November 12, 2020

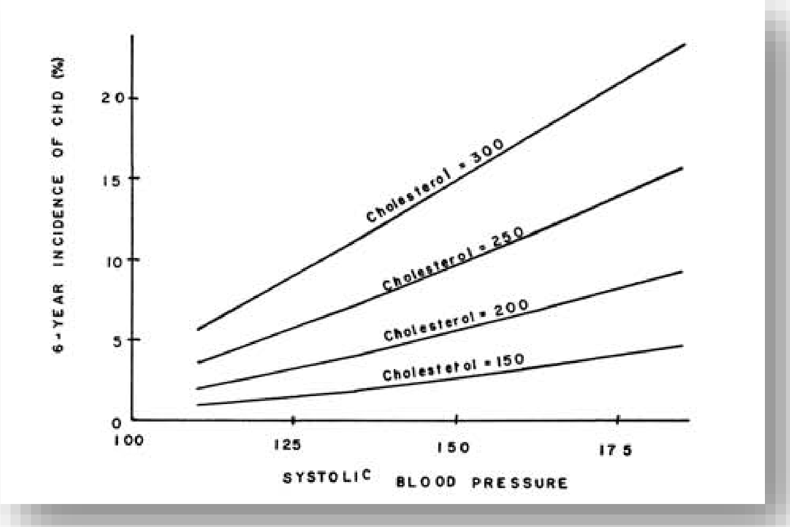
All of Us
Researcher Onramp

All of Us Research Program: Advancing Precision Medicine for All Populations

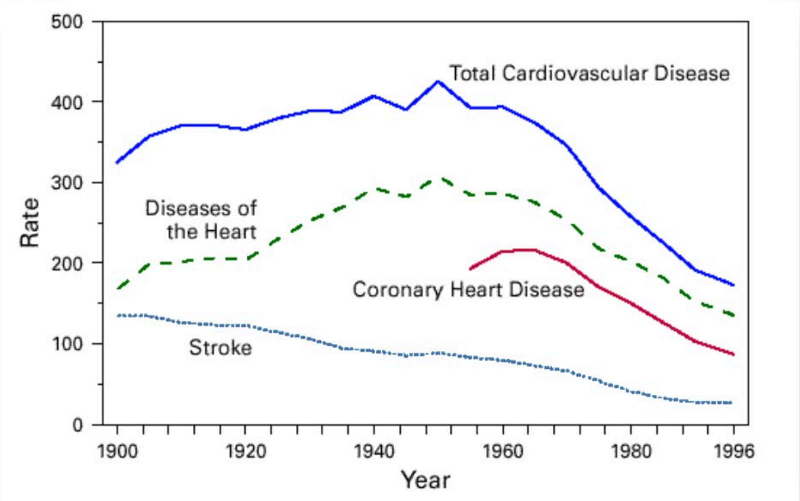
Josh Denny, M.D., M.S.
All of Us CEO

Large Cohort Studies Have Transformed Disease Treatment

Framingham and other cohorts have taught us much about heart disease...



...and cardiovascular deaths have decreased

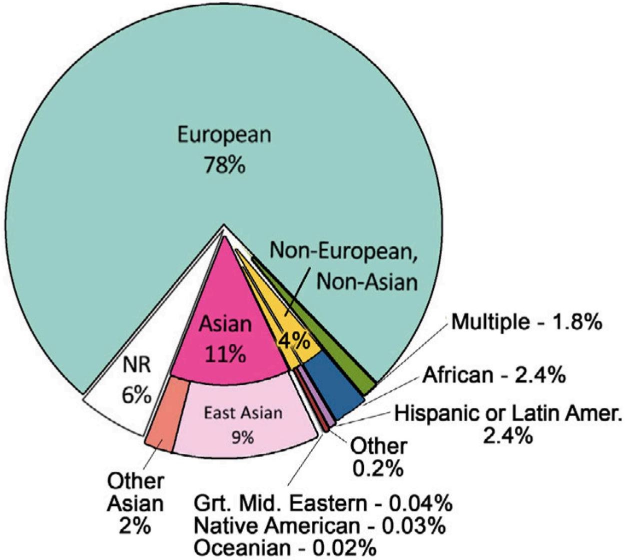


Kannel et al., *Ann Intern Med.*, 1961; CDC, *MMRW Morb Mortal Wkly Rep.*, 1999

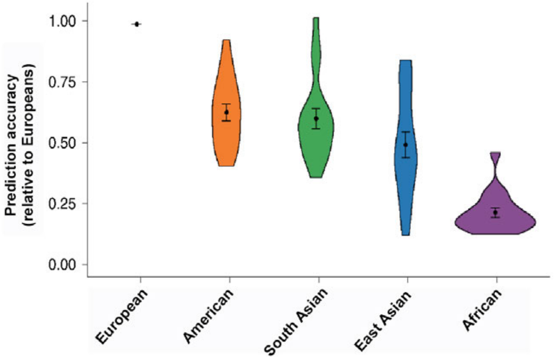
Could we have a similar experience with precision medicine in the next 40 years?

The Lack of Diversity in Research Contributes to Health Disparities

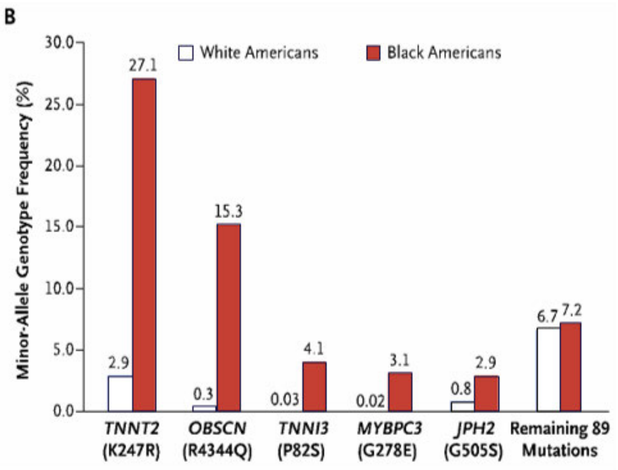
There is a lack of diversity in genome-wide association studies.



This affects polygenic risk scores for diseases…



…and interpretation of genetic variants



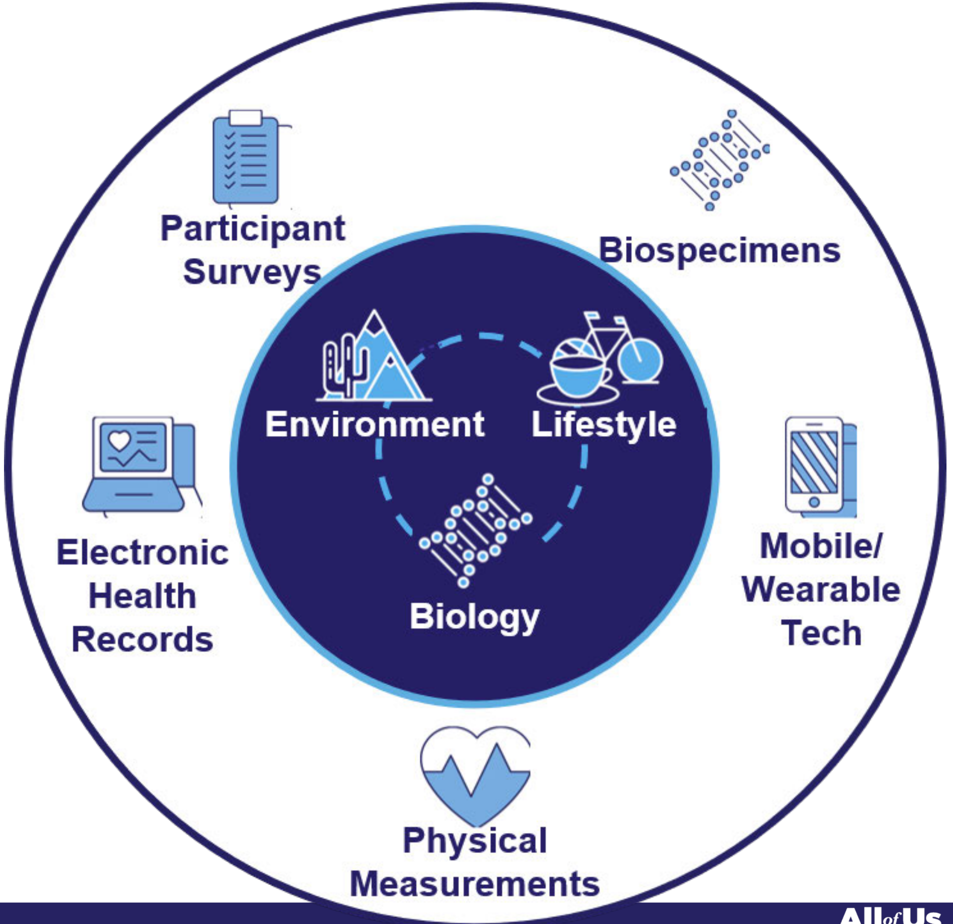
Morales *et al.*, *Genome Biol*, 2018; Martin et al., *Nat Genetics*, 2019; Manrai et al., *NEJM*, 2016

All of Us Will Provide Answers and Make Connections

- National launch in 2018
- Goal is to have one million participants that reflect the diversity of the U.S.
- **Participants are partners**
- Researchers will use data to better understand health and disease

JoinAllofUs.org

ResearchAllofUs.org



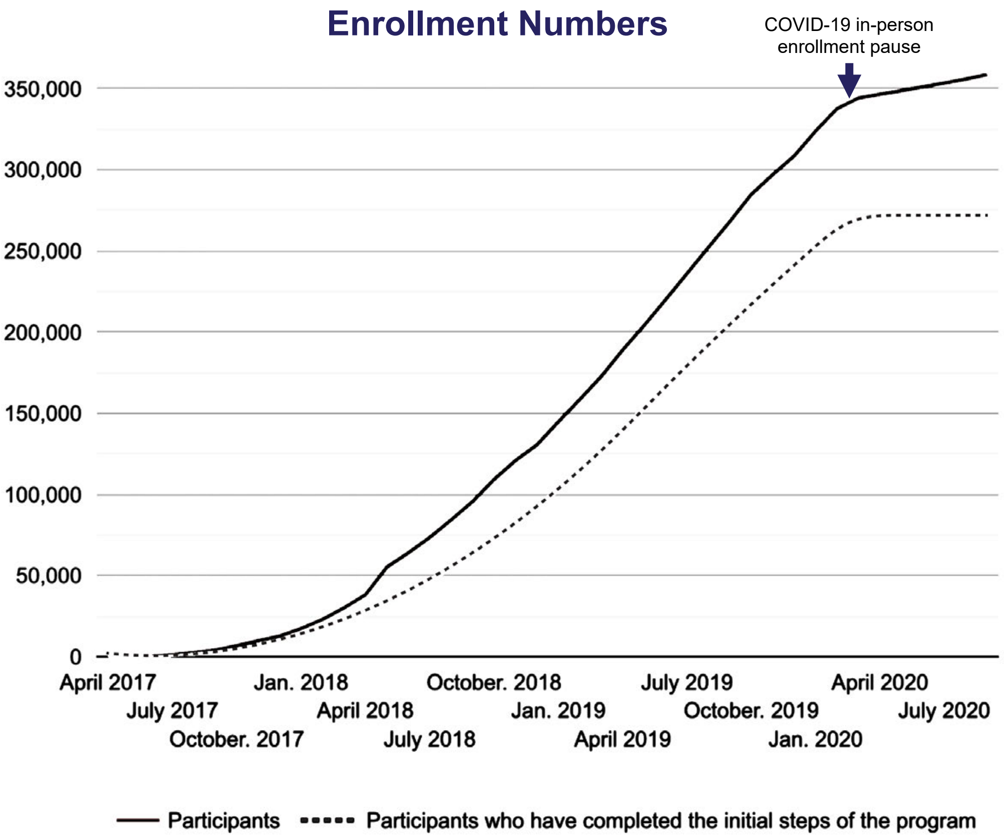
Status of the Program (as of November 8, 2020)

360,000+
Participants

230,000+
Electronic Health Records

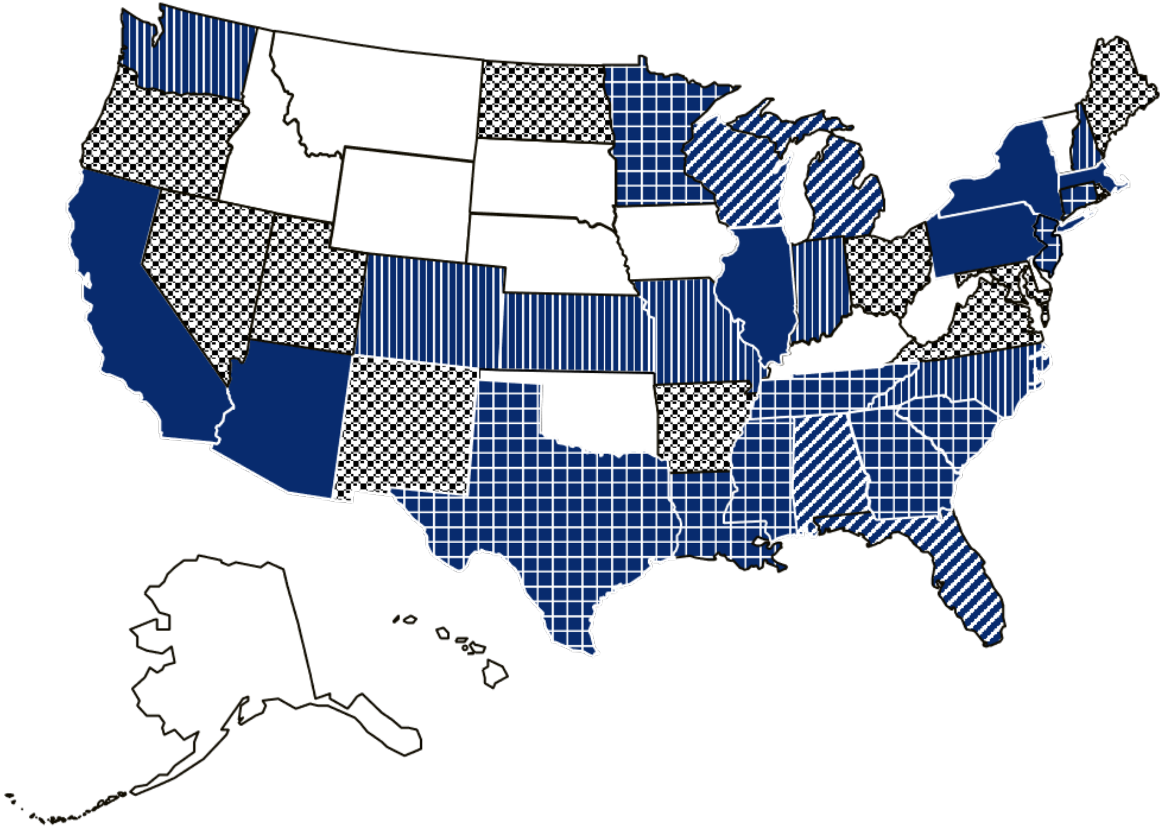
271,000+
Participants who have completed initial steps of the program

277,000+
Biosamples



Status of the Program (as of November 8, 2020)

Participant Representation Across the U.S.

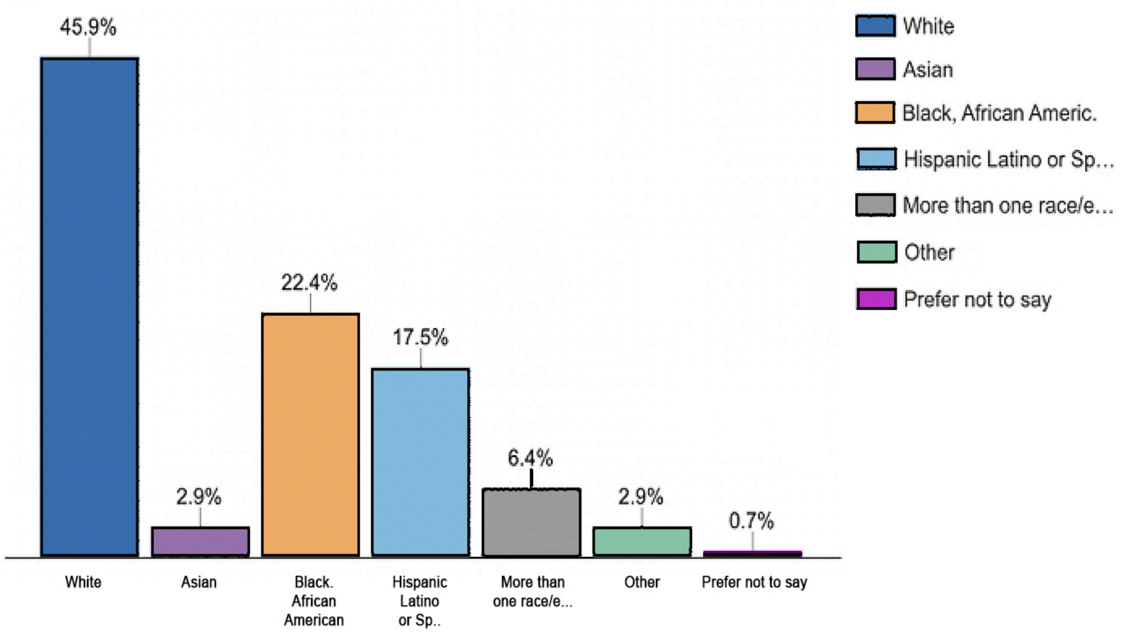


Number of Participants

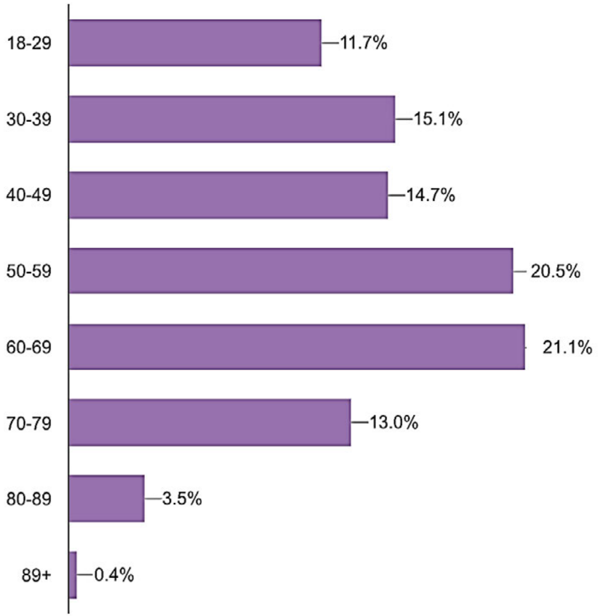
- □ **Under 100:** Alaska, Delaware, Hawaii, Idaho, Iowa, Kentucky, Montana, Nebraska, Oklahoma, South Dakota, Vermont, West Virginia, Wyoming,
- ▩ **100 to 300:** Arkansas, Maine, Maryland, Nevada, New Mexico, North Dakota, Ohio, Oregon, Rhode Island, Utah, Virginia
- ▥ **301 to 999:** Colorado, Indiana, Kansas, Missouri, New Hampshire, North Carolina, Washington
- ▦ **1,000 to 9,999:** Connecticut, Georgia, Louisiana, Minnesota, Mississippi, New Jersey, South Carolina, Tennessee, Texas
- ▨ **10,000 to 19,999:** Alabama, Florida, Michigan, Wisconsin
- ■ **More than 20,000:** Arizona, California, Illinois, Massachusetts, New York, Pennsylvania

Status of the Program Continued

Race and Ethnicity



Age



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

Enrollment, Consent, and EHR Authorization

Enrollment, Consent, EHR Authorization

Participant Surveys

Physical Measurements

Biosamples

Mobile and Wearable Technologies

Participants must be 18 years or older

Online video consent

Includes authorization to share EHR data with researchers

Plans to include children in future

Current Data Collected: Surveys, Measurements, Biosamples

Enrollment, Consent, EHR Authorization

Participant Surveys

Current and future surveys focused on:
- Lifestyle (e.g., Diet)
- Personal and Family Medical History
- Healthcare Access

Other surveys:
- COVID Participant Experience (COPE)

Physical Measurements

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

Biosamples

- Blood
- Saliva (if blood draw is unsuccessful)
- Urine

Mobile and Wearable Technologies

Future Data Collected: Mobile and Wearable Technologies

Enrollment, Consent, EHR Authorization

Participant Surveys

Physical Measurements

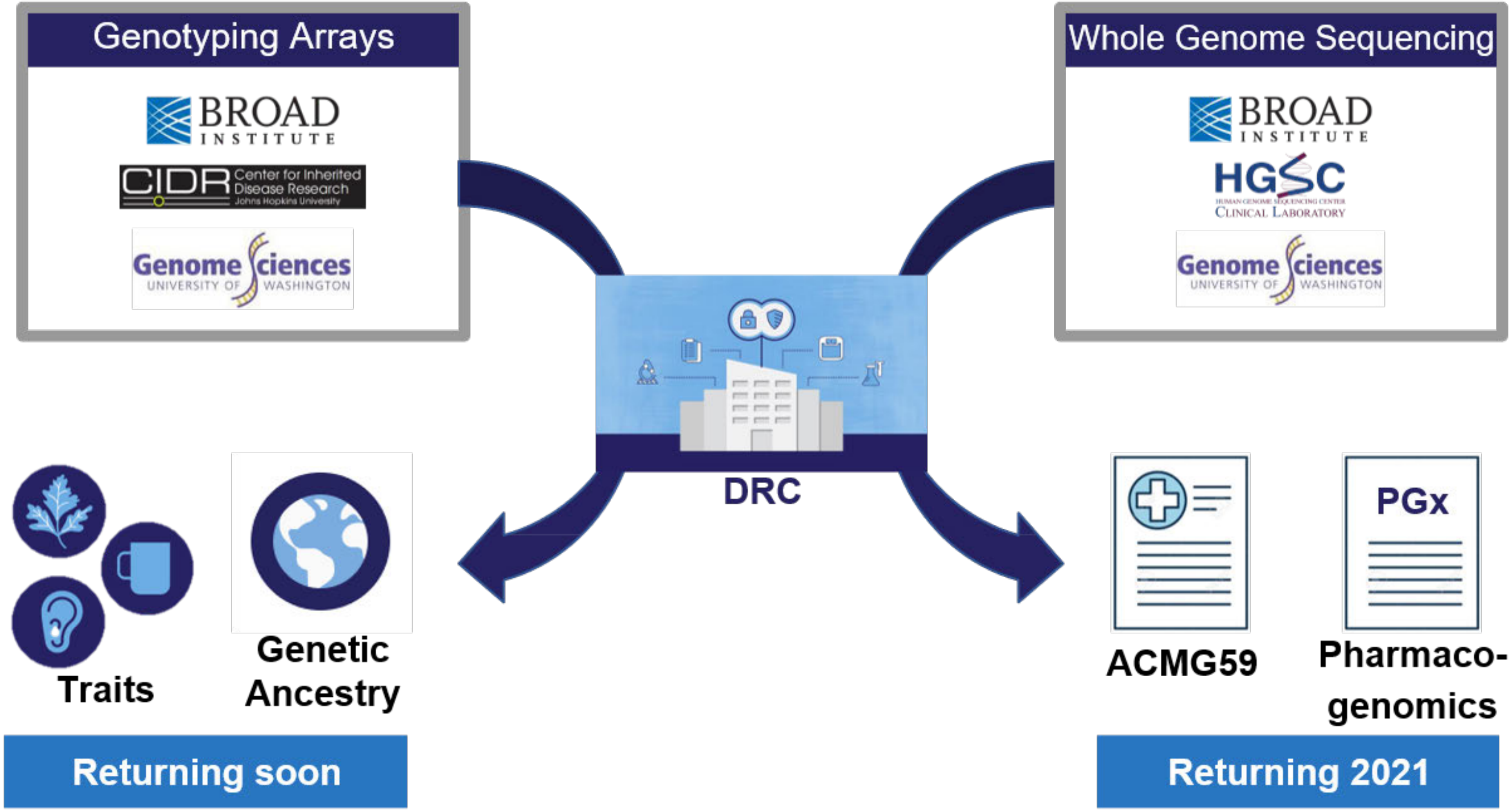
Biosamples

Mobile and Wearable Technologies

Data from wearable fitness devices, starting with Fitbit and Apple HealthKit

Other data collection from integrated apps that track mood and cardio-respiratory fitness

