 were seropositive, with 7 before the first confirmed cases in the states of Illinois, Massachusetts, Wisconsin, Pennsylvania, and Mississippi



Clinical Infectious Diseases

MAJOR ARTICLE

Antibodies to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) in *All of Us* Research Program Participants, 2 January to 18 March 2020

Keri N. Althoff,[1,2] David J. Schlueter,[3] Hoda Anton-Culver,[4] James Cherry,[5] Joshua C. Denny,[2] Isaac Thomsen,[6] Elizabeth W. Karlson,[7] Fiona P. Havers,[8] Mine S. Cicek,[9] Stephen N. Thibodeau,[9] Ligia A. Pinto,[10] Douglas Lowy,[5] Bradley A. Malin,[6] Lucila Ohno-Machado,[11] Carolyn Williams,[12] David Goldstein,[13] Aymone Kouame,[6] Andrea Ramirez,[2] Adrienne Roman,[6] Norman E. Sharpless,[5] Kelly A. Gebo,[14,a] and Sheri D. Schully[2,a]; on behalf of the *All of Us* Research Program

[1]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; [2]*All of Us* Research Program, National Institutes of Health, Bethesda, Maryland, USA; [3]National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA; [4]Department of Medicine, School of Medicine, University of California, Irvine, Irvine, California, USA; [5]National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA; [6]Vanderbilt University Medical Center, Nashville, Tennessee, USA; [7]Harvard Medical School, Brigham and Women's Hospital, Boston, Massachusetts, USA; [8]Centers for Disease Control and Prevention, Atlanta, Georgia, USA; [9]Mayo Clinic, Rochester, Minnesota, USA; [10]Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, Maryland, USA; [11]School of Medicine, University of California, San Diego, San Diego, California, USA; [12]National Institute of Allergy and Infectious Disease, National Institutes of Health, Bethesda, Maryland, USA; [13]Columbia University Medical Center, New York, New York, USA; and [14]Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Background. With limited severe acute respiratory syndrome coronavirus (SARS-CoV-2) testing capacity in the United States at the start of the epidemic (January–March 2020), testing was focused on symptomatic patients with a travel history throughout February, obscuring the picture of SARS-CoV-2 seeding and community transmission. We sought to identify individuals with SARS-CoV-2 antibodies in the early weeks of the US epidemic.

Methods. *All of Us* study participants in all 50 US states provided blood specimens during study visits from 2 January to 18 March 2020. Participants were considered seropositive if they tested positive for SARS-CoV-2 immunoglobulin G (IgG) antibodies with the Abbott Architect SARS-CoV-2 IgG enzyme-linked immunosorbent assay (ELISA) and the EUROIMMUN SARS-CoV-2 ELISA in a sequential testing algorithm. The sensitivity and specificity of these ELISAs and the net sensitivity and specificity of the sequential testing algorithm were estimated, along with 95% confidence intervals (CIs).

Results. The estimated sensitivities of the Abbott and EUROIMMUN assays were 100% (107 of 107 [95% CI: 96.6%–100%]) and 90.7% (97 of 107 [83.5%–95.4%]), respectively, and the estimated specificities were 99.5% (995 of 1000 [98.8%–99.8%]) and 99.7% (997 of 1000 [99.1%–99.9%]), respectively. The net sensitivity and specificity of our sequential testing algorithm were 90.7% (97 of 107 [95% CI: 83.5%–95.4%]) and 100.0% (1000 of 1000 [99.6%–100%]), respectively. Of the 24 079 study participants with blood specimens from 2 January to 18 March 2020, 9 were seropositive, 7 before the first confirmed case in the states of Illinois, Massachusetts, Wisconsin, Pennsylvania, and Mississippi.

Conclusions. Our findings identified SARS-CoV-2 infections weeks before the first recognized cases in 5 US states.

Keywords: SARS-CoV-2; United States; Epidemic; Immunoglobulin G antibodies; *All of Us* Research Program.

All of Us Genomics Pipeline and Return of Information

- **Genomic data generation**

	Samples Sent to Genome Centers	Samples in Process	Samples Processed and Passed QC
WGS	217,000	39,000	174,000
Array	263,000	34,500	228,000

- **Genomic data return to participants**

	Genetic RoR Consent "Yes"	Samples Processed	Participants Notified	Participants Completed Genetics Informing Loops	Participants Viewed Genetic Ancestry or any Trait
Array	176,000	101,000	99,000	68,000	67,000

All of Us Biobank

Background and Scale

- **What is a Biobank?**
  - A secure center(s) that functions to collect, process, store, and distribute biospecimens

- **Why is a Biobank Important?**
  - Standardizes collection procedures to minimize variation across sites and specimen types
  - Ensures the quality of the collected biospecimens
  - Support present and future research studies and investigations
    - Specimens can be difficult and expensive to collect

- **Scope and Scale of AoU**
  - 1M or more participants who reflect the diversity of the US
  - ~35M aliquots stored in Biobank
  - Longitudinal cohort with biospecimen-related reassessments

