What is the *All of Us* Research Program and How Can I Get Started?





8 November 2023

Margaret M Farrell MPH RD Science Communications Lead Division of Communications All of Us Research Program



1

At the end of this presentation, you will be able to

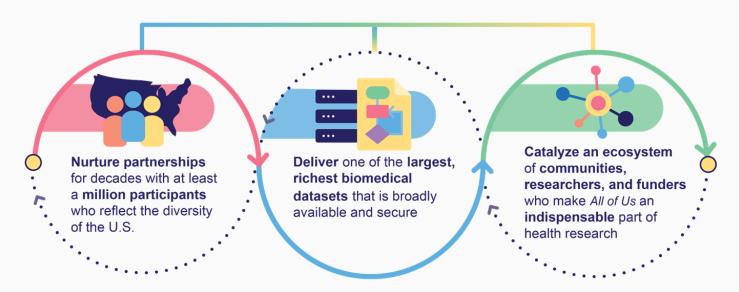
- Tell others about the All of Us Research Program
- Describe some of the research underway
- Imagine possibilities for your research questions
- Know how to register and get started



What is the All of Us Research Program?

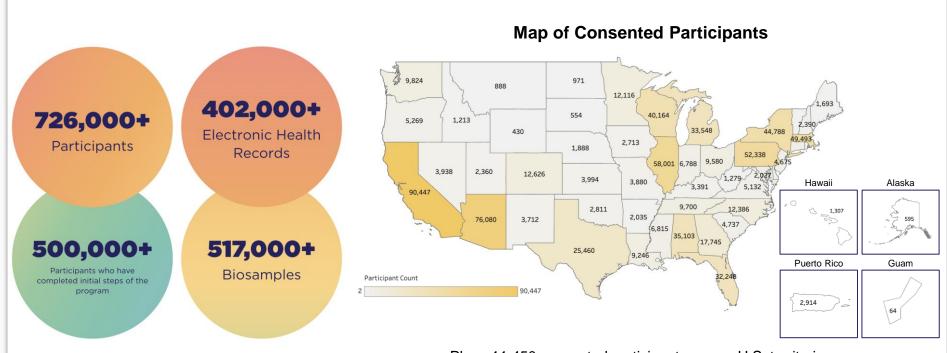
The All of Us Research Program Mission

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



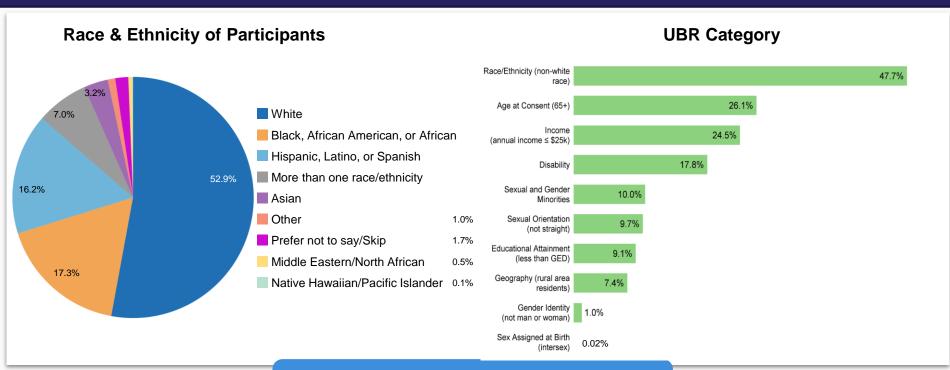
Made possible by a team that maintains a culture built around the program's core values

Enrolled 716K+ Participants With Continued Growth



Plus >11,450 consented participants across U.S. territories

Participant Diversity



Over 80% of *All of Us* participants are underrepresented in biomedical research

Numbers current as of November 3, 2023

Our participants generously contribute a variety of data



Electronic Health Records

Data types collected from EHR include:

- Demographics
- Vital signsDiagnoses
- ProceduresMedications
- Doctor and Laboratory Visits



Participant Surveys

The Basics

Health Care Access & Utilization

Overall Health Lifestyle Personal and Family Medical History Social Determinants of Health