Returning Value to Participants: Genetic Return of Results



Genetic Ancestry and Traits

Genotyping Arrays

Traits

Genetic Ancestry

Returning soon

DRC

Health-Related Genetics

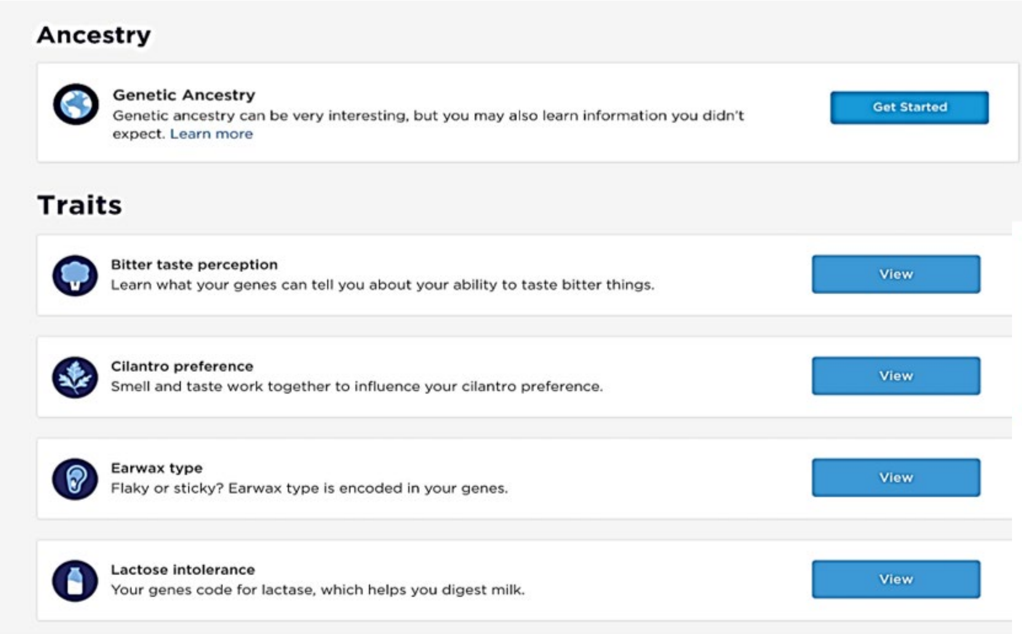
Whole Genome Sequencing

ACMG59

Pharmaco-genomics

Returning 2021

Trait Results



Ancestry

Genetic Ancestry
Genetic ancestry can be very interesting, but you may also learn information you didn't expect. Learn more

Traits

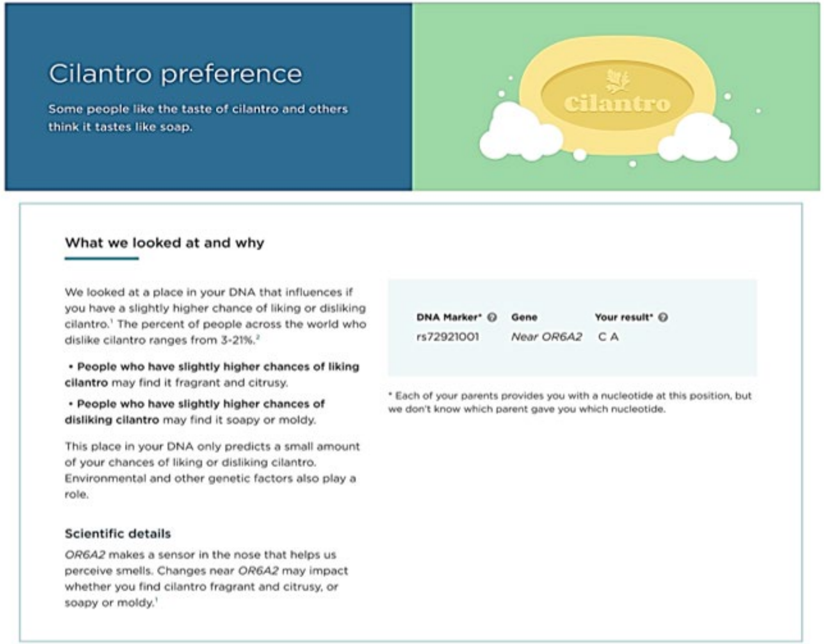
Bitter taste perception
Learn what your genes can tell you about your ability to taste bitter things.

Cilantro preference
Smell and taste work together to influence your cilantro preference.

Earwax type
Flaky or sticky? Earwax type is encoded in your genes.

Lactose intolerance
Your genes code for lactase, which helps you digest milk.

Planning to publicly launch next month



Cilantro preference

Some people like the taste of cilantro and others think it tastes like soap.

What we looked at and why

We looked at a place in your DNA that influences if you have a slightly higher chance of liking or disliking cilantro.[1] The percent of people across the world who dislike cilantro ranges from 3-21%.[2]

- **People who have slightly higher chances of liking cilantro** may find it fragrant and citrusy.
- **People who have slightly higher chances of disliking cilantro** may find it soapy or moldy.

This place in your DNA only predicts a small amount of your chances of liking or disliking cilantro. Environmental and other genetic factors also play a role.

DNA Marker*	Gene	Your result*
rs72921001	*Near OR6A2*	C A

* Each of your parents provides you with a nucleotide at this position, but we don't know which parent gave you which nucleotide.

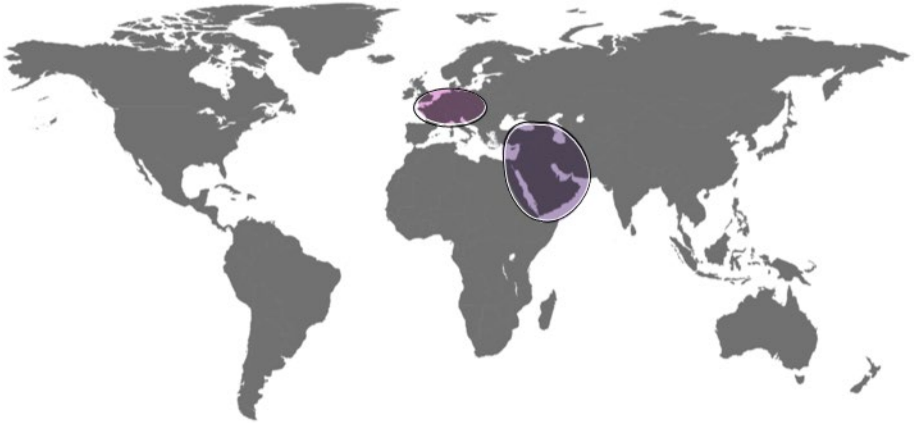
Scientific details

OR6A2 makes a sensor in the nose that helps us perceive smells. Changes near *OR6A2* may impact whether you find cilantro fragrant and citrusy, or soapy or moldy.[1]

Genetic Ancestry Results

The Middle East and North Africa	50%
The Middle East Such as the Arabian Peninsula and Egypt	**50%**
Europe	50%

All tested populations



The Middle East and North Africa

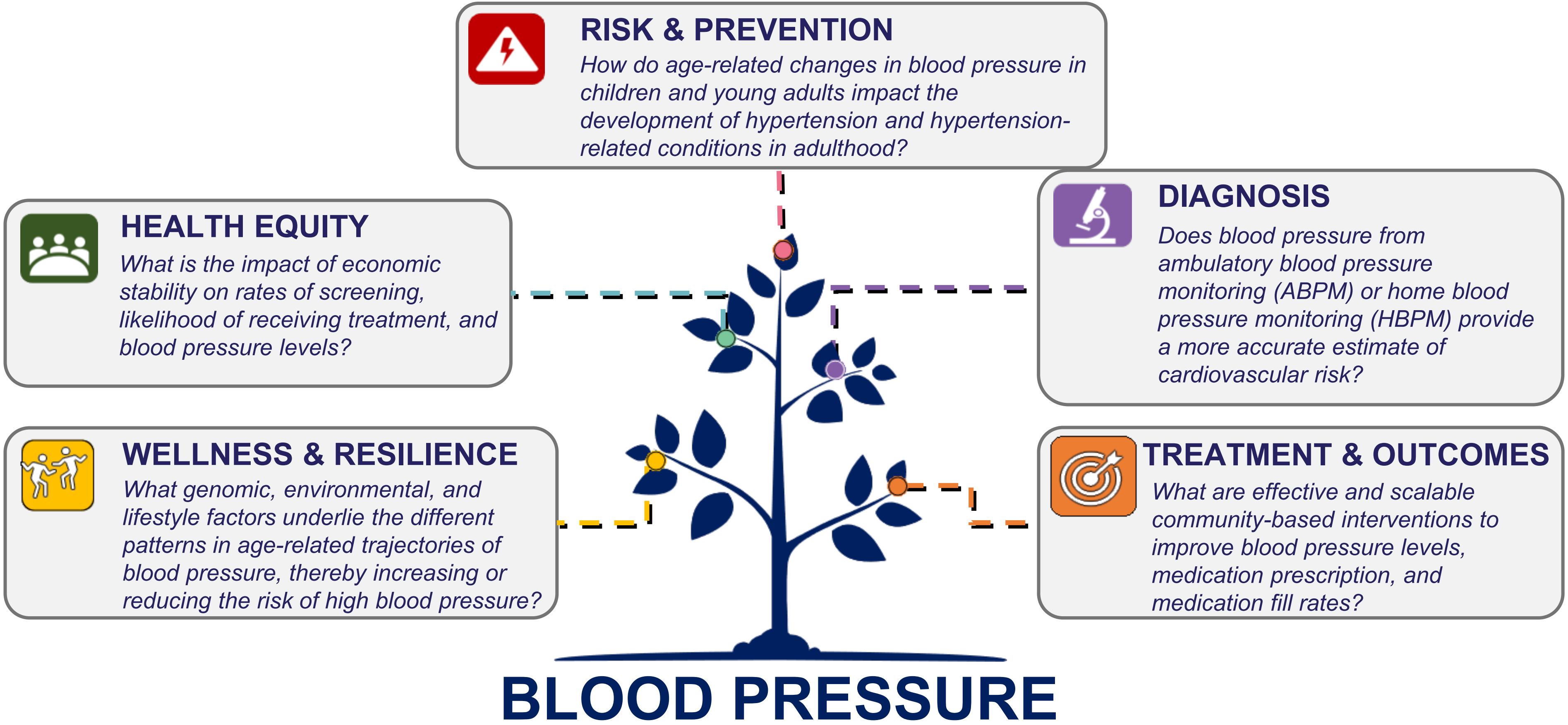
This genetic group represents people from these areas:

- The Middle East
- North Africa
- Western Africa
- The Caucasus ?

Connections near and far

People with recent ancestors from Asia, Europe, and sub-Saharan African may have patterns of DNA from this genetic ancestry group. This is likely because of significant trade and migration through the region that continues to this day. The Silk Road and Incense Route connected the Middle East and North Africa to Europe and Asia. Trans-Siberian trade routes connected North Africa to sub-Saharan Africa.

Framework Example: Blood Pressure



RISK & PREVENTION

How do age-related changes in blood pressure in children and young adults impact the development of hypertension and hypertension-related conditions in adulthood?

HEALTH EQUITY

What is the impact of economic stability on rates of screening, likelihood of receiving treatment, and blood pressure levels?

DIAGNOSIS

Does blood pressure from ambulatory blood pressure monitoring (ABPM) or home blood pressure monitoring (HBPM) provide a more accurate estimate of cardiovascular risk?

WELLNESS & RESILIENCE

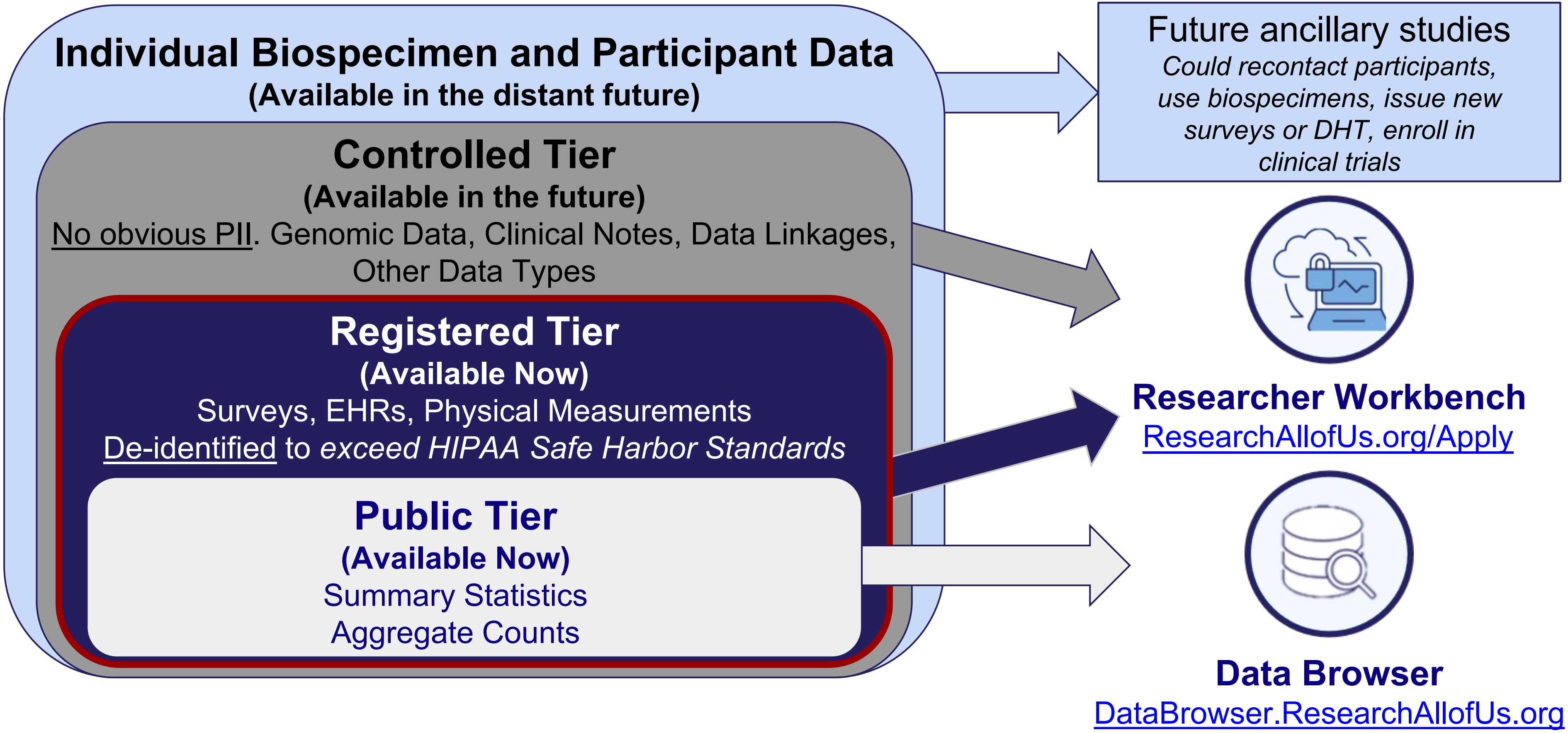
What genomic, environmental, and lifestyle factors underlie the different patterns in age-related trajectories of blood pressure, thereby increasing or reducing the risk of high blood pressure?

TREATMENT & OUTCOMES

What are effective and scalable community-based interventions to improve blood pressure levels, medication prescription, and medication fill rates?

BLOOD PRESSURE

Tiered Data and Resource Access



Individual Biospecimen and Participant Data
(Available in the distant future)

Controlled Tier
(Available in the future)
No obvious PII. Genomic Data, Clinical Notes, Data Linkages, Other Data Types

Registered Tier
(Available Now)
Surveys, EHRs, Physical Measurements
De-identified to *exceed HIPAA Safe Harbor Standards*

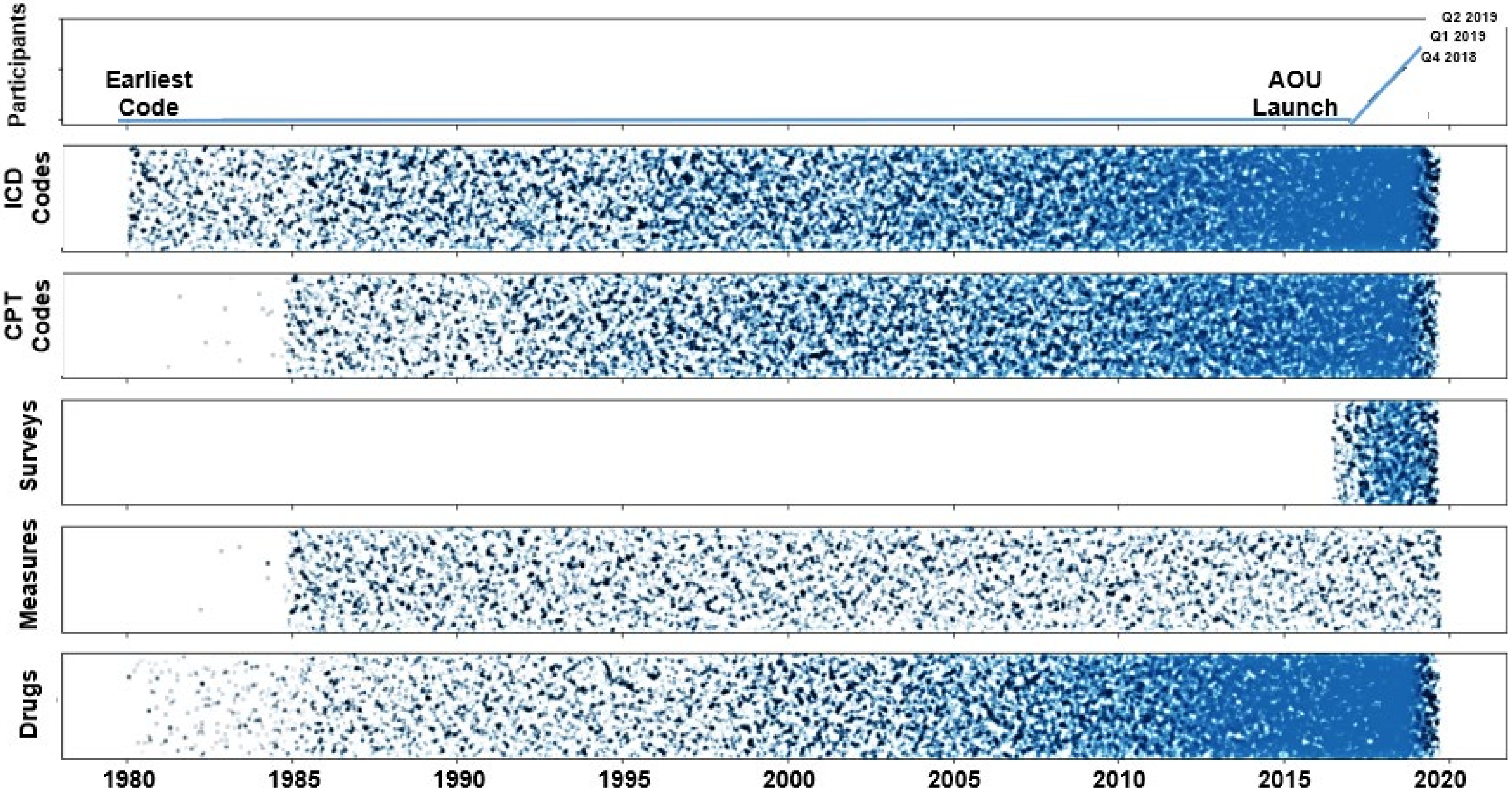
Public Tier
(Available Now)
Summary Statistics
Aggregate Counts

Future ancillary studies
Could recontact participants, use biospecimens, issue new surveys or DHT, enroll in clinical trials

Researcher Workbench
ResearchAllofUs.org/Apply

Data Browser
DataBrowser.ResearchAllofUs.org

The Curated Data Includes Longitudinal EHR (As of May 2020)

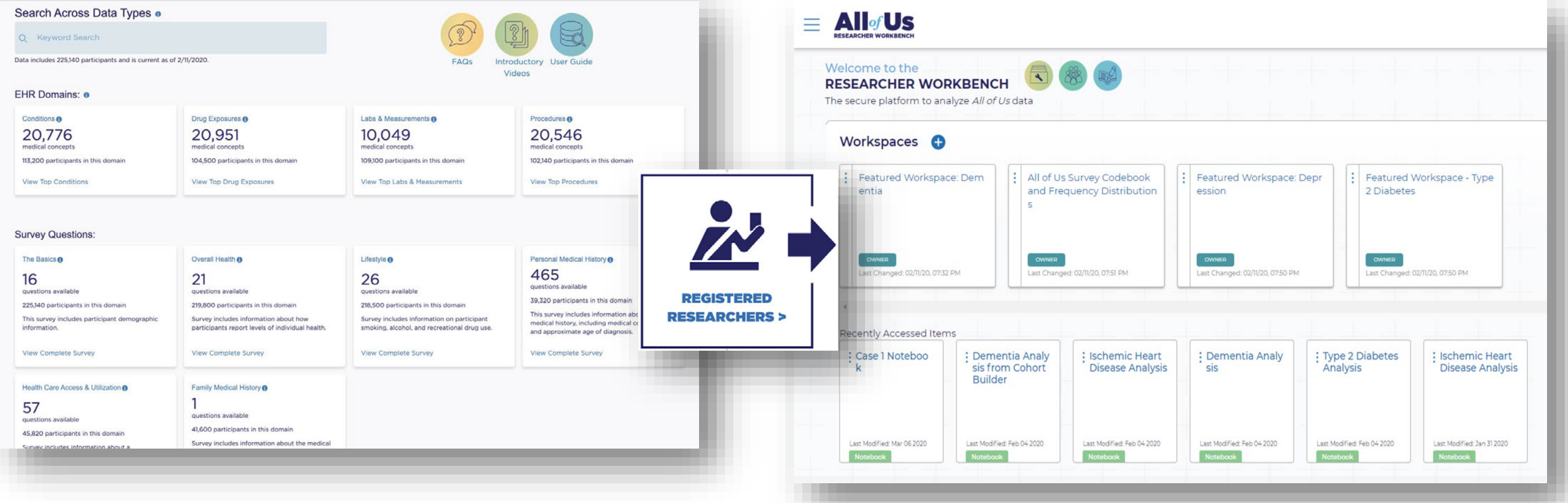


224k participants in dataset
24 million conditions
9 million procedures
>224,000 major surveys
100 million labs and measurements
18 million drug exposures

Returning Value to Researchers: Research Hub

Public: Data Browser
DataBrowser.ResearchAllofUs.org

Restricted: Researcher Workbench
ResearchAllofUs.org/Apply



Search Across Data Types

Keyword Search

Data includes 225,140 participants and is current as of 2/11/2020.