Roles and Responsibilities

- **Shipping biospecimen** collection supplies to enrollment partners and receiving/accessioning programmatic biospecimens
- **Processing biospecimens** using established standardized protocols, to produce high-quality materials for subsequent analyses
- **Storing biospecimens** under optimal conditions to minimize loss, damage, or contamination
- **Distributing biospecimens** to research investigators
- **Supporting** the development and validation of novel bioassays, and evaluation in **various pilot studies**

Roles and Responsibilities

- **Protecting** research participant **confidentiality**
- Supporting the Biobank's role in **adhering to the commitments** made to Tribes and American Indian and Alaska Native (AI/AN) peoples, which are **outlined** in the *All of Us* Research Program **Tribal Consultation Final Report**
- **Maintaining** high quality, robust, flexible and **secure information systems**
- Providing security and back-up systems and a **plan for disaster recovery**
- **Performing** all activities **in a high-quality manner**, with rigorous quality control (QC) and quality assurance (QA) programs

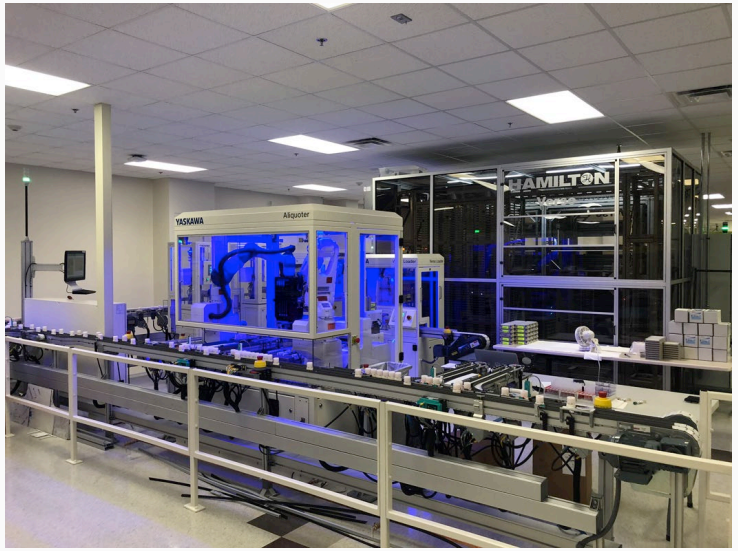
Scalability: Infrastructure Project

All of Us Core Protocol: Biospecimen Collection

Type of sample and collection tube (Collection Priority)	Volume Collected (ml)	Transport Temp °C	Fraction and (number) of aliquots created	Aliquots -80°C	LN2
(1) Clot Activator (SST)	8.5	4	(4) Serum	1.0 ml	--
(2) Plasma Separator (PST)	8	4	(4) Plasma	1.0 ml	--
(3) Na-Heparin	4	4	(2) WB (+DMSO)		1.0 ml
(4) EDTA	4	4	(2) DNA	0.5 ml	--
(5) EDTA	10	4	(5) Plasma (1) WBC (2) RBC (+glycerol)	1.0 ml	--
(6) cfDNA	9	4	(1) Plasma	5.0 ml	--
(7) PAXgene - RNA	2.5	4	1	Entire tube	--
Urine	10	4	(8) Urine	1.0 ml	--
Saliva (back-up)	2	4	(2) DNA	0.5 ml	--

30 aliquots

Automated Sample Pipeline



Current Focus: 2021 and Beyond

All of Us Research Program: 5-year Goals (2021-2026)

By 2026 we will …

Enroll 1 million participants who reflect the diversity of the US, cover the lifespan, and have shared all baseline elements. Of these participants, 500,000 are actively engaged in the program.

Expand data available for 1 million participants, to include surveys, health data streams, a whole genome sequence, environmental data, and physical measures.

Launch ancillary studies as a core and scalable capability, expanding the cohort and delivering phenotypic, lifestyle, environmental, and biologic data.

Establish a diverse global community of 10,000 researchers productively using *All of Us* data.

Incorporate participant return of value into data collections and assess its impact, including return of information to participants on genomics and EHR.

Biospecimen Related Programmatic Priorities

- Partner Support and Expansion
- Biospecimen Access for Researchers
- Bioassay Pilots
- Programmatic Reassessments
- Enhanced Withdrawal Process for American Indian and Alaska Native Participants
- Programmatic and Process Refinements/Iterations
  - Standard Operating Procedures
  - Training
  - Logistics
- Automated Biospecimen Quality Data Dashboards
- Pediatric Enrollment
- Nutrition for Precision Health

Concept Clearance

Council motion and vote to continue support for the *All of Us* Research Program Biobank through reissue of the Biobank specific FOA

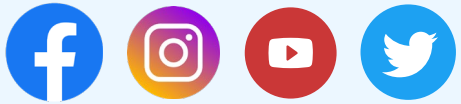
Discussion

Thank You!

joinallofus.org

National Institutes
of Health

AllofUs.nih.gov



@AllofUsResearch
#JoinAllofUs