Physical Measurements

- Blood pressure
- Heart rate
- Height
- Weight

- BMI
- Hip circumference
- Waist circumference



Biosamples

- Blood
- Saliva
- Urine



Wearable Data

Fitbit data, including:

- Heart Rate
- Activity (Daily Summary)
- Activity Intraday Steps
- Sleep data

The Basics

 Collect data on home and work locations

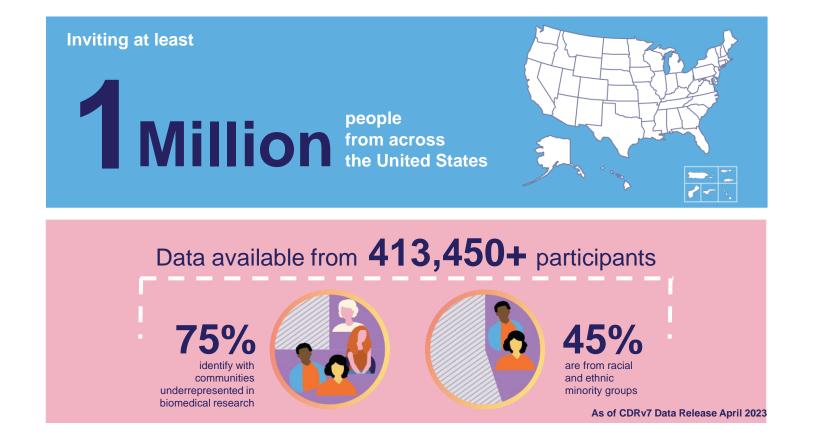
Social Determinants of Health

- Neighborhood Safety & Environment
- Social Support, Discrimination, Loneliness, Stress
- Health Care Discrimination
- Housing, Food Security
- Spiritual Experiences, Religious Attendance
- English Proficiency



What makes the *All of Us* dataset unique?

One of the world's largest, most diverse biomedical datasets of its kind



Summary of Data Available in the Researcher Workbench



413,350+ Survey Responses



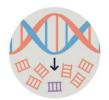
337,500+
Physical
Measurements



312,900+ Genotyping Arrays



287,000+
Electronic Health
Records



245,350+
Whole Genome
Sequences



15,600+
Fitbit
Records
NEW! Sleep Data





Data as of CDRv7 Release, April 2023

Including genomic data

As of April 2023, the *All of Us* Researcher Workbench contains the **largest set of whole genome sequences widely available for research.*** The whole genome sequence dataset includes variation at more than **1 billion** locations, which is nearly **one-third** of the entire human genome.



312,900+ Genotyping Arrays



1,000
Long-Read Sequences



245,350+
Whole Genome Sequences



Genomics Analysis Tools
Hail and PLINK in addition to R,
Python, and Jupyter Notebooks

Genomic Data Is Paired With Rich Phenotypic Data



206.100+

Have Whole Genome Sequences + Electronic Health Records + Physical Measurements + Survey Responses



245,100+

Have Whole Genome Sequences + Physical Measurements + Survey Responses



206.150+

Have Whole Genome Sequences + Electronic Health Records



+008,8

Have Whole Genome Sequences + Fitbit Records
Fitbit data may include physical activity, step counts, heart rate, and
sleep data

11

All of Us Data Can Power Research Across Range of Diseases

Conditions	Domain	Participants
Heart Disease	Heart	73,100
Obesity	Endocrine	65,740
Type 1 Diabetes	Endocrine	5,620
Type 2 Diabetes	Endocrine	45,360
Any cancer	Cancer	42,080
Asthma	Pulmonary	40,560
Chronic Obstructive Pulmonary Disorder	Pulmonary	19,740
Epilepsy	Neuro	7,440
Stroke	Neuro	640
Rheumatoid Arthritis	Autoimmune	6,980
Osteoarthritis	Autoimmune	81,980

Conditions	Domain	Participants
Depressive Disorder	Mental Health	67,380
Bipolar Disorder	Mental Health	12,820
Dementia	Mental Health	4,760
Human Immunodeficiency Virus	Infectious Disease	4,640
COVID-19*	Infectious Disease	58,000*
Alcohol Abuse (AUDIT-C)	Abuse	84,000
Opioid Usage	Medication	155,000
Age-Related Macular Degeneration	Eye	4,740
Hearing loss	Hearing	30,400
Falls	Aging/Nursing	4,860

^{*}using combination of diagnosis code, lab test, and COVID survey answers



What research is underway?