EHR Domains:

Conditions
20,776
medical concepts
113,200 participants in this domain
View Top Conditions

Drug Exposures
20,951
medical concepts
104,500 participants in this domain
View Top Drug Exposures

Labs & Measurements
10,049
medical concepts
109,100 participants in this domain
View Top Labs & Measurements

Procedures
20,546
medical concepts
102,140 participants in this domain
View Top Procedures

Survey Questions:

The Basics
16
questions available
225,140 participants in this domain
This survey includes participant demographic information.
View Complete Survey

Overall Health
21
questions available
219,800 participants in this domain
Survey includes information about how participants report levels of individual health.
View Complete Survey

Lifestyle
26
questions available
218,500 participants in this domain
Survey includes information on participant smoking, alcohol, and recreational drug use.
View Complete Survey

Personal Medical History
465
questions available
39,320 participants in this domain
This survey includes information abc medical history, including medical c and approximate age of diagnosis.
View Complete Survey

Health Care Access & Utilization
57
questions available
45,820 participants in this domain
Survey includes information about a

Family Medical History
1
questions available
41,600 participants in this domain
Survey includes information about the medical

Welcome to the
RESEARCHER WORKBENCH
The secure platform to analyze *All of Us* data

Workspaces

Featured Workspace: Dementia
OWNER
Last Changed: 02/11/20, 07:32 PM

All of Us Survey Codebook and Frequency Distributions
OWNER
Last Changed: 02/11/20, 07:51 PM

Featured Workspace: Depression
OWNER
Last Changed: 02/11/20, 07:50 PM

Featured Workspace - Type 2 Diabetes
OWNER
Last Changed: 02/11/20, 07:50 PM

Recently Accessed Items

Case 1 Notebook
Last Modified: Mar 06 2020
Notebook

Dementia Analysis from Cohort Builder
Last Modified: Feb 04 2020
Notebook

Ischemic Heart Disease Analysis
Last Modified: Feb 04 2020
Notebook

Dementia Analysis
Last Modified: Feb 04 2020
Notebook

Type 2 Diabetes Analysis
Last Modified: Feb 04 2020
Notebook

Ischemic Heart Disease Analysis
Last Modified: Jan 31 2020
Notebook

Summary statistics of participant data

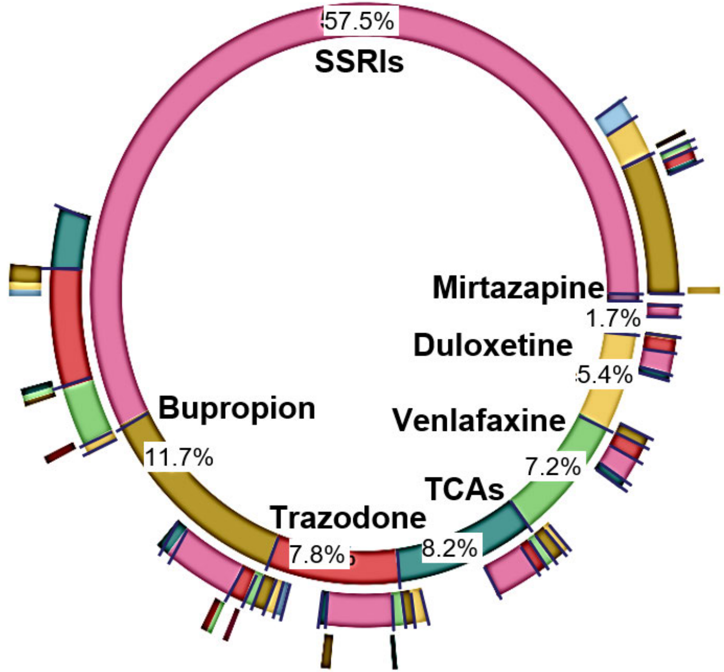
- **EHR Data**, **Survey Questions**, **Physical Measurements**

Beta Launch on May 27, 2020

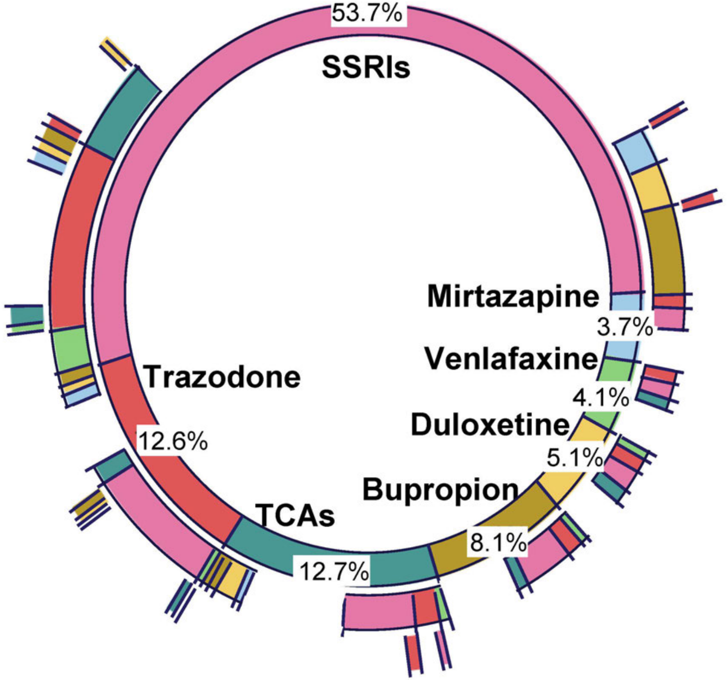
- Currently restricted to U.S. researchers with eRA Commons accounts

Demonstration Project: Depression Medication Sequencing

White (n=9,235)



Non-White (n=6,541)



Response to COVID-19

Coronavirus Disease 2019

CDC is responding to the novel coronavirus outbreak.

How can *All of Us* further our understanding of COVID-19?

Future Researcher Workbench Data: COVID-19

COVID Participant Experience (COPE Survey)

Gather data about participants' experiences with the pandemic. Repeated questionnaire released May–July. Shorter questionnaire released in November.

Seroprevalence Survey

Check for COVID19 antibodies in samples collected **January to March 2020** to better understand COVID-19's arrival in the U.S.

EHR data from COVID-positive participants

Collect and standardize COVID-19 EHR data.

Current and Future Researcher Workbench Data

Current Data

- > 224,000 Participants
- > 224,000 w/ Survey
- > 188,000 w/ Physical Measurement
- > 127,000 w/ EHR

Late 2020/Early 2021 Data Release

- > 315,000 participants
- 40% more Surveys
- 40% more Physical Measurements
- 60% more EHRs
- Fitbit data
- Some COPE Survey data
- Unlinked COVID Serology data

Late 2021/Early 2022 Data Release

- First Controlled Tier release
- More participants
- All COPE Survey data
- COVID EHR data
- Linked COVID Serology data
- Genomics

All of Us Research Hub

Kelsey Mayo, Ph.D.
Mgr, Scientific Product and Portfolio Manager
All of Us Data and Research Center

Thank you to our participants and community partners!

All of Us Community and Provider Partner Network *(as of July 2020)*

AAFP
AAPA
American Academy of Physician Assistants
AAHIVM
American Academy of HIV Medicine
ANAC
Association of Nurses in AIDS Care
American Association of Colleges of Nursing
American Association on Health and Disability
ACCESS
APHA
American Public Health Association
Greensboro AHEC
Area Health Education Center
Texas AHEC East
Waco Region
Central Colorado AHEC
National Pan-Hellenic Council
Black Women's Health Imperative
CenterLink
The Community of LGBT Centers
W. Montague Cobb / NMA Health Institute
Community Council
Advancing Solutions...Empowering Lives
HACU
LULAC
NACCHO
National Association of County & City Health Officials
NAHN
National Association of Hispanic Nurses
National Baptist Convention, U.S.A. Incorporated
National Black Nurses Association, Inc.
NBNA
NHCOA
National Hispanic Council on Aging
NHMA
National Hispanic Medical Association
NRHA
National Rural Health Association
UnidosUS
Stronger Communities. Stronger America.
National Minority Quality Forum
the Y
YMCA
NIH U.S. National Library of Medicine
Asian Health Coalition
Delta Research and Educational Foundation
Creating Possibilities... Improving Lives
Fifty Forward
National Alliance for Hispanic Health
Stanford Medicine

Thank you to our consortium partners!

The Participant Center

America's Blood Centers
The Behavioural Insights Team
careevolution Healthcare Technology
DXC.technology
Medscape
Montage Marketing
waves
patientslikeme
QTC A Leidos Company
Sage Bionetworks
Sync for Science
Quest Diagnostics
Scripps Research Translational Institute
Urban One
Walgreens
BlueCross BlueShield
fitbit
National Blood Collaborative
PicnicHealth
San Diego Blood Bank Foundation
SENSIS
WebMD

Communications & Engagement

WONDROS
The University of Utah
PYXIS Growth Partners

HPO Network

(Health Care Provider Organizations)

RMCs

All of Us California

UC San Diego Health
UCI Health
UCSF University of California San Francisco
UC Davis Health
Cedars-Sinai
USC University of Southern California

Illinois Precision Medicine Consortium

Northwestern Medicine
The University of Chicago
NorthShore University HealthSystem
Rush
Cook County Health
The University of Illinois College of Medicine
UI Health UIC

All of Us New England

Brigham Health Brigham and Women's Hospital
Partners HealthCare
Massachusetts General Hospital
Boston Medical Center

Trans America Consortium

Henry Ford Health System
Essentia Health
Baylor Scott & White Health
Spectrum Health
Reliant Medical Group Part of OptumCare
HealthPartners

New York City Consortium

Columbia University Medical Center
NYC Health + Hospitals
NewYork-Presbyterian
Weill Cornell Medicine

All of Us Southern Network

UAB School of Medicine The University of Alabama at Birmingham Selma Family Medicine
UAB School of Medicine The University of Alabama at Birmingham Montgomery Regional Medical Campus
The University of Mississippi Medical Center
UAB Medicine
Tulane University
LSU Health New Orleans
The University of Alabama University Medical Center
UAB School of Medicine The University of Alabama at Birmingham Huntsville Regional Medical Campus
Cooper Green Mercy Health Services
USA Health
Huntsville Hospital

All of Us SouthEast Enrollment Center

University of Miami Miller School of Medicine
Emory University
Morehouse School of Medicine
UF University of Florida
Grady

All of Us Wisconsin

Marshfield Clinic Health System
Gundersen Health System
Wisconsin
Froedtert & Medical College of Wisconsin

All of Us Pennsylvania

University of Pittsburgh
Temple University

University of Arizona and Banner Health

The University of Arizona
Banner Health

FQHCs (Federally Qualified Health Centers)

Cherokee Health Systems
Jackson-Hinds Comprehensive Health Center
San Ysidro Health
Community Health Center, Inc.
Cooperative Health
HRHCare

VA Medical Centers

VA U.S. Department of Veterans Affairs

Participant Technology Systems Center (PTSC)

vibrent

Biobank

Mayo Clinic

Data & Research Center (DRC)

Vanderbilt University Medical Center
Broad Institute
verily
Columbia University Medical Center
Michigan School of Public Health
Northwestern Medicine
UTHealth The University of Texas Health Science Center at Houston

Genomics Partners

Baylor College of Medicine
Johns Hopkins Medicine
UTHealth The University of Texas Health Science Center at Houston
HudsonAlpha Institute for Biotechnology
Broad Institute
color
Partners HealthCare
University of Washington

Thank you Data and Research Center (DRC) and NIH teams!

All of Us
DATA & RESEARCH CENTER

BROAD INSTITUTE

COLUMBIA UNIVERSITY MEDICAL CENTER

verily

NIH National Institutes of Health
All of Us Research Program

VANDERBILT UNIVERSITY MEDICAL CENTER

All of Us Research Program Objectives

Nurture relationships
with **one million or more** participant partners, from all walks of life, for decades

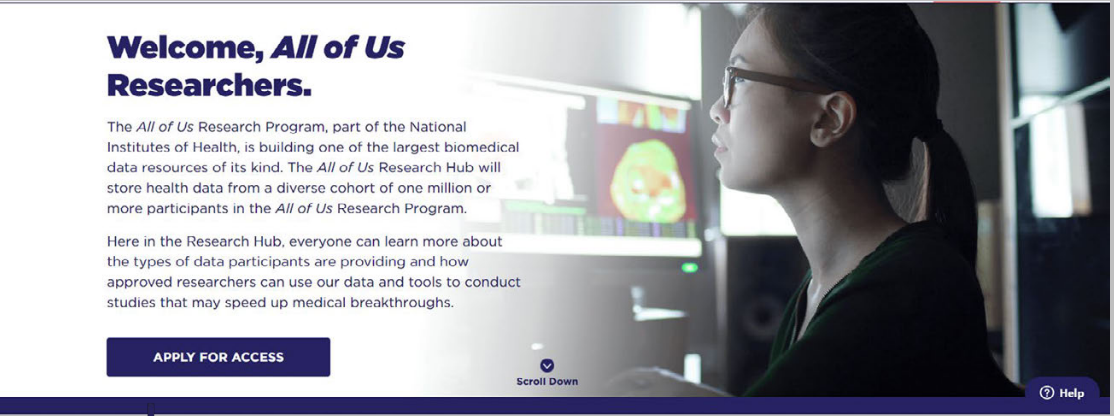
Our mission
To accelerate health research and medical breakthroughs, enabling individualized prevention, treatment, and care for all of us

Deliver the largest, richest biomedical dataset
that is easy, safe, and free to access

Catalyze a robust ecosystem
of researchers and funders hungry to use and support it

All of Us
DATA & RESEARCH CENTER

All of Us Research Hub



Welcome, *All of Us* Researchers.

The *All of Us* Research Program, part of the National Institutes of Health, is building one of the largest biomedical data resources of its kind. The *All of Us* Research Hub will store health data from a diverse cohort of one million or more participants in the *All of Us* Research Program.

Here in the Research Hub, everyone can learn more about the types of data participants are providing and how approved researchers can use our data and tools to conduct studies that may speed up medical breakthroughs.

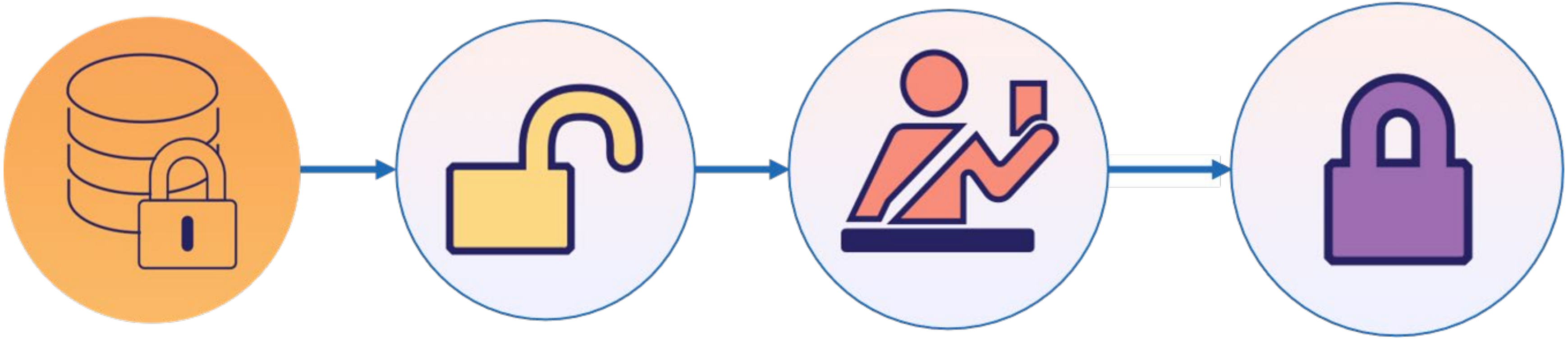
All of Us

RESEARCHER WORKBENCH

Launched Beta in May 2020!

Outline for Today

1. Data Overview



2. Live demo of researchallofus.org & Researcher Workbench

Outline for Today

1. Data Overview

Our Data are Growing. Here are the Current Data Types

Electronic Health Records (EHR)

Participant Surveys

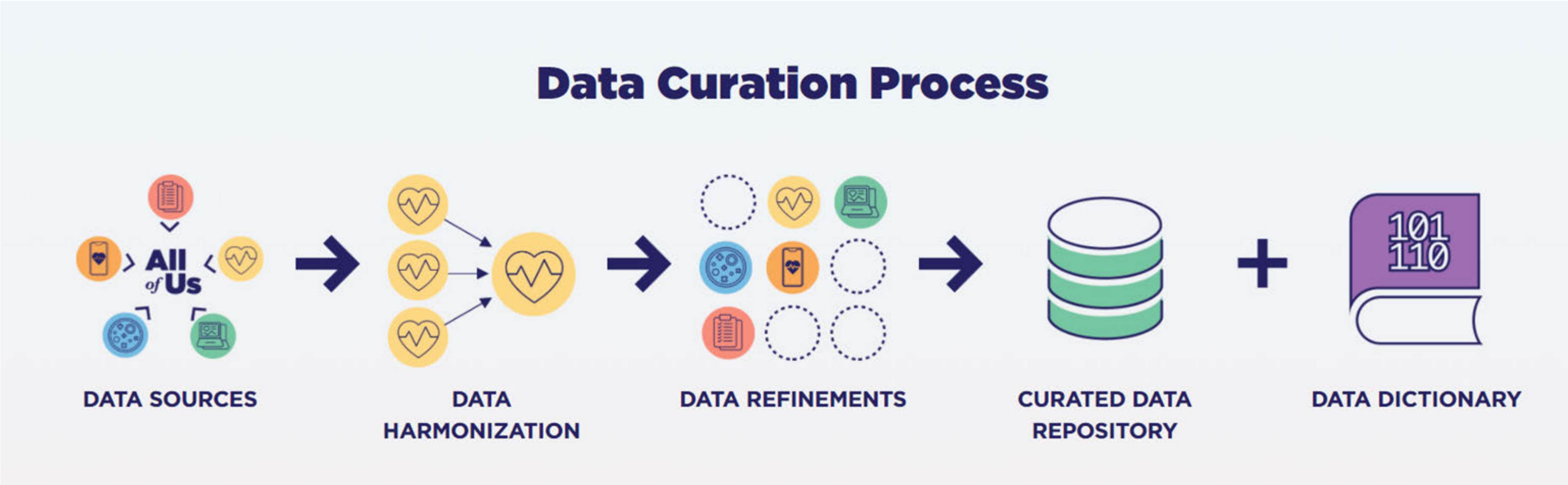
Physical Measurements

Available in the Current Research Dataset

Biosamples

Mobile and Wearable Technologies

Data are harmonized and standardized to a common data model



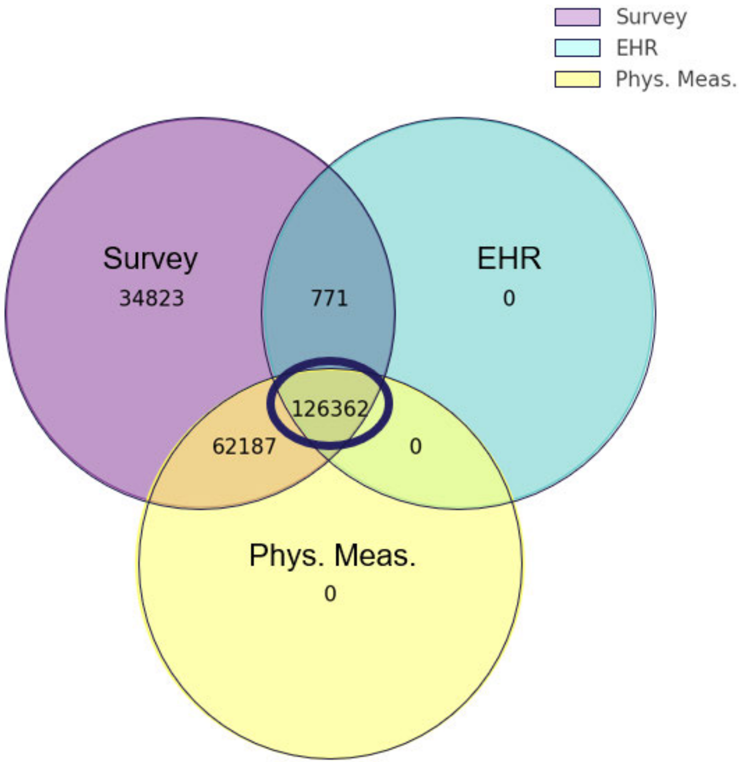
We use the OHDSI collaborative's OMOP Common Data Model

Research Data Available Now

Data Type	Participant Count
Survey	>224,000
Physical Measurement	>188,000
Electronic Health Record	>127,000

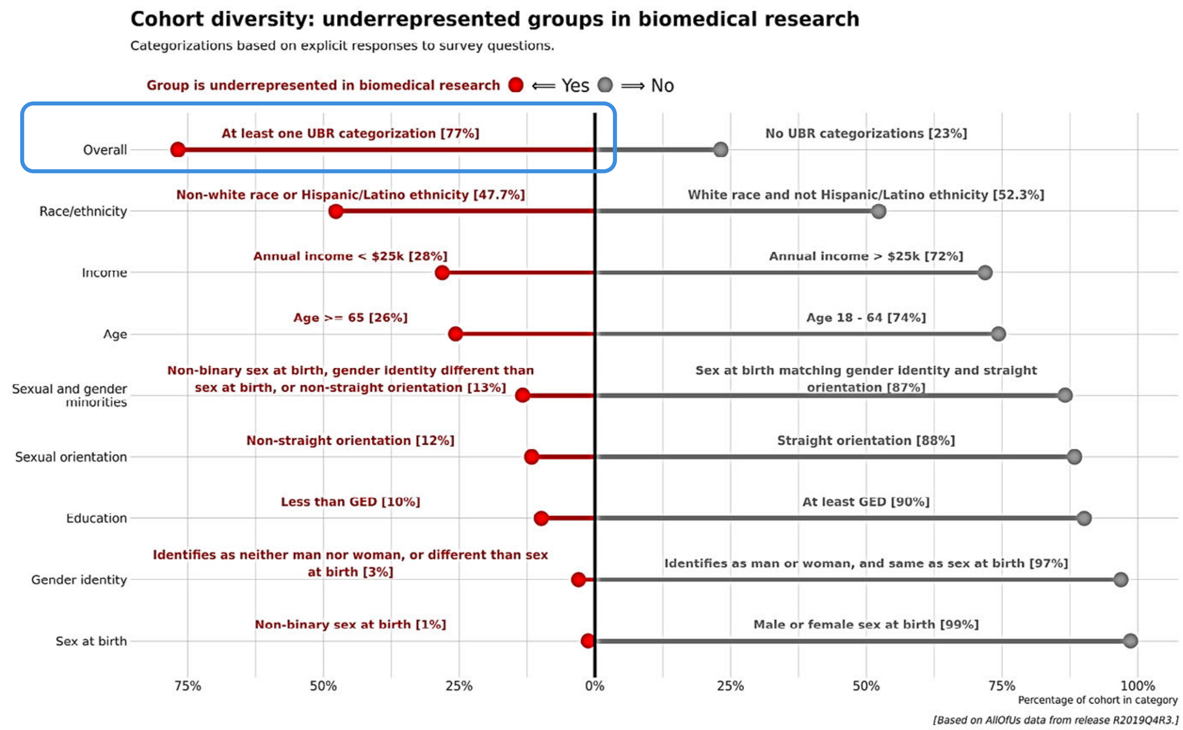
*Counts reflect unique participants with ANY of data of the specified type.