Researcher Workbench Usage and Diversity (data as of November 1, 2023)



7,550+ Registered Researchers



7,650+ Active Projects



210+ Publications using All of Us data



600+ Organizations

40 Historically Black Colleges & Universities 57 Hispanic Serving Institutions



Top Conditions

studied in the Researcher Workbench include:

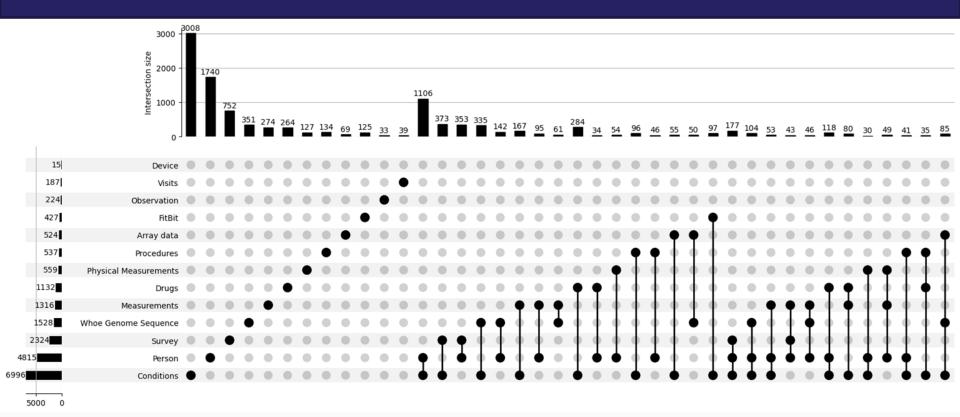
Cardiovascular disease, Hypertension, Mental Health, Cancer, Diabetes

- Creating a diverse researcher cohort that promotes responsible and ethical use of data, returns value to participant communities, and accelerates research impact.
- Encouraging student assemblies and early-stage investigators to bring fresh, creative perspectives & innovative research outcomes.
- Ensuring access for researchers from various institutions/ organizations to establish a truly equitable resource for all.

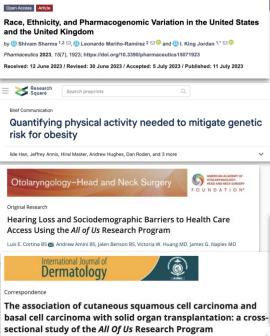


Over 72% of our researchers are underrepresented in the biomedical workforce - including over 27% diverse by race and ethnicity

Most Research Projects Draw From Several Data Types



More than 210 Publications Detail the Scientific Impact of All of Us



Annika Belzer BS, Audrey C. Leasure MD, MHS, Jeffrey M. Cohen MD, Sara H. Perkins MD

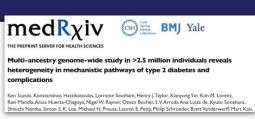
First published: 05 May 2023 | https://doi.org/10.1111/ijd.16700

PLOS ONE

Family and personal history of cancer in the All of Us research program for precision medicine

Lauryn Keeler Bruce . Paulina Paul, Katherine K, Kim, Jihoon Kim, Theresa H, M, Keegan, Robert A, Hiatt, Lucila Ohno-Machado, On behalf of the All of Us Research Program Investigators

Published: July 17, 2023 + https://doi.org/10.1371/journal.pone.0288496



Using machine learning to develop a clinical prediction model for SSRI-associated bleeding: a feasibility study

<u>Jatin Goyal, Ding Quan Ng, Kevin Zhang, Alexandre Chan, Joyce Lee, Kai Zheng, Keri Hurley-Kim, Lee</u> Nguyen, Lu He, Megan Nguyen, Sarah McBane, Wei Li & Christine Luu Cadiz

JAMA Network

BMC Medical Informatics and Decision Making 23, Article number: 105 (2023) Cite this article

Original Investigation | Oncology

August 10, 2023

Alcohol Consumption Among Adults With a Cancer Diagnosis in the All of Us Research **Program**

Mengyao Shi, MBBS, MPH1; Chongliang Luo, PhD1; Oluseye K. Oduyale, MD1; et al.