Count of participants with multiple data types

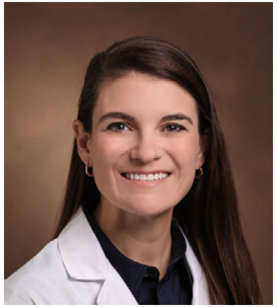


Participants included in this dataset are diverse.

All of Us
DATA & RESEARCH CENTER



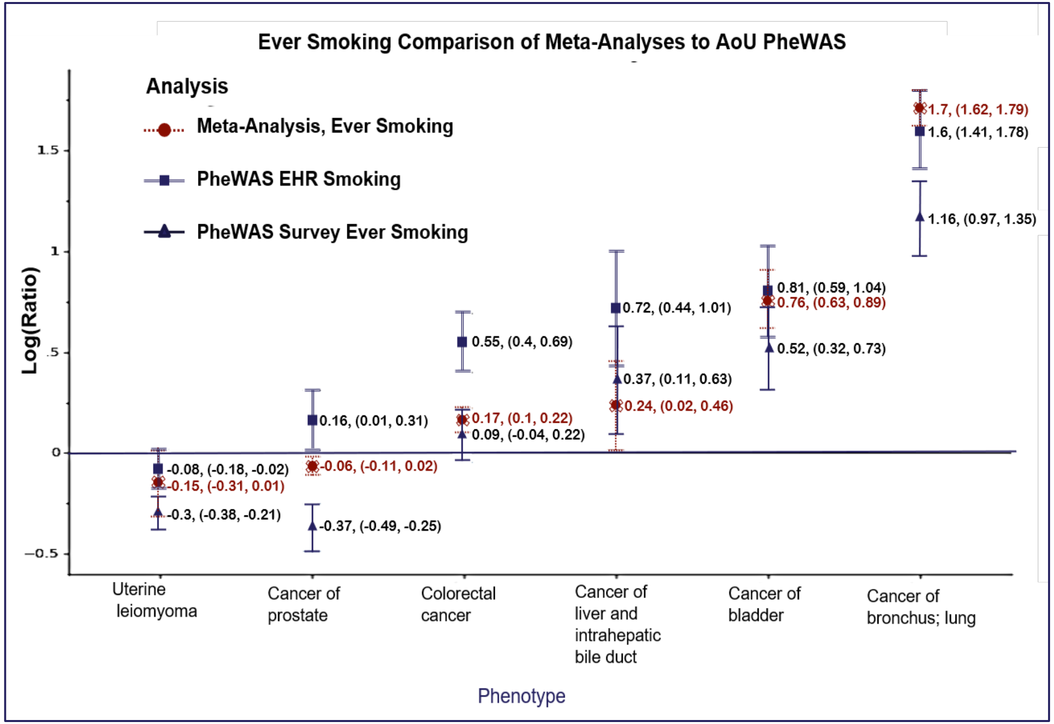
[Based on AllOfUs data from release R2019Q4R3.]



A. Ramirez, L. Suleiman, D. Schlueter, et al., *The All of Us Research Program: data quality, utility, and diversity*, medRxiv 2020.05.29.20116905; doi: https://doi.org/10.1101/2020.05.29.20116905

This dataset reproduces known associations.

Description	EHR Ever Smoking OR (95% CI)	Survey Ever Smoking OR (95% CI)
Top 3 Increased risk effects		
Cancer of the bronchus; lung	4.94 (4.11, 5.95)	3.19 (2.65, 3.84)
Cancer within the respiratory system	4.94 (4.12, 5.92)	3.15 (2.62, 3.78)
Malignant neoplasm of bladder	2.36 (1.87, 2.98)	1.76 (1.42, 2.18)
Top 3 Decreased risk effects		
Vascular hamartomas and non-neoplastic nevi	0.51 (0.42, 0.62)	0.55 (0.48, 0.64)
Nevus, non-neoplastic	0.52 (0.43, 0.64)	0.57 (0.49, 0.66)
Benign neoplasm of skin	0.53 (0.49, 0.58)	0.62 (0.58, 0.66)



A. Ramirez, L. Suleiman, D. Schlueter, et al., *The All of Us Research Program: data quality, utility, and diversity*, medRxiv 2020.05.29.20116905; doi: https://doi.org/10.1101/2020.05.29.20116905

Want to learn more?

medRxiv
THE PREPRINT SERVER FOR HEALTH SCIENCES

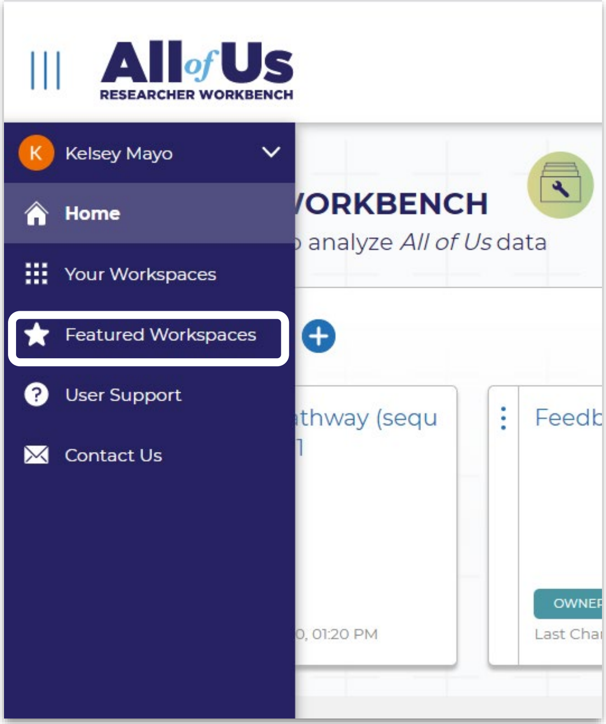
The All of Us Research Program: data quality, utility, and diversity

Andrea H Ramirez, Lina Sulieman, David J Schlueter, Alese Halvorson, Jun Qian, Francis Ratsimbazafy, Roxana Loperena, Kelsey Mayo, Melissa Basford, Nicole Deflaux, Karthik N Muthuraman, Karthik Natarajan, Abel Kho, Hua Xu, Consuelo Wilkins, Hoda Anton-Culver, Eric Boerwinkle, Mine Cicek, Cheryl R Clark, Elizabeth Cohn, Lucila Ohno-Machado, Sheri Schully, Brian K Ahmedani, Maria Argos, Robert M Cronin, Christopher O'Donnell, Mona Fouad, David B Goldstein, Philip Greenland, Scott J Hebbring, Elizabeth W Karlson, Parinda Khatri, Bruce Korf, Jordan W Smoller, Stephen Sodeke, John Wilbanks, Justin Hentges, Christopher Lunt, Stephanie A Devaney, Kelly Gebo, Joshua C Denny, Robert J Carroll, David Glazer, Paul A Harris, George Hripcsak, Anthony Philippakis, Dan M Roden, On behalf of the All of Us Research Program

doi: https://doi.org/10.1101/2020.05.29.20116905

This article is a preprint and has not been peer-reviewed [what does this mean?]. It reports new medical research that has yet to be evaluated and so should *not* be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.29.20116905.this version posted June 3, 2020. The copyright holder for this preprint



A. Ramirez, L. Suleiman, D. Schlueter, et al., *The All of Us Research Program: data quality, utility, and diversity*, medRxiv 2020.05.29.20116905; doi: https://doi.org/10.1101/2020.05.29.20116905

Coming Soon (Late 2020)...

Electronic Health Records (EHR)

Participant Surveys

Physical Measurements

COVID Participant Experience (COPE Survey)

Mobile and Wearable Technologies (Fitbit)

Refreshed + New Data Types

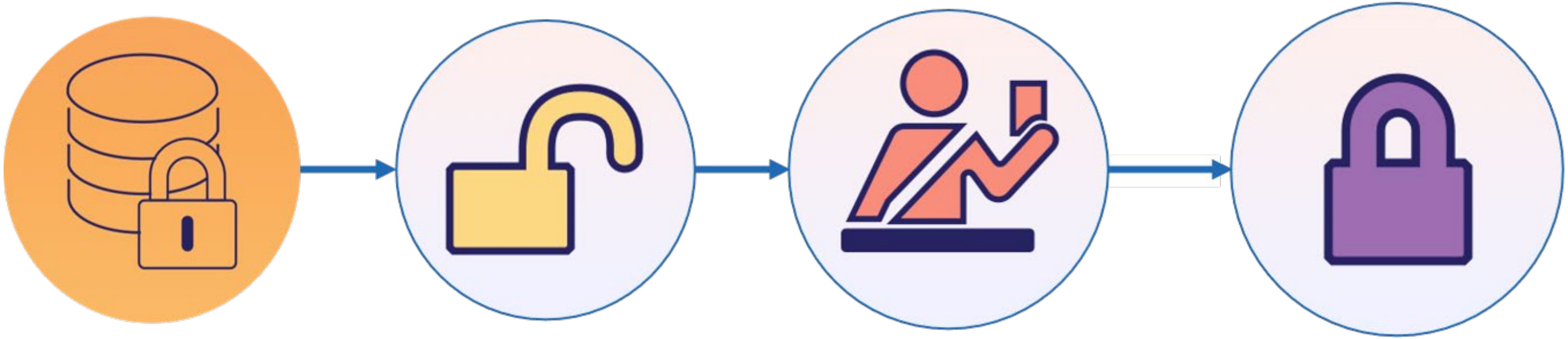
Research Data Available Soon (Winter 2020)

Data Type	Participant Count*	
Survey	>315,000	**~40% increase**
COPE Survey	>63,000	**New!**
Physical Measurement	>263,000	**~40% increase**
Electronic Health Record	>203,000	**~60% increase!**
FitBit	>8,000	**New!**

*Counts reflect unique participants with ANY of data of the specified
Counts are approximate and subject to change upon final release of the dataset.

Outline for Today

1. Data Overview



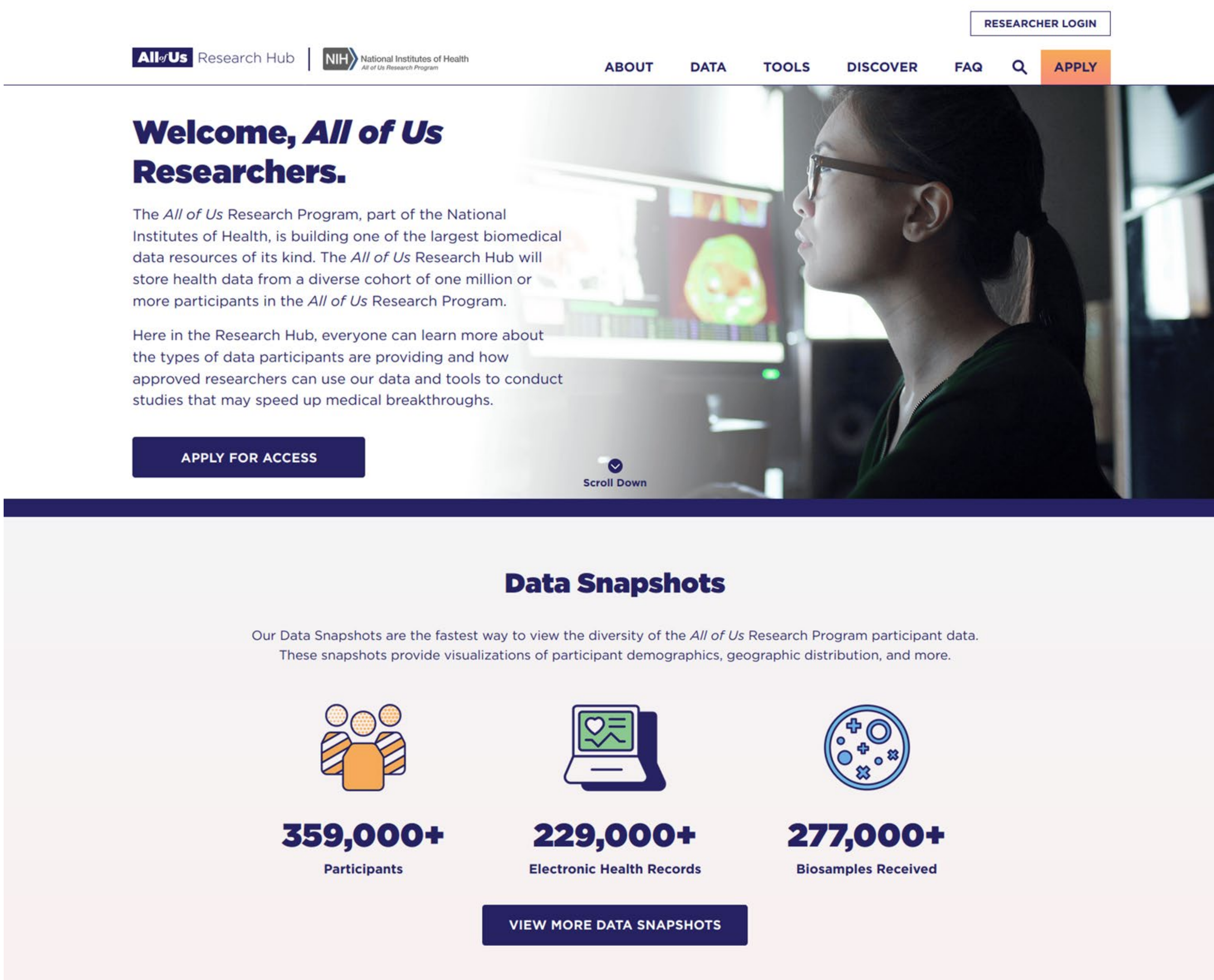
2. Live demo of researchallofus.org & Researcher Workbench

Research Hub

- Data Browser
- Apply Page

Researcher Workbench

- Support Hub
- Ex. Workspace



Welcome, *All of Us* Researchers.

The *All of Us* Research Program, part of the National Institutes of Health, is building one of the largest biomedical data resources of its kind. The *All of Us* Research Hub will store health data from a diverse cohort of one million or more participants in the *All of Us* Research Program.

Here in the Research Hub, everyone can learn more about the types of data participants are providing and how approved researchers can use our data and tools to conduct studies that may speed up medical breakthroughs.

Data Snapshots

Our Data Snapshots are the fastest way to view the diversity of the *All of Us* Research Program participant data. These snapshots provide visualizations of participant demographics, geographic distribution, and more.

359,000+	229,000+	277,000+
Participants	Electronic Health Records	Biosamples Received

Research Hub

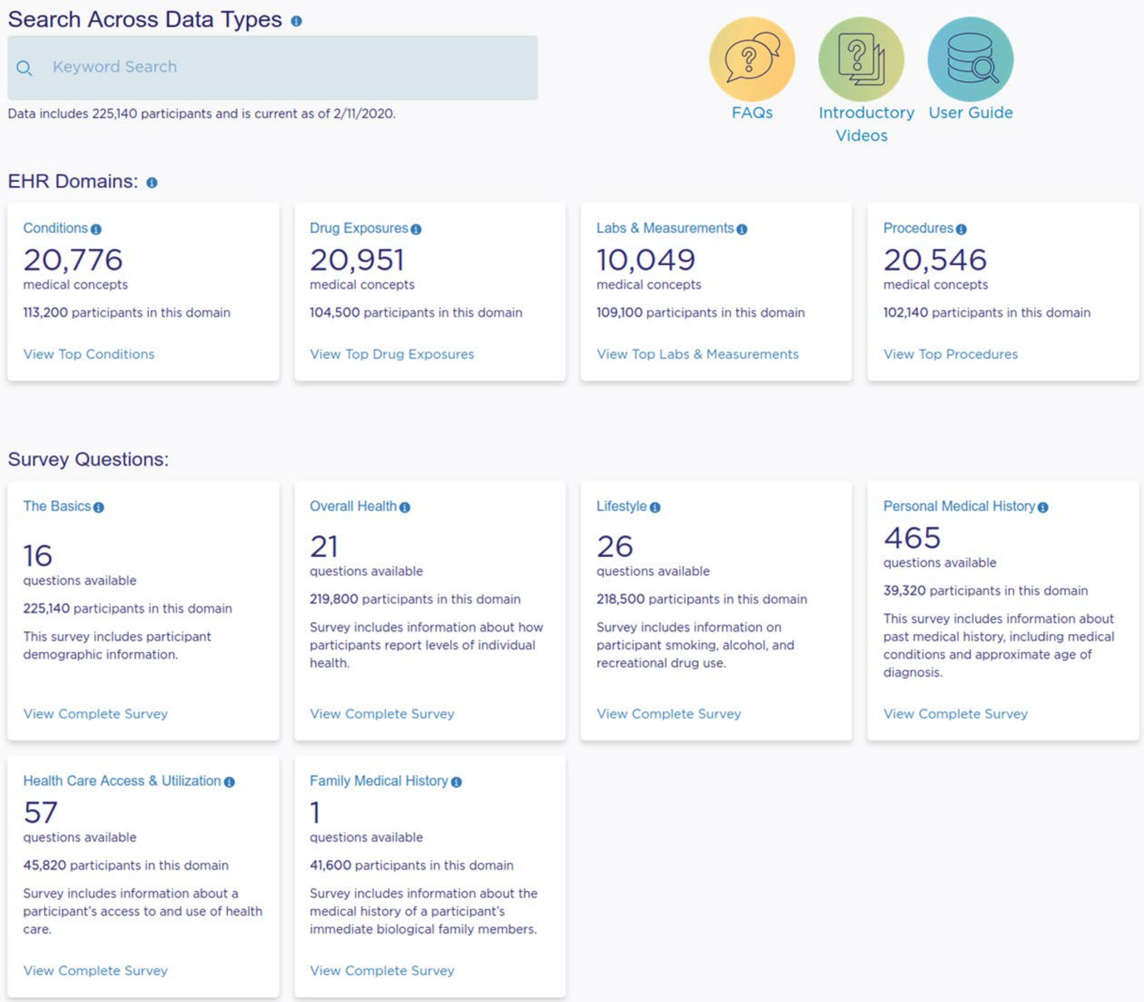
Data Browser

Apply Page

Researcher Workbench

Support Hub

Ex. Workspace



Search Across Data Types

Data includes 225,140 participants and is current as of 2/11/2020.