Research article Open Access | Published: 11 June 2023

Clinical Pharmacology & Therapeutics

Article

Drug-Induced Liver Injury with Commonly Used Antibiotics in the All of Us Research Program

Shaopeng Gu 🔀 Govarthanan Rajendiran, Kennedy Forest, Tam C. Tran, Joshua C. Denny, Eric A. Larson, Russell A. Wilke

First published: 07 May 2023 | https://doi.org/10.1002/cpt.2930

Article | Open Access | Published: 10 October 2022

Association of step counts over time with the risk of chronic disease in the All of Us Research Program

Hiral Master, Jeffrey Annis, Shi Huang, Joshua A. Beckman, Francis Ratsimbazafy, Kayla Marginean Robert Carroll, Karthik Natarajan, Frank E. Harrell, Dan M. Roden, Paul Harris & Evan L. Brittain

Nature Medicine 28, 2301-2308 (2022) Cite this article



Open Access • Published: January 20, 2023 • DOI: https://doi.org/1

Original Investigation | Equity, Diversity, and Inclusion

July 31, 2023

Prevalence of 12 Common Health Conditions in Sexual and Gender Minority Participants in the All of Us Research Program

Nguyen K. Tran, PhD, MPH1,2; Mitchell R. Lunn, MD, MAS1,2,3; Claire E. Schulkey, PhD4; Samantha Tesfaye, BA; Siddhartha Nambiar, PhD5: Snigdhansu Chatteriee, PhD6: Dawn Kozlowski, MEd7: Paula Lozano, PhD8,9: Fornessa T. Randal, MCRP8.9; Yicklun Mo, MSW8.9; Siya Qi, MS8.9; Ell Hundertmark, BS^{1,10}; Chloe Eastburn, BA^{1,11}; Anthony T. Pho, PhD^{1,2} Zubin Dastur, MS, MPH^{1,10}, Micah E, Lubensky, PhD^{1,12}, Annesa Flentie, PhD^{1,12,13}, Juno Obedin-Maliver, MD, MPH, MAS^{1,3,10}

Nuclear genetic control of mtDNA copy number and heteroplasmy in humans

Rahul Gupta , Masahiro Kanai, Timothy J. Durham, Kristin Tsuo, Jason G. McCoy, Anna V. Kotrys, Wei Zhou, Patrick F. Chinnery, Konrad J. Karczewski, Sarah E. Calvo, Benjamin M. Neale 2 & Vamsi K. Mootha

Nature 620, 839-848 (2023) Cite this article

nature

Higher Hospital Frailty Risk Score Is Associated With Increased Risk of Stroke: **Observational and Genetic Analyses**

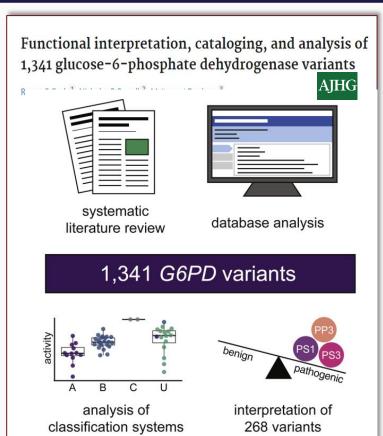
Daniela Renedo, Julián N. Acosta, Andrew B. Koo, Cyprien Rivier, Nanthiya Sujijantarat, Adam de Havenon, Richa Sharma, Thomas M, Gill. Kevin N. Sheth. Guido J, Falcone and Charles C. Matouk F

Originally published 22 May 2023 | https://doi.org/10.1161/STROKEAHA.122.041891 | Stroke. 2023;54:1538-1547



researchallofus.org/publications/

Publications Detail Not Only How All of Us is Poised to Drive New Discoveries...





But Also, How *All of Us* is Stimulating Research With Communities Traditionally Underrepresented in Biomedical Research...





Reference: Tran, N. K., Lunn, M. R., Schulkey, C. E., Tesfaye, S., Nambiar, S., Chatterjee, S., Kozlowski, D., et al. (2023). Prevalence of 12 common health conditions in sexual and gender minority participants in the All of Us Research Program. JAMA Network Open, 6(7), e2324969.

...and Powering New Understandings Within Populations

Concise Research Report | Published: 03 May 2022

Revisiting the Latino Epidemiologic Paradox: an Analysis of Data from the All of Us Research Program

Journal of General Internal Medicine 37, 4013–4014 (2022) | Cite this article

983 Accesses | 1 Citations | 409 Altmetric | Metrics

INTRODUCTION

Black individuals living in the United States have higher rates of cardiovascular disease (CVD) than non-Hispanic White (NHW) people. Among Hispanic/Latino individuals, the data are more complex. Despite higher rates and/or poorer control of CVD risk factors and multiple other social and economic disadvantages, age-adjusted CVD mortality rates are one-third lower in Latino than NHW individuals. L2 Similarly, age-adjusted self-reported rates of CVD are about ten percent lower. To date, the reasons for this "Latino Epidemiologic Paradox" remain unclear. 3-4 Explanatory hypotheses such as misclassification and immigration-related factors have been refuted. Many also question whether the paradox actually exists. In this study, we analyze data from a diverse national cohort which includes > 40,000 self-reported Hispanic/Latino people to detect for evidence of this Paradox.

METHODS

Created under the National Institutes of Health's Precision Medicine Initiative, the All of Us Research Program (AoURP) has enrolled over 280,000 core participants of which 80% are from groups historically under-represented in research.⁵ Data sources include survey data,



Reference: Montanez-Valverde, R., et al. (2022). Revisiting the Latino Epidemiologic Paradox: An analysis of data from the All of Us Research Program. Journal of General Internal Medicine. 37(15), 4013-4014.

"More to come": More data and more data types from more participants

2019

Data Browser

(launched May 2019)

Interactive tool available to the public that provides summary statistics from the program's growing database, allowing researchers to understand the characteristics of our participant population, explore the data types available and plan research questions

2020

Researcher Workbench

(launched May 2020)

Released of the first version of the All of Us dataset and the first set of researcher tools for beta testing

- Survey data
- Physical measurement data
- EHR data

2021

Expansion of data and tools

(throughout 2021)

Expanded of the *All of Us* dataset with additional survey, measurement, and EHR data.

- Start of genotyping and whole genome sequencing
- Initial return of results to participants
- >1,000K researchers registered

2022

Controlled Tier & Genomics

(ongoing)

Launch our Controlled tier, available to approved researchers who have taken additional steps and training to access:

- · Genomics data
- Individual and granularlevel information

Available in Registered and Controlled tiers, with basic information available via our public data browser:

- ACS data linkage
- New surveys

2023 and beyond

Additional data types and tools

(ongoing)

Expect to enroll 1M participants within 5-7 years of launch and continue to expand the All of Us dataset with new data releases, including over time:

- Assay data
- Additional wearables data
 - Fitbit sleep data
 - Apple Healthkit
- New surveys (e.g., Mental Health & Wellbeing)
- New Linkages to external data sources
- Expanded researcher access
- Ancillary study data

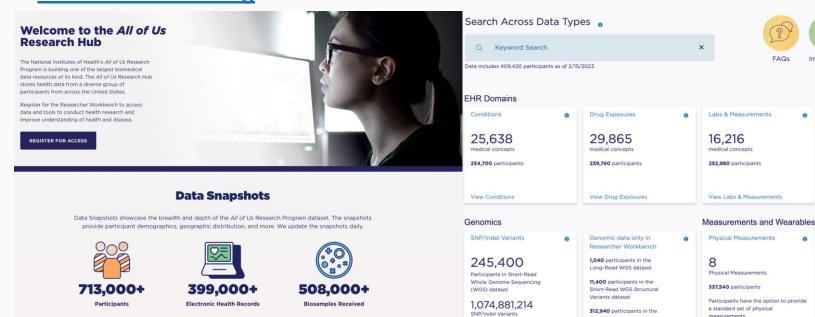


How can you get started?