EHR Domains:

Conditions
20,776
medical concepts
113,200 participants in this domain
View Top Conditions

Drug Exposures
20,951
medical concepts
104,500 participants in this domain
View Top Drug Exposures

Labs & Measurements
10,049
medical concepts
109,100 participants in this domain
View Top Labs & Measurements

Procedures
20,546
medical concepts
102,140 participants in this domain
View Top Procedures

Survey Questions:

The Basics
16
questions available
225,140 participants in this domain
This survey includes participant demographic information.
View Complete Survey

Overall Health
21
questions available
219,800 participants in this domain
Survey includes information about how participants report levels of individual health.
View Complete Survey

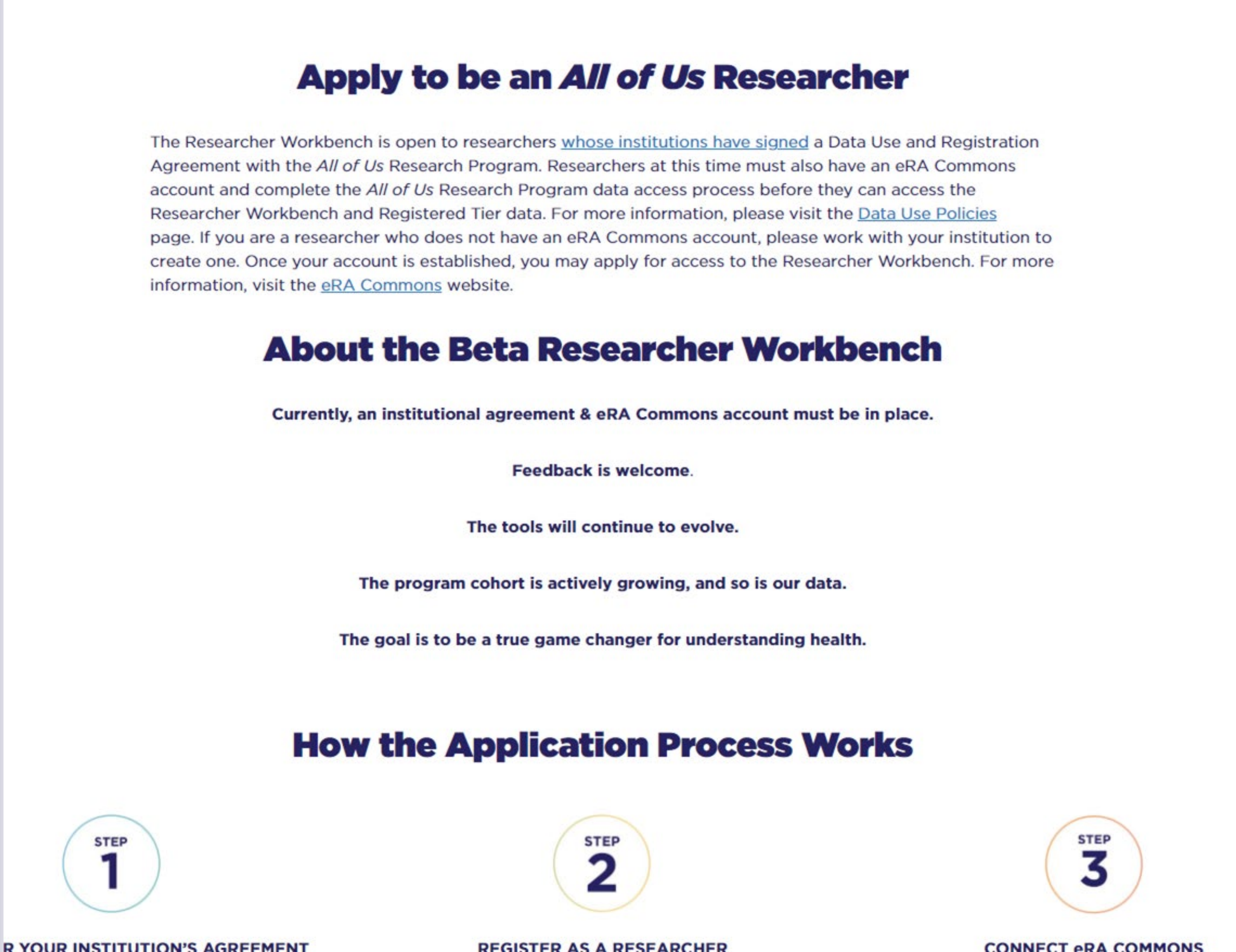
Lifestyle
26
questions available
218,500 participants in this domain
Survey includes information on participant smoking, alcohol, and recreational drug use.
View Complete Survey

Personal Medical History
465
questions available
39,320 participants in this domain
This survey includes information about past medical history, including medical conditions and approximate age of diagnosis.
View Complete Survey

Health Care Access & Utilization
57
questions available
45,820 participants in this domain
Survey includes information about a participant's access to and use of health care.
View Complete Survey

Family Medical History
1
questions available
41,600 participants in this domain
Survey includes information about the medical history of a participant's immediate biological family members.
View Complete Survey

- Research Hub
  - Data Browser
  - **Apply Page**
- Researcher Workbench
  - Support Hub
  - Ex. Workspace



Apply to be an *All of Us* Researcher

The Researcher Workbench is open to researchers whose institutions have signed a Data Use and Registration Agreement with the *All of Us* Research Program. Researchers at this time must also have an eRA Commons account and complete the *All of Us* Research Program data access process before they can access the Researcher Workbench and Registered Tier data. For more information, please visit the Data Use Policies page. If you are a researcher who does not have an eRA Commons account, please work with your institution to create one. Once your account is established, you may apply for access to the Researcher Workbench. For more information, visit the eRA Commons website.

About the Beta Researcher Workbench

Currently, an institutional agreement & eRA Commons account must be in place.

Feedback is welcome.

The tools will continue to evolve.

The program cohort is actively growing, and so is our data.

The goal is to be a true game changer for understanding health.

How the Application Process Works

Research Hub

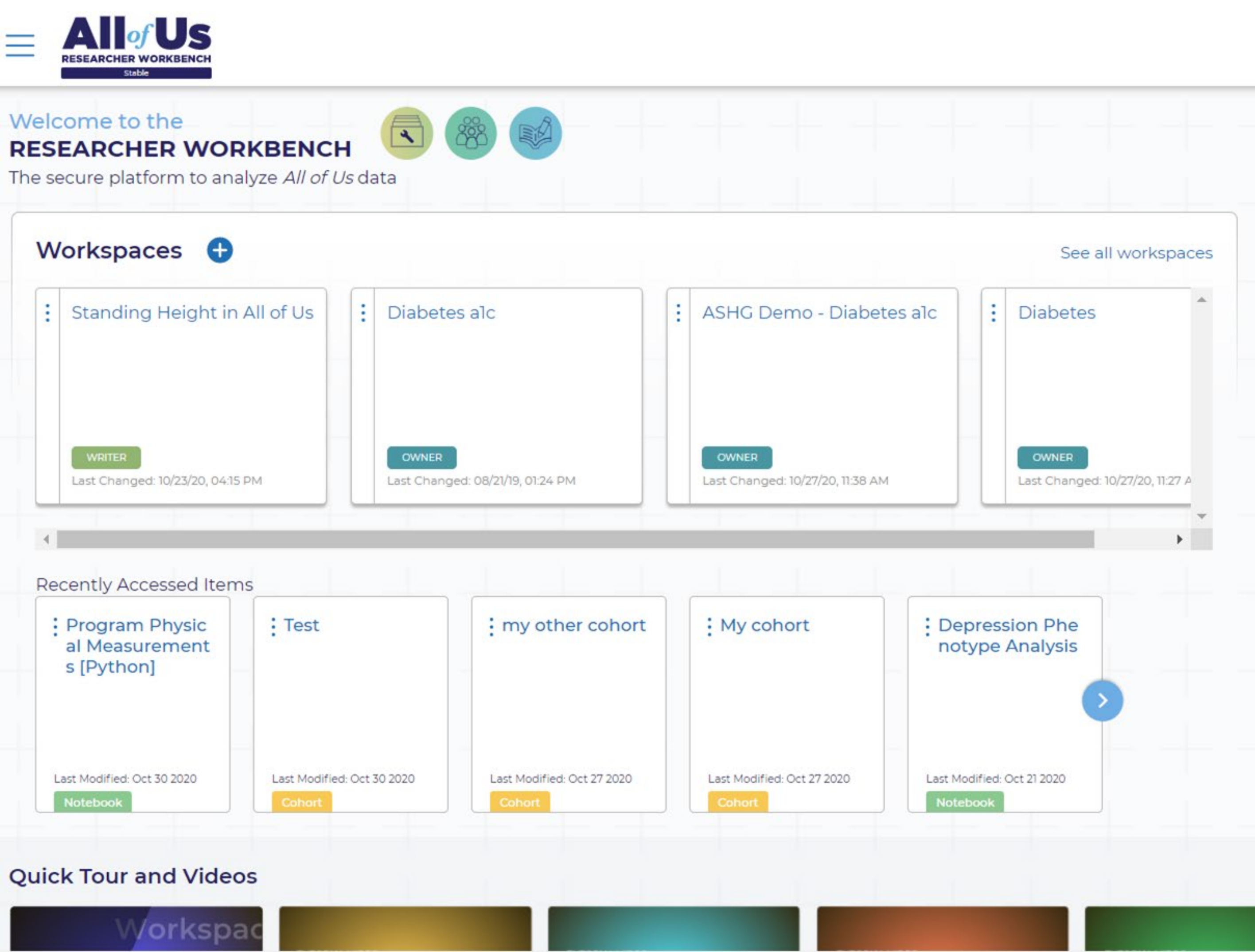
Data Browser

Apply Page

Researcher Workbench

Support Hub

Ex. Workspace



Welcome to the
RESEARCHER WORKBENCH
The secure platform to analyze *All of Us* data

Workspaces

See all workspaces

Workspace	Role	Last Changed
Standing Height in All of Us	WRITER	10/23/20, 04:15 PM
Diabetes a1c	OWNER	08/21/19, 01:24 PM
ASHG Demo - Diabetes a1c	OWNER	10/27/20, 11:38 AM
Diabetes	OWNER	10/27/20, 11:27 A

Recently Accessed Items

Item	Last Modified	Type
Program Physical Measurements [Python]	Oct 30 2020	Notebook
Test	Oct 30 2020	Cohort
my other cohort	Oct 27 2020	Cohort
My cohort	Oct 27 2020	Cohort
Depression Phenotype Analysis	Oct 21 2020	Notebook

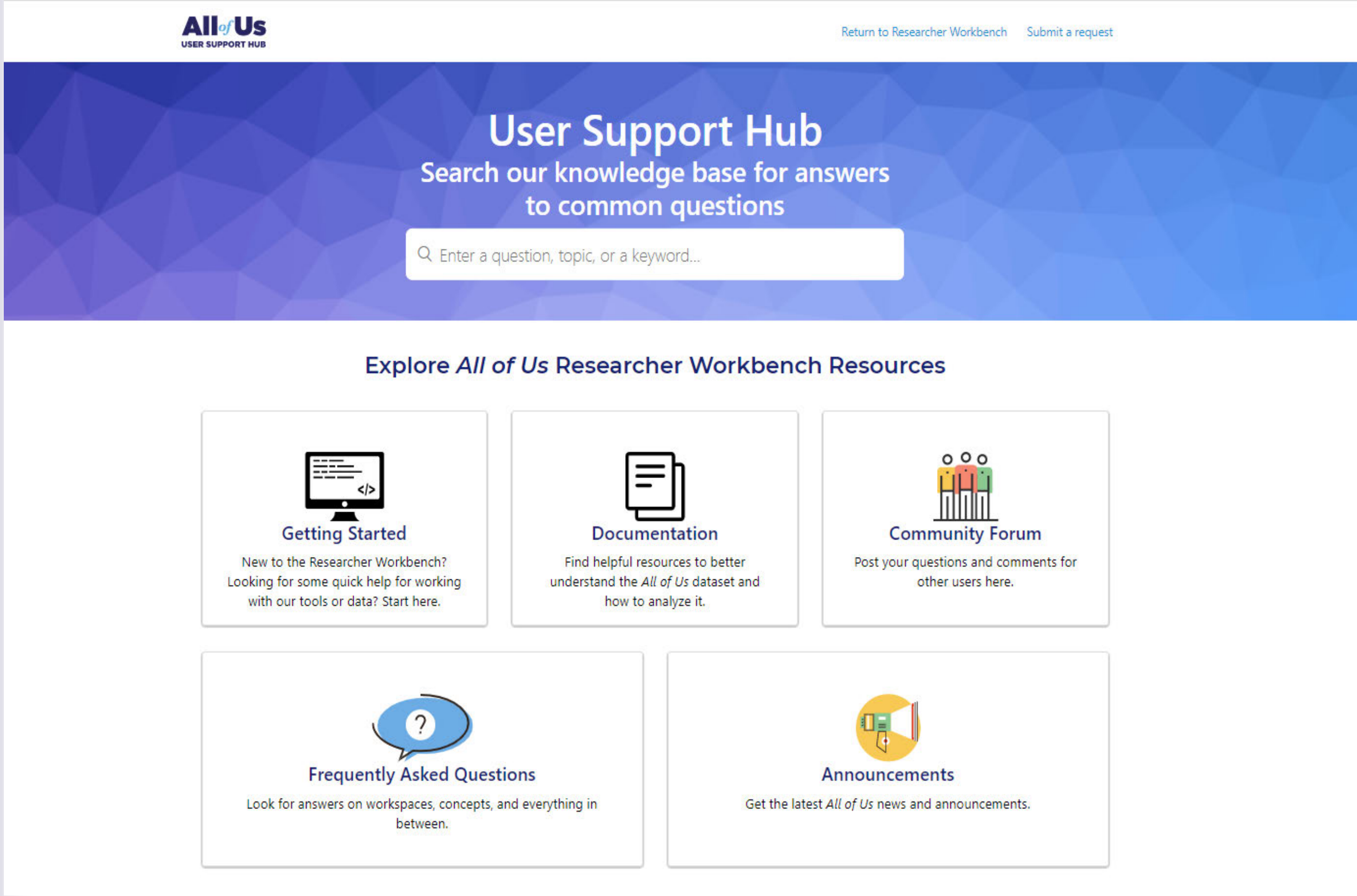
Quick Tour and Videos

Research Hub

- Data Browser
- Apply Page

Researcher Workbench

- **Support Hub**
- Ex. Workspace



User Support Hub

Search our knowledge base for answers to common questions

Explore *All of Us* Researcher Workbench Resources

Getting Started

New to the Researcher Workbench? Looking for some quick help for working with our tools or data? Start here.

Documentation

Find helpful resources to better understand the *All of Us* dataset and how to analyze it.

Community Forum

Post your questions and comments for other users here.

Frequently Asked Questions

Look for answers on workspaces, concepts, and everything in between.

Announcements

Get the latest *All of Us* news and announcements.

Research Hub

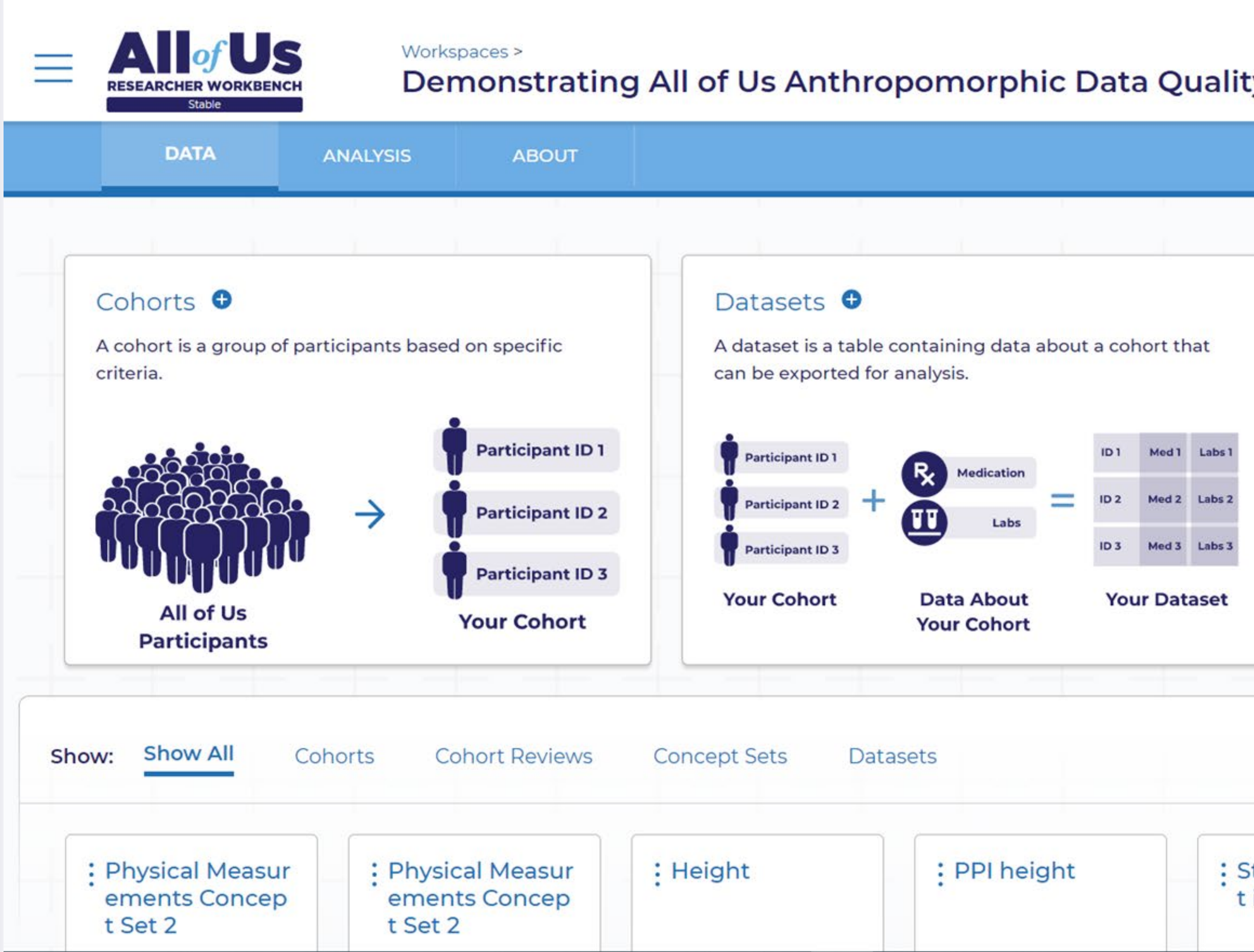
Data Browser

Apply Page

Researcher Workbench

Support Hub

Ex. Workspace



Demonstrating All of Us Anthropomorphic Data Qualit

Cohorts

A cohort is a group of participants based on specific criteria.

Datasets

A dataset is a table containing data about a cohort that can be exported for analysis.

Live Demo on **Synthetic Data***

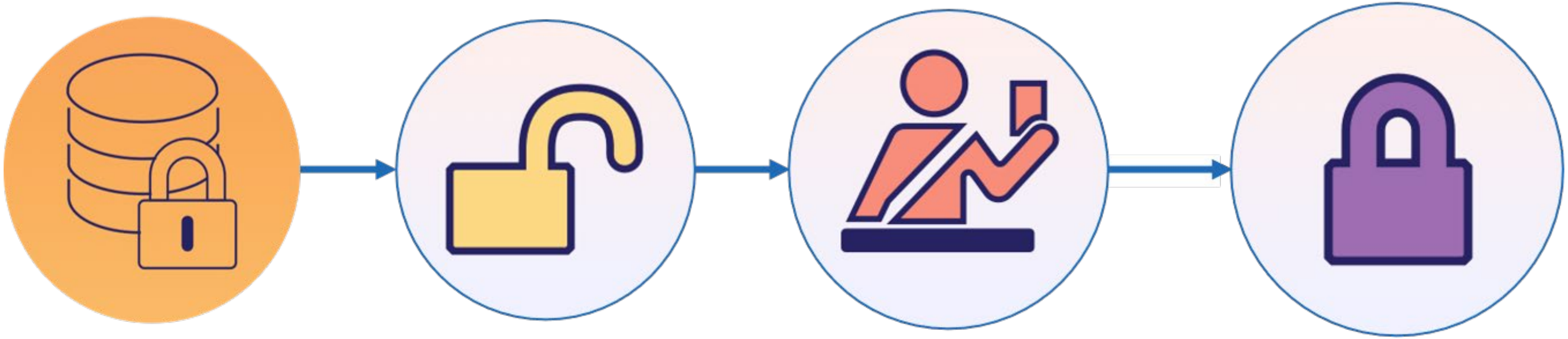
“Demonstrating *All of Us* Anthropomorphic Data Quality”

Primary Investigator	Co-Investigator	Data Scientist	Graduate Student
Kelsey Mayo, Ph.D.	Adrienne Roman, Ph.D.	Francis Ratsimbazafy, Ph.D.	Michael Holmes, B.S.
			

All of Us
DATA & RESEARCH CENTER

Outline for Today

1. Data Overview

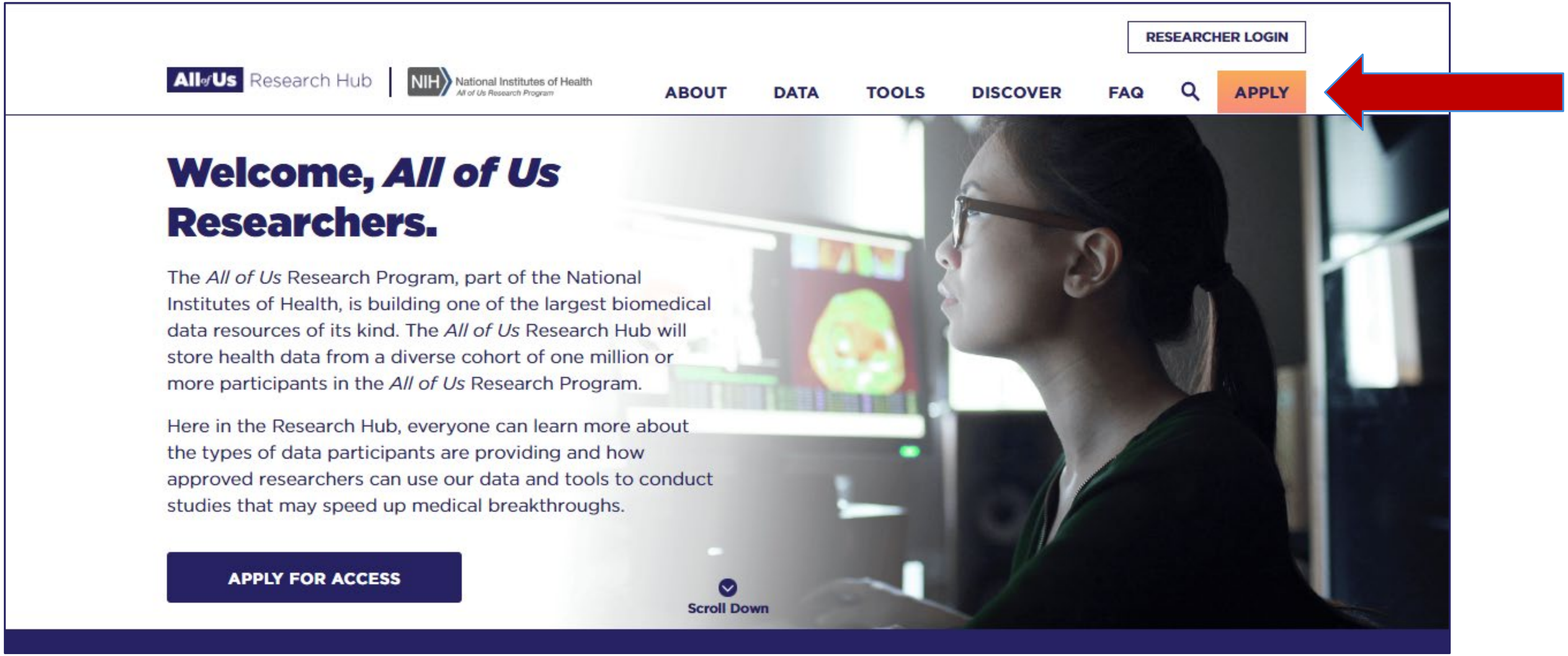


2. Live demo of researchallofus.org & Researcher Workbench

3. Thank you! & Wrap Up

Interested in becoming a Beta Researcher?