First: Is All of Us right for you? Will you find the data you need?

Research Allof Us.org

VIEW MORE DATA SNAPSHOTS



Introductory User Guide

Procedures

30.328

medical concepts

View Procedures

Fitbit Measurements

15,620 participants

activity summaries.

View Fitbit

Fitbit data includes heart rate and

Fitbit

6

View Physical Measurements

Genotyping Arrays dataset

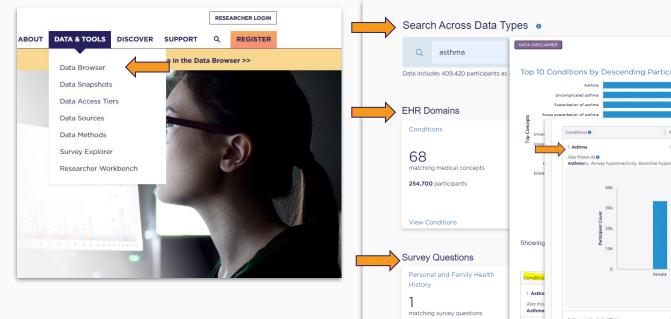
Register for access

View SNP/Indel Variants

242,580 participants

Videos

Quick case study: Driving research around asthma



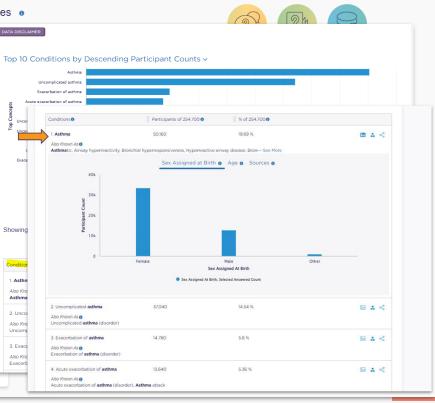
185,240 participants

View Complete Survey

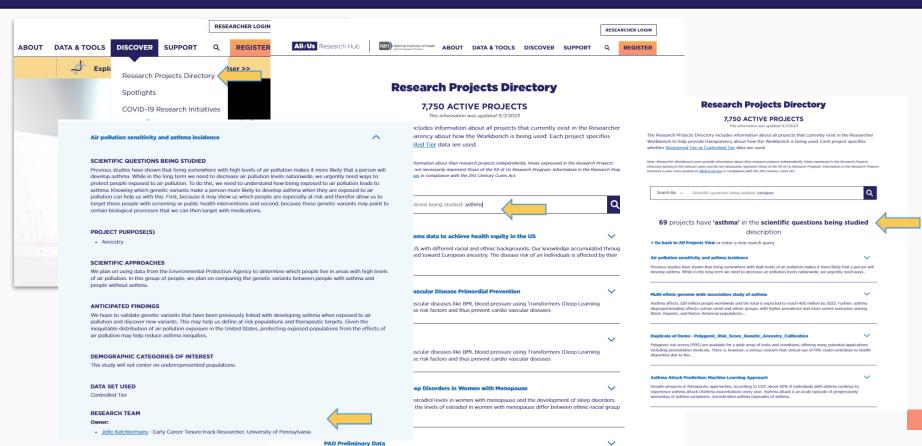
Survey includes information about th

medical history of a participant and

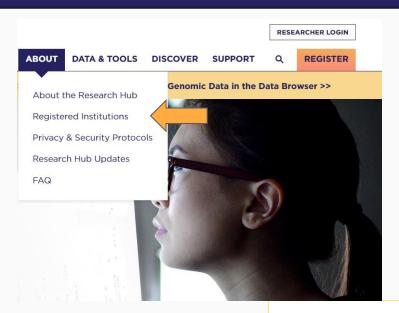
their immediate family members.