Visit researchallofus.org to apply!



Welcome, *All of Us* Researchers.

The *All of Us* Research Program, part of the National Institutes of Health, is building one of the largest biomedical data resources of its kind. The *All of Us* Research Hub will store health data from a diverse cohort of one million or more participants in the *All of Us* Research Program.

Here in the Research Hub, everyone can learn more about the types of data participants are providing and how approved researchers can use our data and tools to conduct studies that may speed up medical breakthroughs.

All of Us
DATA & RESEARCH CENTER

Coming up!

“Researcher Workbench: Perspective from an informatics researcher” by Dr. Lucila Ohno-Machado

“Racial, Ethnic, and Gender Differences in Obesity and Body Fat Distribution: an All of Us Research Program Demonstration Project” by Dr. Jason Karnes

“Health Outcomes of Sexual and Gender Minority People: Data from the *All of Us* Research Program” by Dr. Mitch Lunn

All of Us Researcher Onboarding Walkthrough Breakout Sessions

BREAK TIME!

Please return by 12:45 pm ET

Follow us!

Twitter @AllofUsResearch
Facebook www.facebook.com/AllofUsResearch
Instagram: https://www.instagram.com/allofusresearch/
YouTube: https://www.youtube.com/channel/UCQId1TfpwPaYiDIGlxEhlkA/feed

Researcher Workbench:
Perspective from an Informatics Researcher

All of Us
Research Program

Lucila Ohno-Machado, MD, PhD, AoU Science Committee member

Demonstration Project Co-investigator

November 12, 2020

Demonstration Projects TF Charter

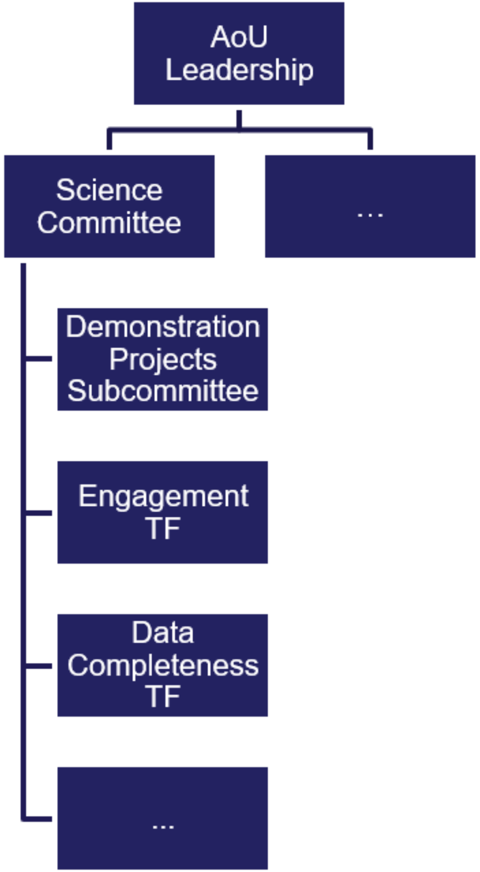
Mission: Demonstrate the quality, utility, and diversity of *All of Us* Research Program data and tools via partnership with awardees

Goal: Fully executed research projects demonstrating the utility and validity of AoU data timed to publish at data platform launch, *not* novel discovery work.

Demonstration Projects Subcommittee

Chair, Demonstration Projects Subcommittee
- Andrea H. Ramirez, MD, MS

Science Committee Demonstration Project Subcommittee
- Mine Cicek, PhD
- Cheryl Clark, MD, ScD
- Elizabeth Cohn, PhD, RN
- Lucila Ohno-Machado, MD, PhD
- Robert Winn, MD
- Eric Boerwinkle, PhD
- Roxana Loperena, PhD
- Kelsey Mayo, PhD
- Stephen Mockrin, PhR
- Sheri Schully, PhD
- Kelly Gebo, MD, MPH
- Replication of known disease/risk factor associations or predictive models
- Implementation of phenotype algorithms
- Cohort characterization



AoU Leadership

Science Committee

…

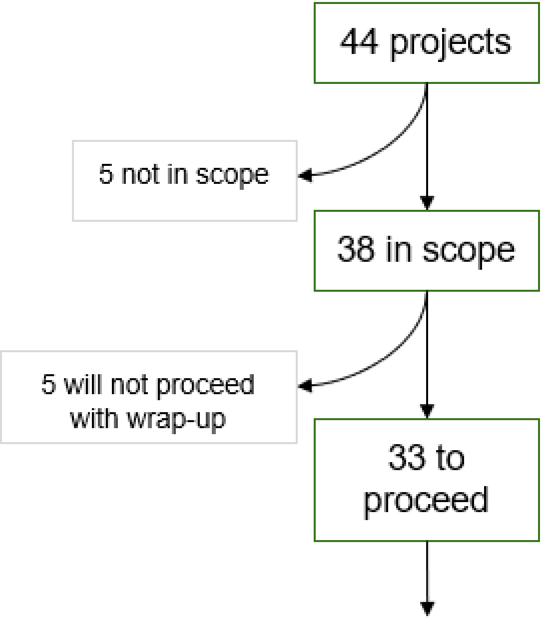
Demonstration Projects Subcommittee

Engagement TF

Data Completeness TF

…

Phase II Demonstration Projects Progress



44 projects

5 not in scope

38 in scope

5 will not proceed with wrap-up

33 to proceed

Step 0 (representative)	Step 1 (concept sheet)	Step 2 (collaborator)	Step 3 (hotfix CDR)	Step 4 (Workspace and notebook clean-up)	Step 5-7
33 projects	33 projects	30 projects	28	29	6

Demonstration Projects

Test the platform

Provide constructive feedback

Investigate reproducibility of previous studies or statistics

- Internal request for proposals among consortium institutions
- Various internal meetings, individual interactions with the Data Research Center
- General meeting to report on all projects
- Publications under review at various journals

AoU Science Committee: Demonstration Projects Subcommittee

NOVEMBER 2019

Mission: Demonstrate the quality, utility, and diversity of *All of Us* Research Program data and tools via partnership with awardees

Goal: Fully executed research projects demonstrating the utility and validity of AoU data timed to publish at data platform launch, *<u>not</u>* novel discovery work.

Slide from Andrea Ramirez

Overall Timeline 2020

Feb 4	Final Project Update (#6)
Week of Feb 24	Final Feedback Survey
Late Feb - Early March	Beta Launch!
Feb - March	Schedule Phase II Demonstration Projects Code Review
April - May	Work with teams to publish Phase II Manuscripts
July	Calls for Phase III Demonstration Projects - Controlled Tier or new data types
July - Oct	Phase III Demonstration Projects

Slide adapted from from Andrea Ramirez’

Engaging a large consortium, covering broad health areas

***All of Us* Awardee**	**No. of Approved Projects**
Trans-American Consortium for the Health Care Systems Research Network (TACH)	4
NYC Precision Medicine Consortium	4
The Participant Center	5
Asian Health Coalition	2
University of Arizona	4
New England Precision Medicine Consortium	4
California Precision Medicine	2
Stanford University	1
Illinois Precision Medicine Consortium	5
South East Enrollment Center (SEEC)	4
Data and Research Center (DRC)	3

Health Areas	**No. Projects in Health Area**
Cardiovascular Disease	7
Cancer	5 (2 overlaps with other area)
Diabetes and Obesity	6 (3 overlap with other areas)
Mental Health	2
Wellness	2
Opioids and Pain	2
Chronic Kidney Disease	1
Chronic lower respiratory disease	1
Neurodegenerative condition and cognition	1
Other areas: Environmental exposures, Health disparities, Infectious disease, Body temperature, Women's health, Pediatric data description, Hematologic disorders	13
Gap addressing projects: Ophthalmology Compare Census and CDC data in AoU Rare diseases	3

Slide from Andrea Ramirez

Examples of proposals that were not accepted

Common reasons for decline for Phase II projects

- **Data not available at the time**
- **Proposal was beyond comparison of AoU results with published data**

D040: Much of the data proposed in the studies 1, 2, 3, 5, 6, 7 is not available in the current CDR […]. However, study 4 related to glaucoma is a gap area

D041: Aims 1 and 2 in this proposal are not in scope with what demonstration projects are, and are considered more operational studies […]. Aim 3 is within scope […]. It will be very interesting to compare the Census data with the AoU data

Aim 1: Compare the level of research participation by demographic characteristics with published results.

Aim 2: Compare the level of technological participation by demographic characteristics.

Does AoU have the data I need?

1. Check Data Browser for existence of variable and count of participants with values
2. Check the Workbench's Data Dictionary for specifics
3. Check the Cohort Definition Library to see if someone already defined the cohort you are interested in
4. Check for Notebooks with code you can reuse

Quick overview of demo projects

Asian Health Coalition Project 1

D004: *Characterizing the Prevalence of Type 2 Diabetes in Asian American Cohort of the All Of Us Research Program*

BACKGROUND

- Asian Americans are more likely to develop Type 2 Diabetes even with a lower body mass index (BMI).
- Asian Americans have a 60% higher risk of Type 2 Diabetes than non-Hispanic whites.
- The U.S. Preventive Services Task Force recommends screening for abnormal blood glucose and type 2 diabetes in adults 40 to 70 years of age who are overweight or obese (BMI ≥ 25).
- American Diabetes Association (ADA) revised its guidelines in 2015, recommending all Asian Americans 45 or older should be screened when they have BMI values starting at 23.

METHODS

1. Cohort Selection

- Demographics = Race (Asian) → 7.607

  <AND>

- Conditions = Type 2 Diabetes → 405

  <OR>

- Conditions = Diabetes → 456

  <OR>

- Measurements = Glycated Hemoglobin ≥ 6.5%* → 602

  <OR>

- Measurements = Fasting Glucose in Serum/Plasma ≥ 126* → 608

  <OR>

- Measurements = Glucose in Serum/Plasma ≥ 126* → 902

* Cheng YJ et al. Prevalence of diabetes by race and ethnicity in the United States, 2011-2016. JAMA, 2019; 322(24): 2389-2398.

METHODS

2. **Variables**
    - Physical Measurements
        - Height
        - Weight
    - PPI (The Basics)
        - Age
        - Gender
        - Employment
        - Health Insurance Status
    - PPI (Overall Health)
        - Overall Health
        - Rating of Physical Health
        - Rating of Mental Health

Asian Health Coalition Project 2

D005: *Describe the Demographic Characteristics of the Asian American Cohort in the All of Us Research Program*

SPECIFIC AIM

- Compare the demographic characteristics of the Asian American cohort with the 2017 Population Estimates and 2017 American Community Survey 5-year estimated data

Illinois Precision Medicine

Illinois Project 1

1. Describe prevalence of HTN in the *AoU* Research Program by demographic factors
2. Among those with HTN, describe prevalence of awareness, treatment and control of HTN by demographic factors
3. Compare HTN prevalence, awareness, treatment and control in *AoU* Research Program to the 2015-2016 National Health and Nutrition Examination Survey (NHANES)
4. Compare HTN prevalence, awareness and treatment to documentation in the EHR.

Illinois Project 2

Cohort characterization of mammography rates in All of Us

- Background: While clinical recommendations vary, between the ages of 45 and 55, most women are recommended for a mammography screening
- Not all women receive the recommended screening

New England Precision Medicine

New England Research Project 1

What is the prevalence of hypertension (HTN) defined using an electronic health record definition from eMERGE among UBR groups defined by race/ethnicity, income and education?

Replication of HTN rates from NHANES 2015-2016, among total, by sex; by sex and race/ethnicity; by sex and age, by sex and race/ethnicity 2015-2016 (Fryer 2017)

Do treatment pathways for HTN (using medication sequencing analysis) vary by UBR groups defined by race/ethnicity, income and education, and in geographic regions based on grouping states?

Replication of US data on treatment pathways from OHDSI (Hripsak, 2016)

Fryer, CD, Ostchega Y, Hales, CM, Hypertension prevalence and control among adults: United States, 2015-2016
Hripcsak G, Ryan PB, Duke JD, et al. Characterizing treatment pathways at scale using the OHDSI network. Proc Natl Acad Sci U S A 2016;113:7329-36

New England Research Project 2

- **What is the prevalence of World Health Organization (WHO) defined categories for overweight and obesity at the state-level among UBR groups defined by race/ethnicity, income and education for adults aged 18 and older using measured PM and EHR data from the *All of Us* Research Program?**
  - Replication of obesity rates from NHANES 2015-2016 by overall, by sex; by sex and race/ethnicity; by sex and age
  - Examine the prevalence of World Health Organization (WHO) defined categories for overweight and obesity at the state-level for adults aged 18 and older using BMI calculated from PM height and weight measurements, adjusted for age, gender, race/ethnicity, income, and education data from the *All of Us* Research Program to replicate Ward et al NEJM, 2019

Projected U.S. State-Level Prevalence of Adult Obesity and Severe Obesity. Ward et al. NEJM (2019)

New England Research Project 3

Describe the epidemiology of fibroids in the *AoU* female cohort

- Defining uterine fibroids in the *All of Us* Cohort using 1 SNOMED code vs Feingold Algorithm
- Compare results obtained from the *AoU* female cohort to published data by Stewart and by Feingold

Criteria for Eligibility

- Assigned female sex at birth
- Has both EHR and survey data

New York City Precision Medicine

Association between Blood Pressure and Stroke in Diverse Populations

South East Enrollment Center (SEEC)

South East Enrollment Center (SEEC) - Emory

D028: Epidemiology of atrial fibrillation

- Background
  - Atrial fibrillation is a common arrhythmia (lifetime risk between 1 in 3 and 1 in 5)
  - Associated with increased risk of stroke, heart failure, dementia
  - Prior epidemiologic studies show:
    - Higher risk in older individuals
    - Higher risk in men than women
    - Higher risk in non-Hispanic whites vs other racial/ethnic groups in the US
    - Risk factors: hypertension, obesity, diabetes, other CVDs

Trans-American Consortium for the Health Care Systems Research Network (TACH)

TACH - Henry Ford Health System

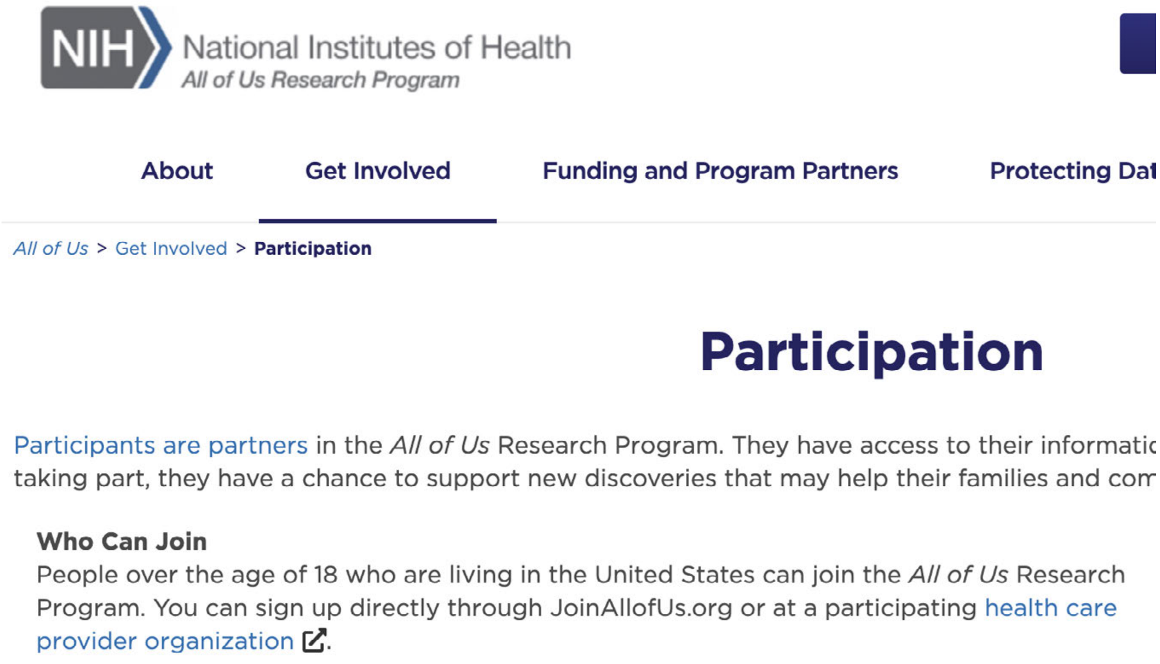
D014 - Examining Opioid Use in the All of Use Research Program Cohort

- Identify the prevalence of opioid use as recorded by electronic health record data
- Examine the lifetime and recent prevalence of opioid misuse, both street opioid use and prescription opioid use, as derived from participant provided information surveys.

Data Research Center

Data Research Center Project 1

D38: Exploration of Pediatric Data from All of Us



All of Us > Get Involved > **Participation**

Participation

Participants are partners in the *All of Us* Research Program. They have access to their informatic
taking part, they have a chance to support new discoveries that may help their families and com

Who Can Join

People over the age of 18 who are living in the United States can join the *All of Us* Research Program. You can sign up directly through JoinAllofUs.org or at a participating health care provider organization.

- When patients enroll, EHRs from childhood may come along
- Questions:
    - How much data is there?
    - What data types are there?
    - Can the data be used for pediatric research?

Data Research Center Projects 2 and 3

D43: Replicate a Phenotype Risk Scores approach for three mendelian diseases, including cystic fibrosis (CF), hereditary hemochromatosis (HH) and sickle cell (SC) anemia in allofus (AOU) cohort.

D44: Comparison of family health history gathered in survey and electronic health records in *All of Us* Registered Tier Data

California Precision Medicine

California Precision Medicine Project 1

- **D040: Validation of Published Clinical Descriptive and Predictive Models on AoU Data**

  Predictive Modeling of Glaucoma Surgery using Data from All of Us

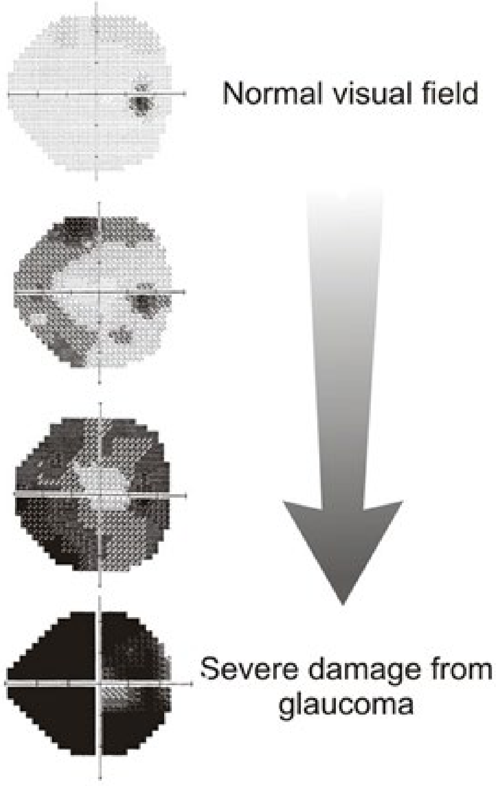
  - PI Sally Baxter

  Background

  - Glaucoma is the world's leading cause of irreversible blindness
  - Projected to affect 80 million people worldwide in 2020 and over 110 million in 2040

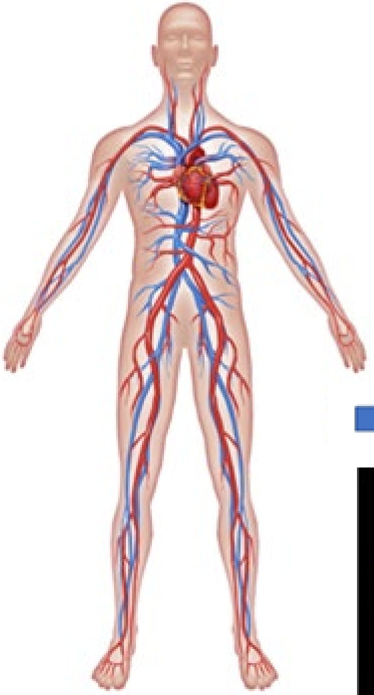
Photos: National Eye Institute, National Institutes of Health

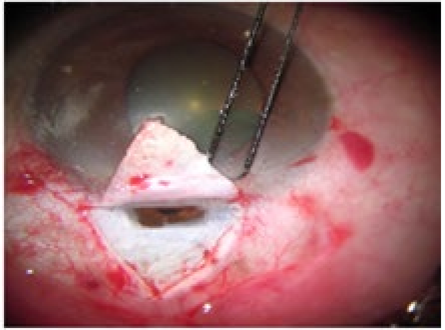
Glaucoma Simulation

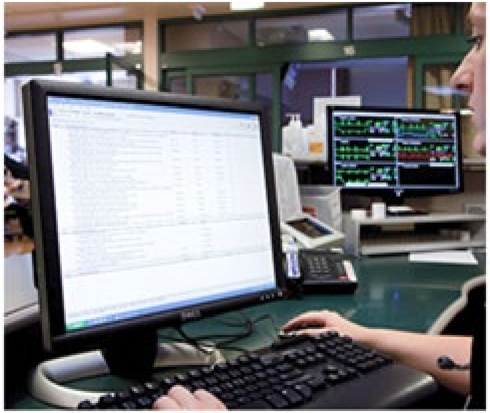


California Precision Medicine

- **Study Question**



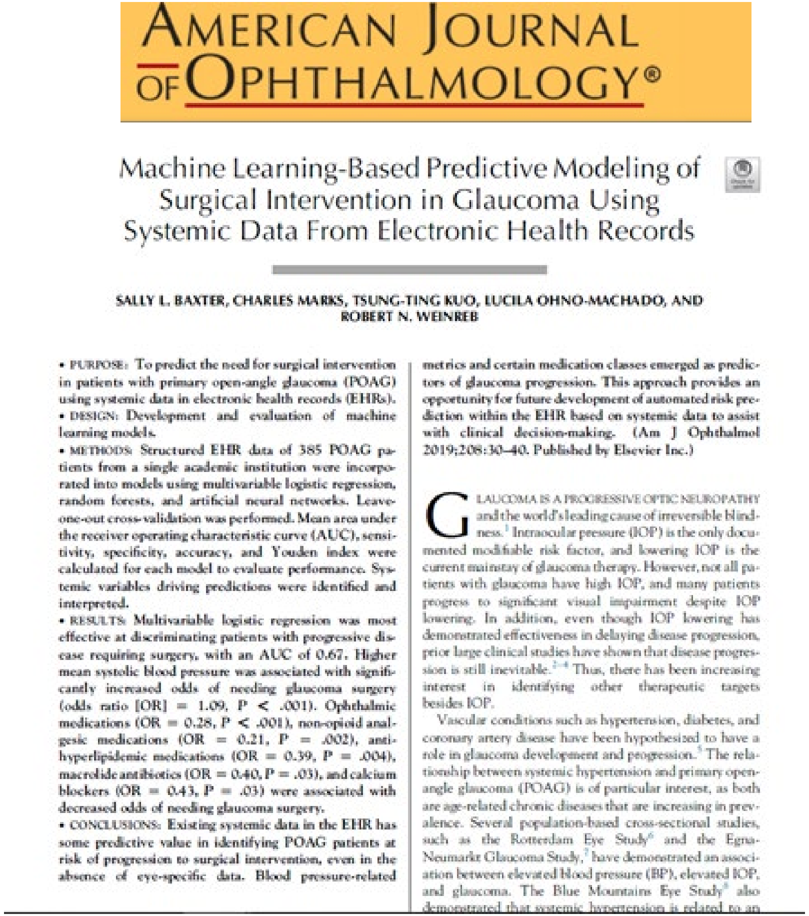




Can data-drive modeling of systemic attributes obtained from existing EHR datapredict the risk of glaucoma progression, as indicated by the need for a surgical intervention within 6 months?

California Precision Medicine

- **Prior Published Predictive Model**
    - Existing systemic data in the EHR have some predictive value in identifying patients with glaucoma at risk of progression to surgical intervention *even in the absence of eye-specific endpoints*
    - Blood pressure-related metrics and certain medication classes emerged as important predictors
    - Evaluation of systemic EHR data likely underutilized in clinical practice



AMERICAN JOURNAL OF OPHTHALMOLOGY®

Machine Learning-Based Predictive Modeling of Surgical Intervention in Glaucoma Using Systemic Data From Electronic Health Records

SALLY L. BAXTER, CHARLES MARKS, TSUNG-TING KUO, LUCILA OHNO-MACHADO, AND ROBERT N. WEINREB

- PURPOSE: To predict the need for surgical intervention in patients with primary open-angle glaucoma (POAG) using systemic data in electronic health records (EHRs).
- DESIGN: Development and evaluation of machine learning models.
- METHODS: Structured EHR data of 385 POAG patients from a single academic institution were incorporated into models using multivariable logistic regression, random forests, and artificial neural networks. Leave-one-out cross-validation was performed. Mean area under the receiver operating characteristic curve (AUC), sensitivity, specificity, accuracy, and Youden index were calculated for each model to evaluate performance. Systemic variables driving predictions were identified and interpreted.
- RESULTS: Multivariable logistic regression was most effective at discriminating patients with progressive disease requiring surgery, with an AUC of 0.67. Higher mean systolic blood pressure was associated with significantly increased odds of needing glaucoma surgery (odds ratio [OR] = 1.09, $P < .001$). Ophthalmic medications (OR = 0.28, $P < .001$), non-opioid analgesic medications (OR = 0.21, $P = .002$), anti-hyperlipidemic medications (OR = 0.39, $P = .004$), macrolide antibiotics (OR = 0.40, $P = .03$), and calcium blockers (OR = 0.43, $P = .03$) were associated with decreased odds of needing glaucoma surgery.
- CONCLUSIONS: Existing systemic data in the EHR has some predictive value in identifying POAG patients at risk of progression to surgical intervention, even in the absence of eye-specific data. Blood pressure-related metrics and certain medication classes emerged as predictors of glaucoma progression. This approach provides an opportunity for future development of automated risk prediction within the EHR based on systemic data to assist with clinical decision-making. (Am J Ophthalmol 2019;208:30–40. Published by Elsevier Inc.)

GLAUCOMA IS A PROGRESSIVE OPTIC NEUROPATHY and the world's leading cause of irreversible blindness.[1] Intraocular pressure (IOP) is the only documented modifiable risk factor, and lowering IOP is the current mainstay of glaucoma therapy. However, not all patients with glaucoma have high IOP, and many patients progress to significant visual impairment despite IOP lowering. In addition, even though IOP lowering has demonstrated effectiveness in delaying disease progression, prior large clinical studies have shown that disease progression is still inevitable.[2–4] Thus, there has been increasing interest in identifying other therapeutic targets besides IOP.

Vascular conditions such as hypertension, diabetes, and coronary artery disease have been hypothesized to have a role in glaucoma development and progression.[5] The relationship between systemic hypertension and primary open-angle glaucoma (POAG) is of particular interest, as both are age-related chronic diseases that are increasing in prevalence. Several population-based cross-sectional studies, such as the Rotterdam Eye Study[6] and the Egna-Neumarkt Glaucoma Study,[7] have demonstrated an association between elevated blood pressure (BP), elevated IOP, and glaucoma. The Blue Mountains Eye Study[8] also demonstrated that systemic hypertension is related to an

California Precision Medicine

- **Prior Published Predictive Model - 15 predictors**

Table 4. Relative contribution of various predictor variables in the multivariable logistic regression model predicting need for surgical intervention within 6 months among patients with primary open-angle glaucoma

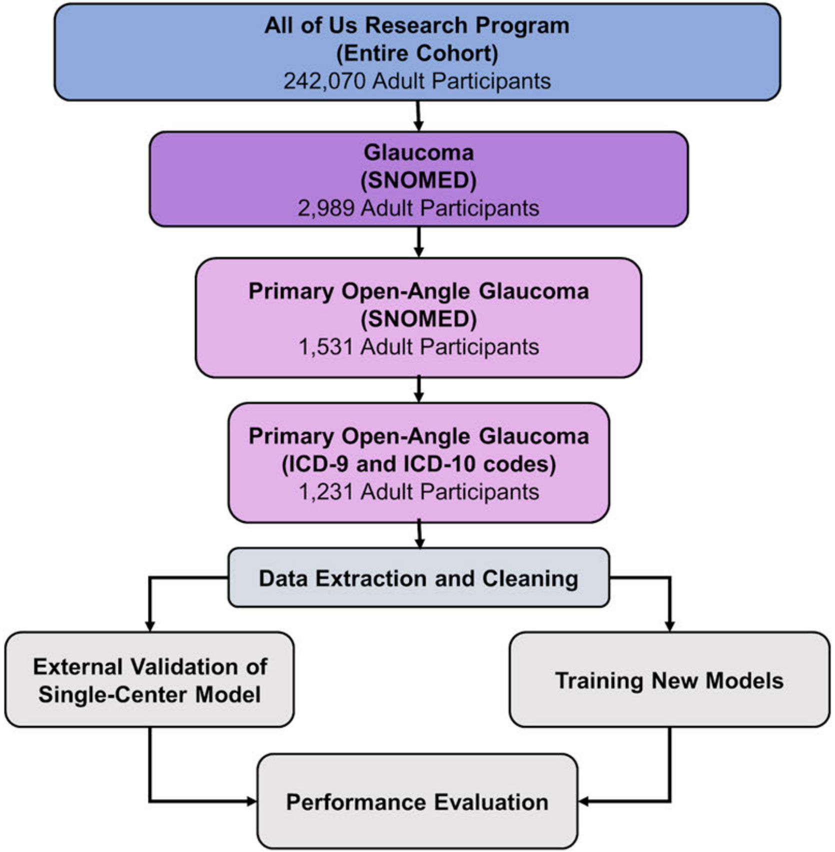
Variable	Adjusted Odds Ratio (95% Confidence Interval)	P-Value[a]
Ophthalmic medication	0.28 (0.17, 0.46)	<0.001
Minimum systolic blood pressure	0.92 (0.89, 0.95)	<0.001
Mean systolic blood pressure	1.09 (1.06, 1.13)	<0.001
Non-opioid analgesic medication	0.21 (0.07, 0.52)	0.002
Anti-hyperlipidemic medication	0.39 (0.21, 0.73)	0.004
Number of days hospitalized	0.97 (0.94, 0.99)	0.006
Calcium blocker medication	0.43 (0.21, 0.89)	0.025
Macrolide antibiotic medication	0.40 (0.17, 0.93)	0.034
Anticoagulant medication	2.75 (1.05, 7.46)	0.042
Male gender	1.52 (0.94, 2.47)	0.089
Cold/cough medication	2.22 (0.83, 6.06)	0.115
Minimum diastolic blood pressure	0.98 (0.95, 1.01)	0.117
Dementia	0.26 (0.04, 1.38)	0.141
Antidepressant medication	0.56 (0.50, 1.21)	0.143
Metastatic disease	0.31 (0.06, 1.43)	0.149

[a] The threshold for statistical significance was $p<0.05$

California Precision Medicine

- **Aim of this Demonstration Project:**
    - To validate the prior predictive model using participants' EHR data from the All of Us Research Program
- **Methods:**
    - **Defined patient cohort** with same ICD criteria as original cohort
    - Used the **Workbench interface** to build datasets based on queries for predictors in the model (concept sets)
        - Supplemented these with **custom SQL queries** directly within the notebook when necessary (i.e. for visits, medications)
    - **Cleaned the data t**o produce a final dataset with the same structure as the dataset used for the original model
    - **Evaluated model performance** using the AoU dataset as a test dataset

California Precision Medicine



All of Us Research Program
(Entire Cohort)
242,070 Adult Participants

Glaucoma
(SNOMED)
2,989 Adult Participants

Primary Open-Angle Glaucoma
(SNOMED)
1,531 Adult Participants

Primary Open-Angle Glaucoma
(ICD-9 and ICD-10 codes)
1,231 Adult Participants

Data Extraction and Cleaning

External Validation of Single-Center Model

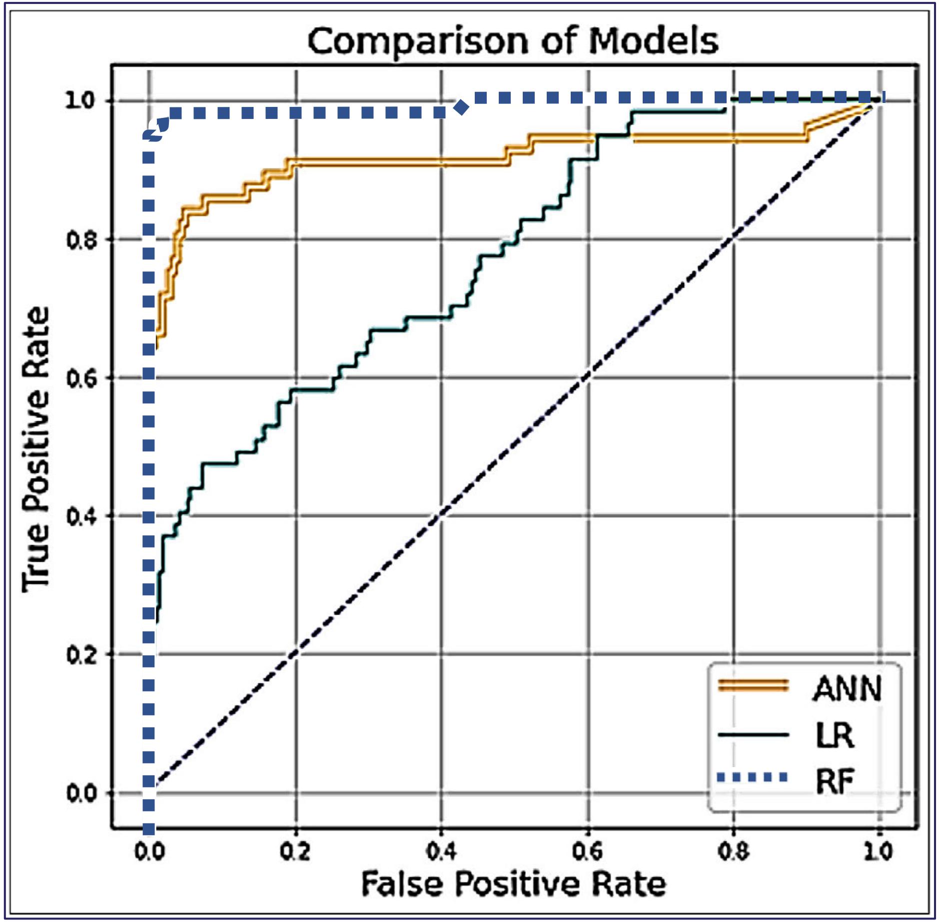
Training New Models

Performance Evaluation

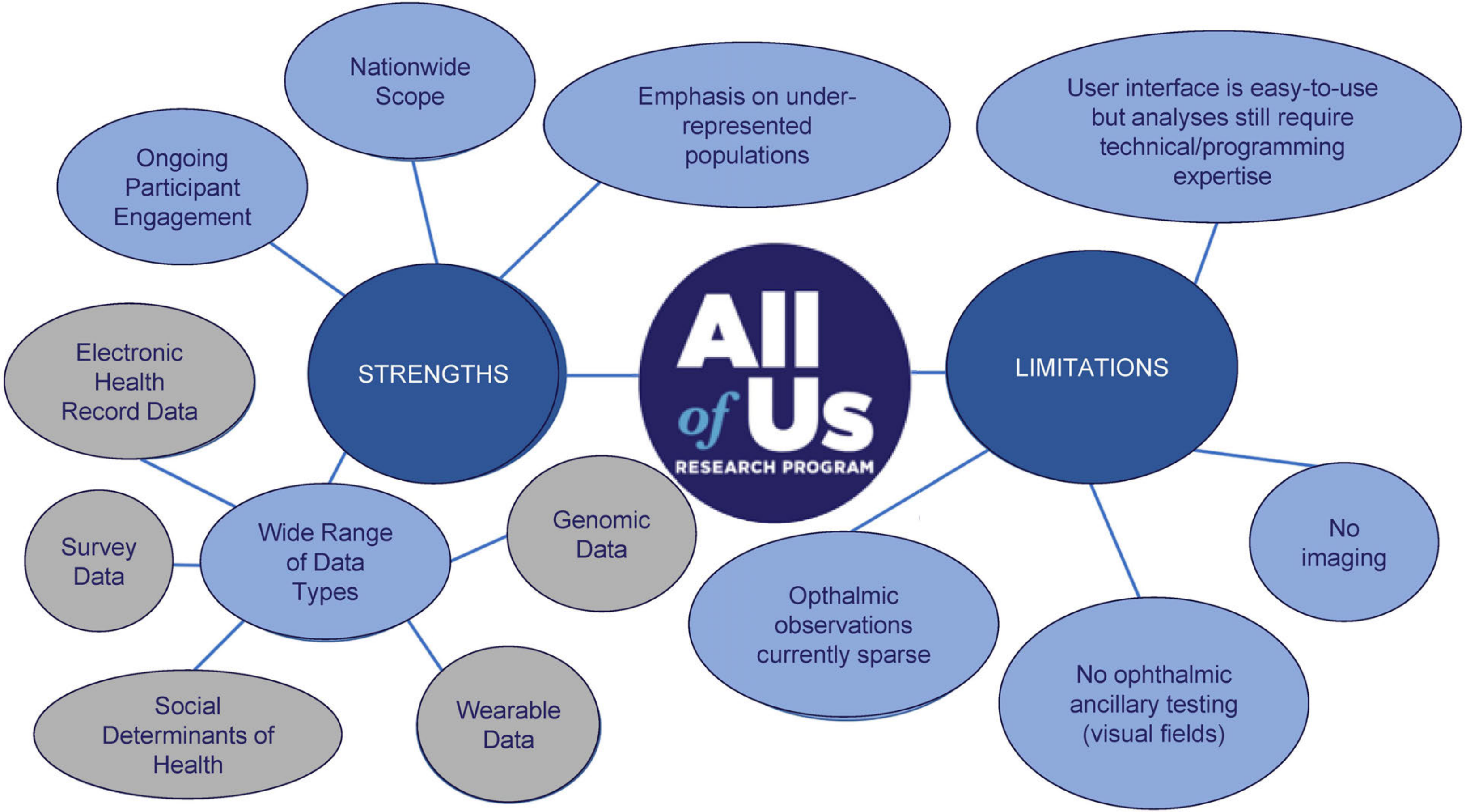
Table 2. Comparison of demographic characteristics and primary outcome between cohorts of adults with primary open-angle glaucoma derived from a single academic center and derived from the All of Us data repository.

	Single-center Cohort (N=385)	***All of Us* Cohort** (N=1231)	**p-value**
Mean (SD) Age	73.1 (12.2) years	69.1 (10.5) years	<0.001
Female	198 (51.4%)	705 (57.3%)	0.04
Self-Reported Race			<0.001
White	214 (55.6%)	508 (41.3%)	
Black or African-American	23 (6.0%)	412 (33.5%)	
Asian	49 (12.7%)	27 (2.2%)	
Other Race or Mixed Race	70 (18.2%)	21 (1.7%)	
None or Skipped	29 (7.5%)	263 (21.4%)	
Participants who underwent glaucoma surgery	174 (45.2%)	286 (23.2%)	<0.001

California Precision Medicine



California Precision Medicine



STRENGTHS

- Nationwide Scope
- Emphasis on under-represented populations
- Ongoing Participant Engagement
- Wide Range of Data Types
  - Electronic Health Record Data
  - Survey Data
  - Social Determinants of Health
  - Wearable Data
  - Genomic Data

LIMITATIONS

- User interface is easy-to-use but analyses still require technical/programming expertise
- Opthalmic observations currently sparse
- No ophthalmic ancillary testing (visual fields)
- No imaging

California Precision Medicine Project 2

D041: Understanding Regional, Demographic, and Other Characteristics in Incidence and Prevalence of Cancer

PI Katherine Kim

Explore cancer prevalence and incidence in the AoU participant population compared with national population estimates.

- 5-year cancer prevalence rate for 2010-2015 by demographics, geography, and self-report compared to EHR
- Compare AoU to NHIS reported prevalence of family history of cancer
- 1-year incidence of cancer for 2016 by demographics, geography, and self-report compared to EHR

Phase II Demo Project Culmination Meeting

- Thursday, Jan 30th in Nashville, TN
- 50 overall attendees, 26 demo project team members
- 9 out of 11 Consortium Awardees represented (The Participant Center, Asian Health Coalition)
- 22 different projects presented both on-site and remotely (of 38 total)
- Discussions around feedback mechanism and new beta features
- Discussions around authorship and alpha access next steps and publication of projects internally in Featured Workspaces

Slide from Andrea Ramirez

Participants' privacy is a priority: What can/cannot be done

No individual-level data can leave the AoU platform, only aggregates

Data have been obfuscated to protect privacy

- Geographic information: State is the unit of analysis (no zip code)
- Counts below 20 must not be reported

Not all AoU data from EHRs are available yet (e.g., clinical notes)

Operational data are not yet available for research

Feedback gathered

- **Feedback on process, data and tools**
    - Easy to use overall (but need to have a programmer with OMOP knowledge on the team for EHRs)
    - Concept sets straightforward to build for most variables
    - Very nice visualization tools
    - Helpful code snippets for analyses

  Hope these could come soon

    - More documentation about what EHR data is/will be available and how it is structured (e.g, will there be additional lab results? histology codes?)
    - Clear mapping of PPI medical conditions to ICD to allow cross-analytics
    - Would have liked to see medication classes included in the interface
    - Would be great to have OMOP Visits Table included in the interface
    - Would like to be able to directly visualize Google Bucket files

Racial, Ethnic, and Gender Differences in Obesity and Body Fat Distribution: an *All of Us* Research Program Demonstration Project

Jason H Karnes, Amit Arora, Lina Sulieman, Eric Boerwinkle, Mine Cicek, Cheryl Clark, Elizabeth Cohn, Kelly Gebo, Roxana Loperena, Kelsey Mayo, Steve Mockrin, Lucila Ohno-Machado, Andrea Ramirez, Sheri Schully, Yann C Klimentidis* on behalf of the *All of Us* Research Program Investigators

NIH
National Institutes of Health

Background

- Type-2 diabetes and cardiovascular disease among the most pressing health issues of our time
- These diseases and their risk factors, such as abdominal obesity, exhibit notable disparities across different racial/ethnic groups[1,2]
  - Differences in body weight and shape across gender and race/ethnicity extensively described
- Used the first interim release of AoU data to replicate previous findings
  - Assessed suitability, breadth, depth, and quality of AoU data.
- Set out to replicate previous findings and examine how nationally representative AoU data is with regard to obesity and body fat distribution3
- We focused on racial/ethnic and gender differences in body weight and body fat distribution
  - Also evaluated levels of alanine aminotransferase (ALT), a surrogate measure of hepatic steatosis

1. Hales C, Carroll M, Fryar C, Ogden C. Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017-2018. NCHS Data Brief. 2020;360.
2. Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. Lancet (London, England). 2017 Dec;390(10113):2627–42.
3. Denny JC, Rutter JL, Goldstein DB, Philippakis A, Smoller JW, Jenkins G, et al. The “All of Us” Research Program. N Engl J Med. 2019 Aug;381(7):668–76.

METHODS: Design and Data Collection

- AoU data is a patchwork of datasets with varying numbers of total patients[3]
    - Participant-provided information (PPI), Physical measurements, Electronic health records (EHRs)
- Self-reported Non-Hispanic Whites (NHW), Non-Hispanic Blacks (NHB), Asians, and Hispanics
    - Based on response to: Which categories describe you? (Select all that apply): 1) American Indian and Alaska Native; 2) Black, African American, or African; 3) Asian; 4) Hispanic, Latino, or Spanish; 5) Middle Eastern or North African; 6) Native Hawaiian or Other Pacific Islander; 7) White; 8) None of these describe me; 9) Prefer Not To Answer.
    - Due to privacy methodology, categories condensed to White; Black; Asian; Hispanic, Latino or Spanish; Other (if one of options 5-6 or None); Two or more races (if multiple options 1-7 selected); and Prefer not to answer.
    - All participants reporting American Indian and Alaska Native race/ethnicity removed
- Biological sex at birth assessed at baseline
    - only considered biological sex assigned at birth
- Standardized physical measurements obtained at baseline
    - Height, weight, waist circumference (WC), and hip circumference
- ALT levels from EHR records (LOINC code 1742-6)
    - Most recent ALT (if multiple labs available) and excluded ALT measured prior to 2005
    - Heavy drinkers (≥14 drinks per week based on PPI) excluded from ALT analyses

3. Denny JC, Rutter JL, Goldstein DB, Philippakis A, Smoller JW, Jenkins G, et al. The "All of Us" Research Program. N Engl J Med. 2019 Aug;381(7):668–76.