Use the Research Project Directory to Explore Work Underway

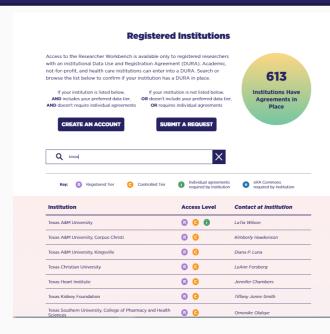


Institutions Must Have a Data Use and Registration Agreement with All of Us



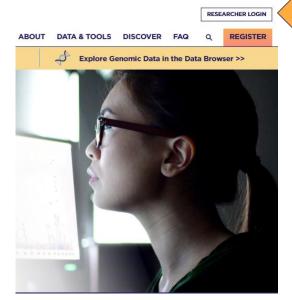
https://allof-us.org/DURAForm

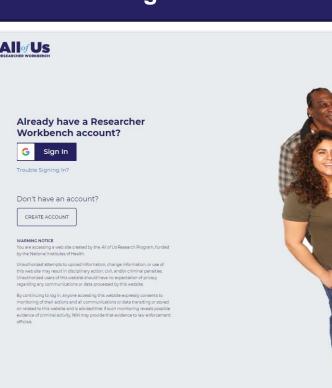




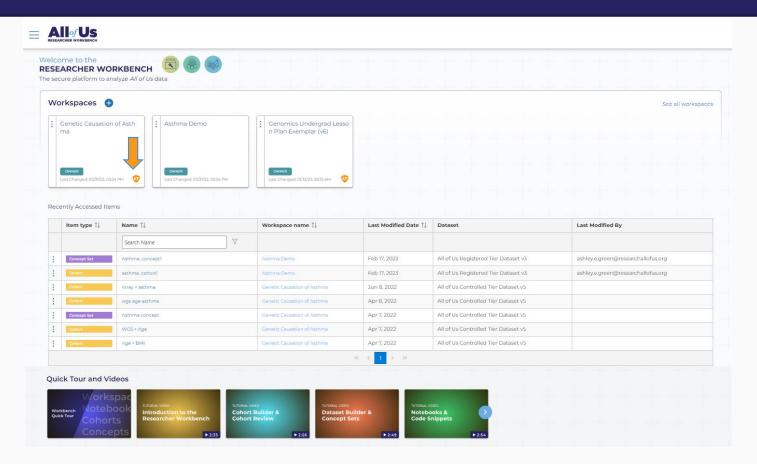
All 8 Texas public health institutions and half of the public general academic universities have these agreements in place

With the DURA in hand, registration is straightforward

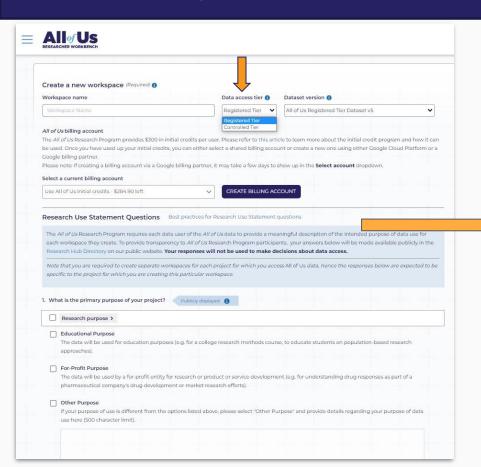




Within the workbench, you can leverage collaborative workspaces, cohort building tools, interactive notebooks, and more



All of Us employs a data passport model; transparency is a core value



COPE Covid-19 Mitigation Behaviors

PROJECT PURPOSE(S)

Population Health

SCIENTIFIC QUESTIONS BEING STUDIED

Answers to COPE survey questions regarding adherence to COVID-19 mitigation recommendations will be used to answer the following questions:

- 1. Do responses vary by demographic characteristics including age, gender, race/ethnicity, household characteristics, employment status, and geographical variables such as urban/rural and state of residence?
- 2. Are differences in responses by state associated with contemporaneous state mandates regarding COVID-19 mitigation behavior, controlling for other factors identified in #1 above?
 3. Do mitigation behaviors change over time?

Identification of variables associated with mitigation behaviors during the current pandemic may help improve further mitigation efforts, including for potential future pandemics.

SCIENTIFIC APPROACHES

Using the datasets and tools within the AoURP Research Workbench for cohort creation and statistical analyses (including the R statistical program), we will:

 Generate descriptive statistics on the sample of participants who answered the COPE surveys and compare with those of the overall AOURP cohort to assess generalizability of findings to the cohort and nationally.

- Compare responses re: COVID-19 mitigation behaviors by demographic and geographic characteristics overall and between states, using bivariate statistics appropriate for the variable types.
- 3. Examine responses by level of contemporaneous state mandates regarding mitigation behaviors in multivariable regression models controlling for all other variables found to be associated with these behaviors in 2.
- Perform a longitudinal analysis of changes in mitigation behaviors for individuals answering all 3 COPE surveys.

ANTICIPATED FINDINGS

We expect to that mitigation behaviors are associated with many different demographic and geographic areas and change over time. Identification of the associated variables will contribute to the understanding of the effectiveness of public health messages and mandates, within demographic, geographic, and temporal contexts.

DEMOGRAPHIC CATEGORIES OF INTEREST

This study will not center on underrepresented populations.

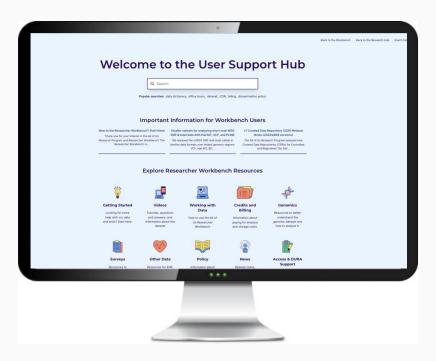
RESEARCH TEAM

Owne

• Jill Waalen - Mid-career Tenured Researcher, Scripps Research



Robust support available via the User Support Hub



Find instructional materials about the *All of Us*Researcher Workbench

Includes video tutorials, educational resources, release notes, and more

Search 290+ articles*

Find information about data dictionaries, the Controlled Tier directory, how WGS and array data are organized, and more

Connect with experts during weekly LIVE office hours

Explore our calendar of dedicated office hours where researchers can talk about data types and tools with experts



researchallofus.org/support

Stay in touch to learn more



Subscribe to Research Roundup

Stay up to date on the latest news and insights from *All of Us* through our **bimonthly email newsletter.**



allof-us.org/RRSignup

The *All of Us* Program wouldn't be possible without the generosity of our participants and the dedication of our researchers to enable health discoveries.





All of Us Community and Provider Partner Network (as of April 2023)













































































































All of Us Consortium Members (beyond community partners, as of April 2023)

The Participant Center











All of Us

New England

mass General Brigham









New York City

COLUMBIA

Weill Cornell

HEALTH+ HOSPITALS Harlem

Consortium





Communications & Engagement

IS HEALTH

TUSKEGEE

LSU Health

Nutrition for

WONDROS

All of Us Southern Network



















Trans America

Consortium

HONNY FORD

Essentia Health

BaylorScott&White

HealthPartners Institute

RELIANT MEDICAL GROUP



Spectrum







LIGE MEDICINE

MEDICAL CENTER

Tulane University

HEERSINK



HPO Network

(Health Care Provider Organizations)

HPO Lite









All of Us California

UCI Health







Precision Medicine Consortium

UC San Diego Health









Northwestern Medicine CHICAGO **♦NorthShore**





Illinois



All of Us Pennsylvania

University of Arizona and Banner Health

THE UNIVERSITY
OF ARIZONA









Sun River Health



NewYork-

☐ Presbyterian







VA Medical Centers







Precision Health (NPH)

All of Us Southeast **Enrollment Center**



















All of Us Wisconsin Marshfield Clinic









WISCONSIN

































































Thank You!



Research Allof Us.org



National Institutes of Health

AllofUs.nih.gov



The All of Us Research Program

Engages people & communities who have been left out of medical research in the past



Combines biological factors and social determinants on a large, inclusive scale



Follows participants

as they move, age, and grow



Easily accessible

to any researcher with a secure internet connection