METHODS: Statistical Analysis

- All registered tier data used if existed on a participant (maximum number of individuals included in each analysis)
- Distribution of raw and calculated anthropomorphic measurements examined visually and outliers removed
  - height >250 cm (n≤20), BMI>150 kg/m2 (n≤20), BMI<12 kg/m2 (n≤20), WHR<0.2 (n≤20), and WHR>6 (n≤20)
  - Due to privacy restrictions, no single individual or groups of individuals below 20 are reported
  - ALT outliers also excluded (>99th percentile)
- Obesity was defined as having a BMI ≥ 30 kg/m2
- T-tests for differences in continuous variables between males and females within race/ethnicity categories
- Age-adjusted obesity prevalence rates calculated by race/ethnicity and gender and compared to NHANES age-adjusted obesity rates1
- Linear regressions for BMI and ALT using race/ethnicity, gender, age, and BMI (for ALT modelling)
  - Additional regressions with gender x race interaction term
- Analyses performed in R version 3.6.2 within the AoU Researcher Workbench Jupyter Notebook
- Code available within Researcher Workbench at https://workbench.researchallofus.org/workspaces/aou-rw-54ae5687/racialethnicdifferencesanthropolipidalt/notebooks/Notebook_addressingComments.ipynb

---

1. Hales C, Carroll M, Fryar C, Ogden C. Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017-2018. NCHS Data Brief. 2020;360.

RESULTS

- Total of 242,285 participants included in AoU Registered Tier data at time of code review
  - 224,350 had necessary PPI-derived data on race/ethnicity and gender
  - 176,714 used in analysis of anthropomorphic measurements
  - 46,710 were used in the analysis of ALT
- BMI highest among NHB women (mean 33.32 kg/m^2) and lowest among Asian women (mean 24.52 kg/m^2)
- Gender difference in BMI most pronounced among NHB
  - NHB women 33.32 kg/m^2 versus NHB men 28.40 kg/m^2 ($p<2.22\text{x}10^{-308}$)
  - similar among Hispanics but almost absent in NHWs
- Increased BMI with age ($\beta=0.01$ [0.01], $p=4.49\text{x}10^{-28}$), in women versus men ($\beta=1.47$[0.04], $p<2.22\text{x}10^{-308}$), NHBs versus NHWs ($\beta=2.25$[0.05], $p<2.22\text{x}10^{-308}$) and Hispanics versus NHWs ($\beta=1.55$[0.05], $p=3.59\text{x}10^{-231}$)
- Significant race/ethnicity x gender interaction terms observed for all race/ethnic groups

Table 1: Characteristics of age, anthropometric traits and ALT levels by race/ethnicity and gender[a]

Race/Ethnicity and Gender	N	Age (years)	BMI (kg/m^2)	Prevalence of obesity[b]	Waist Circumference (cm)	Hip Circumference (cm)	WHR	N[c]	ALT (IU/L)
NHW Women	53,771	54.1 (16.9)	29.09 (7.69)	0.37	91.34 (17.9)	108.6 (16.2)	0.839 (0.094)	17382	22.5 (14.34)
NHW Men	35,060	57.1 (16.8)	29.02 (6.14)	0.33	101.3 (16.2)	106.5 (12.4)	0.95 (0.099)	9356	27.3 (16.7)
NHB Women	23,432	49.8 (14.8)	33.32 (9.11)	0.58	100.5 (18.6)	115.4 (18)	0.871 (0.1)	6149	19.3 (12.8)
NHB Men	17,804	50.9 (13.7)	28.4 (7.13)	0.31	96.94 (17.7)	105.2 (14.1)	0.92 (0.11)	2686	25.1 (16.61)
Hispanic Women	25,432	44.4 (15.8)	30.98 (7.4)	0.49	94.83 (16.8)	109.8 (15.3)	0.863 (0.099)	7214	23.3 (17.36)
Hispanic Men	12,206	46 (15.8)	29.75 (6.69)	0.40	100.1 (16.6)	105.4 (12.9)	0.949 (0.098)	2571	29.5 (21.11)
Asian Women	3,472	42.8 (16.9)	24.52 (4.99)	0.12	79.3 (12.5)	96.34 (10.9)	0.822 (0.08)	840	21.3 (13.18)
Asian Men	2,243	44.8 (17.2)	26.12 (4.62)	0.15	89.7 (13.1)	99.43 (9.49)	0.902 (0.11)	429	28.9 (17.53)

ALT indicates alanine aminotransferase; BMI, body mass index; cm, centimeters; IU/L, international units per liter; NHB, Non-Hispanic Black; NHW, Non-Hispanic White; WHR, waist-to-hip ratio.

[a] Continuous variables are presented as mean (standard deviation).

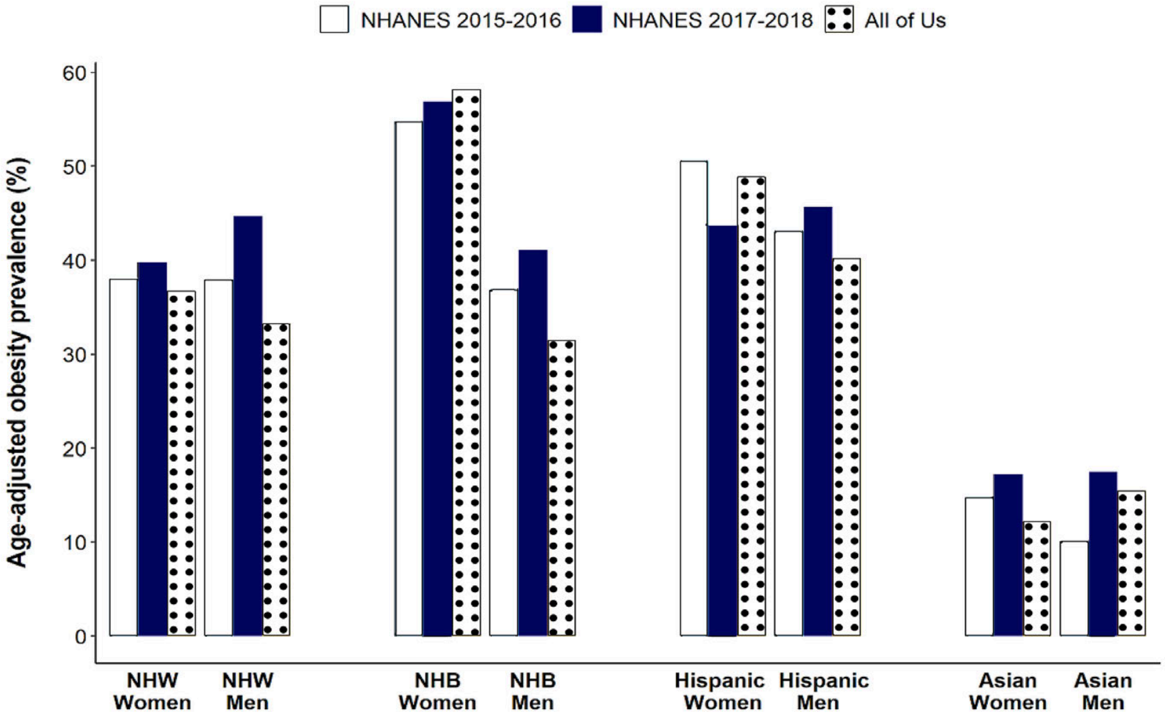
[b] Age-adjusted obesity prevalence.

[c] Number of individuals with available ALT data that were included form the AoU EHR dataset.

Age-adjusted prevalence of obesity in AoU participants

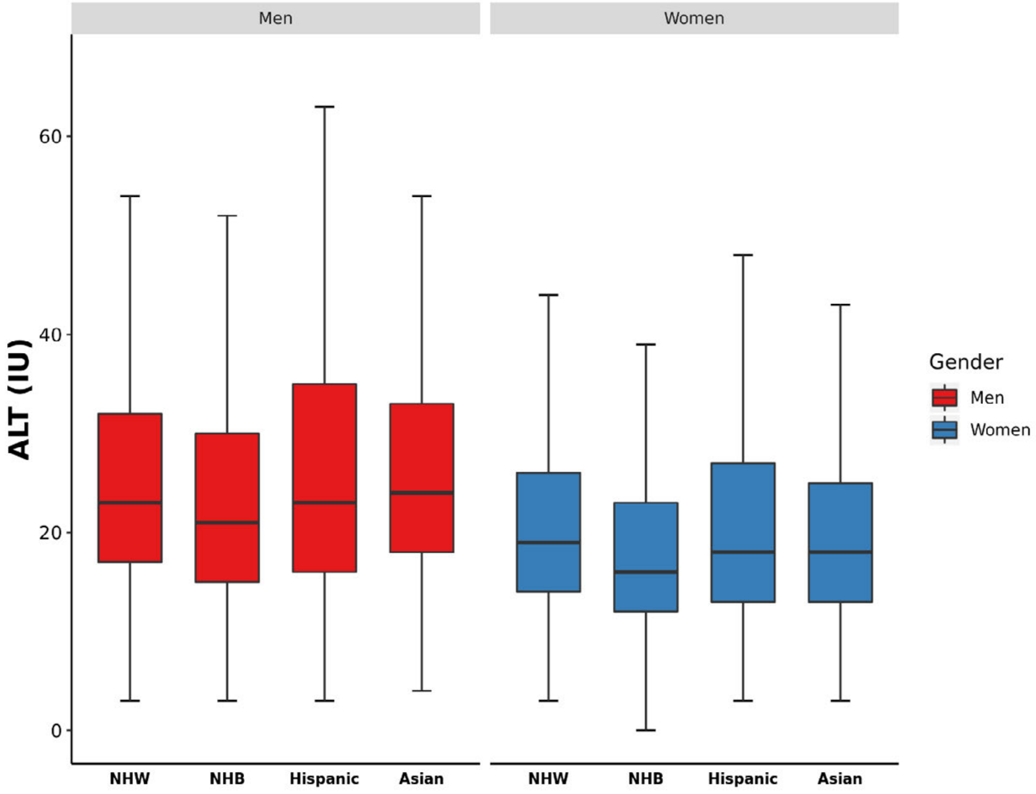
- Age-adjusted prevalence of obesity varied from 58% in NHB women to 12% in Asian women

- Compared to NHANES data, AoU participants had similar but lower obesity prevalence in all race/ethnic and gender groups with the exception of NHB women



ALT levels in AoU participants by gender and race/ethnicity

- Men exhibited higher levels of ALT than women in all racial/ethnic groups
- In both men and women, NHB exhibited lower levels of ALT compared to all three other ethnic/racial groups
- ALT decreased with increasing age (β=-0.08[0.01], p=$7.98\text{x}10^{-47}$), in women versus men (β =-6.18[0.18], p=$2.38\text{x}10^{-268}$), and in NHBs versus NHWs (β =-3.93[0.22], p=$1.40\text{x}10^{-72}$).
- ALT increased with increasing BMI (β=0.21 [0.01], p=$5.47\text{x}10^{-85}$)
- Significant race/ethnicity x gender interaction terms were also observed for all race/ethnic groups



CONCLUSIONS and DISCUSSION

- Replicate known racial/ethnic and gender disparities related to obesity
    - Gender differences in obesity very pronounced in NHB due to genetic factors as well as social and economic factors
    - NHB women have lower ALT despite higher BMI: “metabolically healthy” obesity phenotype?
- Our data suggest AoU averages representative of national averages related to obesity
    - Subtle healthy participant bias (age-adjusted obesity prevalence generally, but not consistently, lower than NHANES)
- Sample size, diversity, availability of PPI and EHR data, suggest AoU will be major resource for future work in racial/ethnic disparities
    - whole genome sequence data, specimens, and the ability to re-contact also likely to bolster AoU’s scientific utility
- Limitations: no exclusion of pregnancy; did not account for comorbid disease or drug treatments for obesity; patchwork of datasets, creating inconsistency in availability of data

Workbench Considerations for AoU data

- Cohort builder: saving derived datasets was initially challenging, but now works quite well
- Link between datasets and contents of the data browser contents is quite good
  - jupyter notebook system works well
  - Loading datasets into jupyter is quicker after “BigQuery” client library
  - Sometimes “BigQuery” cannot detect cohort data
- We encountered some issues of data consistency with same data from multiple sources (i.e. height and weight): recommend cross-checking & plotting data
- Feedback on Data
  - Data are (relatively) clean and standardized (terminologies, OMOP)
  - Would be nice to more easily query and analyze program vs EHR data
- Feedback on Tools
  - Cohort builder, Concept sets, and overall workbench mimic the research process
  - Visualizations are generally helpful, but need more cleaner versions that are interpretable
- DRC has been responsive to our questions

Scientific Considerations for AoU and EHR data

- Missing data and outliers
- Time stamps for labs, diagnoses, drug exposures, etc.
- Patchwork of datasets
  - Different sources of different data sets (AoU visit, surveys, EHR, etc)
  - Different n’s for different data sets
  - Some patients have decades of EHR data, some patients only have AoU enrollment data
  - Disparate clinical and billing practices between sites
  - Datasets will be periodically updated
- Using cohort builder for association, prevalence, incidence studies
  - Denominator problem when creating case and control datasets separately
  - Systematic differences in availability of data?
- Considerations for privacy – availability of tier dataset
  - n<20 and outliers
- Healthy participant bias?

Acknowledgements:

- AoU Data and Research Center (DRC), led by Vanderbilt University Medical Center
- National Institutes of Health, Office of the Director: Regional Medical Centers: 1 OT2 OD026549; 1 OT2 OD026554; 1 OT2 OD026557; 1 OT2 OD026556; 1 OT2 OD026550; 1 OT2 OD 026552; 1 OT2 OD026553; 1 OT2 OD026548; 1 OT2 OD026551; 1 OT2 OD026555; IAA #: AOD 16037; Federally Qualified Health Centers: HHSN 263201600085U; Data and Research Center: 5 U2C OD023196; Biobank: 1 U24 OD023121; The Participant Center: U24 OD023176; Participant Technology Systems Center: 1 U24 OD023163; Communications and Engagement: 3 OT2 OD023205; 3 OT2 OD023206; and Community Partners: 1 OT2 OD025277; 3 OT2 OD025315; 1 OT2 OD025337; 1 OT2 OD025276. National Heart, Lung, and Blood Institutes (R01-HL136528).
- JHK is supported by grants from the NIH's National Heart Lung and Blood Institute (NHLBI) under award number K01HL143137, the American College of Clinical Pharmacy, and the Flinn Foundation.
- Participants of the All of Us Research Program.

Health Outcomes of Sexual and Gender Minority People: Data from the *All of Us* Research Program

Mitchell R. Lunn, MD, MAS
Stanford University School of Medicine
All of Us Researcher

PRIDEnet @ Stanford University

PRIDEnet @ Stanford is a national community engagement partner for *All of Us*.

Activities:

- Engage sexual and gender minority (SGM; *i.e.*, LGBTQ+) people to participate in AoU
- Train the AoU Consortium to be affirming in their interactions with SGM people
- Engage external SGM researchers in utility of AoU data for health disparities research

Team members:

- Kate Vitale, PhD
- Haley Hedlin, PhD
- Daryl Mangosing, MPH
- Annesa Flentje, PhD
- Micah Lubensky, PhD
- Carolyn Hunt, MPA
- Mahri Bahati, MPH
- Ana Rescate, MBA
- Zubin Dastur, MS, MPH
- Kelly Gebo, MD, MPH
- Juno Obedin-Maliver, MD, MPH, MA
- Mitchell R. Lunn, MD, MAS
- Demonstration Projects Subcommittee
- PRIDEnet Participant Advisory Committee
  - …on behalf of the *All of Us* Research Program investigators

Background

- Sexual and gender minority (SGM) people face numerous health disparities compared to their heterosexual (*i.e.,* straight) and cisgender (*i.e.,* not transgender) peers.
  - E.g. Increased rate of
    - Tobacco use
    - Certain cancers
    - Depression
    - HIV infection
- The diversity of All of Us data makes it an excellent resource for studying SGM health disparities.
- Unprecedented collection of sexual orientation and gender identity data.

Research Questions

- What are the health conditions/behaviors of sexual and gender minority (SGM) participants compared with non-SGM participants in the All of Us Research Program?
    - Are previously-reported health disparities reproduced in the *All of Us* data?
    - What are the sociodemographic and health characteristics of sexual minority (SM) and gender minority (GM) subgroups in *All of Us*?

Methods

Sociodemographics:

- Using “The Basics” PPI module [data are generalized], we categorized participants as SGM or non-SGM, and subdivided SGM participants into four subgroups:

	Sex assigned at birth (SAAB)	Sexual orientation (SO)	Gender Identity (GI)	N (%)
Cisgender sexual minority men	Male	Non-straight or Unknown	Man	9,520
Cisgender sexual minority women	Female	Non-straight or Unknown	Woman	13,517
Gender minority people with female sex assigned at birth	Female	Any	Man or Other/unknown	1,685
Gender minority people with male sex assigned at birth	Male	Any	Woman or Other/unknown	2,288
Total				**27,010**

- We examined each group’s breakdown of responses to sexual orientation, gender identity, and sex assigned at birth and other demographic questions:
  - ***- Age - Race - Ethnicity - Geography - Educational Attainment - Employment Status - Annual Household Income - Insurance Coverage***
- We characterized the size of overall difference between SGM and non-SGM participants on each demographic variable using the absolute standardized difference statistic (ASD).

Results

Sociodemographics:

Demographic variable	Overall SGM tendency (relative to non-SGM)	ASD
Age	Younger	**0.35**
Race	Slightly more people of color	**0.19**
Ethnicity	Slightly more people of Hispanic ethnicity	0.15
Geography	Similar distribution across US	0.04
Education	Slightly lower educational attainment	**0.23**
Employment	Similar proportion employed	0.16
Income	Lower income	**0.31**
Insurance	Insured at a slightly lower frequency	0.17

Absolute Standardized Difference (ASD)

interpretation guidelines:

0.2: small difference
0.5: moderate difference
0.8: large difference

Methods

Health Conditions and Behaviors:

- We characterized the prevalence of the following health conditions among sexual minority (SM) and gender minority (GM) participants using EHR data and physical measurement data (PM&B):
  - ***Anxiety***
  - ***Asthma***
  - ***Cancer***
  - ***Cardiovascular disease***
  - ***Chronic kidney disease***
  - ***Depression***
  - ***Diabetes mellitus***
  - ***HIV infection***
  - ***Hypertension***
  - ***Tobacco use***
  - ***Substance use disorder***
  - ***Body mass index >25.0***

Methods

Health Conditions and Behaviors:

- Condition required one SNOMED CT code instance on at least two different dates.
- Logistic regression models were used to generate an odds ratio for each health condition/behavior in each SM and GM subgroup using an appropriate non-SGM reference group.
    - Due to generalized data, GM groups contained transgender as well as gender-expansive (outside female-male binary). We compared these groups to cisgender women ***and*** cisgender men.
- Models were adjusted for age, race, ethnicity, and income

Results

Health Conditions and Behaviors:

	Cisgender sexual minority men (N=5208) aOR
Reference group	*Straight cisgender men*
Anxiety	1.51 (1.39 - 1.63)
Asthma	1.18 (1.06 - 1.31)
Depression	1.74 (1.62 - 1.87)
HIV infection	15.26 (13.46 - 17.31)
Tobacco use	1.20 (1.07 - 1.34)
Substance use disorder	1.15 (1.05 - 1.27)

Literature estimates

→ **2.7 (1.8, 4.1)**
(aOR, Bostwick 2010)

→ **1.16**
(est OR, CDC 2018)

Results are expressed as odds ratios (aOR) compared to the indicated reference group and adjusted for age, race, ethnicity, and income

Conditions not significantly increased:
Cancer, cardiovascular disease, chronic kidney disease, diabetes mellitus, hypertension, overweight/obesity

Results

Health Conditions and Behaviors:

	Gender minority people assigned female sex at birth (N=1180) aOR
Reference group	*Straight cisgender women*
Anxiety	1.20 (1.04 - 1.37)
Asthma	1.21 (1.04 - 1.42)
Chronic kidney disease	1.46 (1.14 - 1.88)
Depression	1.20 (1.05 - 1.37)
HIV infection	1.68 (1.07 - 2.64)
Overweight / obese	1.24 (1.08 - 1.42)
Substance use disorder	1.34 (1.06 - 1.70)

Results are expressed as odds ratios (aOR) compared to the indicated reference group and adjusted for age, race, ethnicity, and income.

<u>Conditions not significantly different</u>:
Cancer, cardiovascular disease, diabetes mellitus, hypertension, tobacco use

Results

Health Conditions & Behaviors:

	Gender minority people assigned female sex at birth (N=1180) aOR
Reference group	*Straight cisgender men*
Anxiety	2.15 (1.87 - 2.47)
Asthma	2.47 (2.11 - 2.90)
Depression	2.09 (1.82 - 2.39)
Overweight / obese	1.16 (1.01 - 1.33)

Results are expressed as odds ratios (aOR) compared to the indicated reference group and adjusted for age, race, ethnicity, and income.

Conditions not significantly increased:
Cancer, cardiovascular disease, chronic kidney disease, diabetes mellitus, HIV infection, hypertension, tobacco use

Conclusions

- Previously reported health disparities between SGM and non-SGM participants were often replicated in the *All of Us* data.
- SGM subgroups show distinct health disparities and uncovered health disparities that were not previously assessed/reported.
- SGM and non-SGM participants show small demographic differences

Limitations

- Data generalizations
- Non-random sampling
- Observational data
- Demographic differences
- Missing data

Implications

- This demonstration project adds to a body of literature characterizing health disparities for SGM individuals.
- Such findings support the prioritization of research areas for tackling SGM health disparities.
- This work also demonstrates the utility of the *All of Us* data for SGM (and likely many other UBR communities) research.

Lessons Learned / Next Steps

Valuable Skills:

- Understanding the OMOP data model
- Writing SQL queries
- Missing data analysis approaches

How interacted with support:

- Kick-off meeting
- Community support boards
- Help request tickets
- Help hours
- Email

Awareness Points:

- All of Us-specific data curation details
- EHR data coverage
- Other non-obvious data limitations (missing data, data generalizations, constructs being measured)
- Computing limitations

Next steps:

- Repeat analyses with non-generalized data (aka controlled tier) to provide more granular descriptions of sexual orientations and gender identities

Wrap Up Remarks

Claire Schulkey, Ph.D.
All of Us Data Scientist

A HUGE thank you to the Researcher Onramp Team for organizing this event!

The Team

Melissa Bassford
Margaret Farrell
Minnkyong Lee
Roxana Loperena-Cortez
Kelsey Mayo
Andrea Ramirez
Raquel Rivas
Adrienne Roman
Claire Schulkey
Sheri Schully
Katrina Theisz
Courtney Wallin
Joann Smith

Thank You!

Share info about the *All of Us* Researcher Workbench

Share this recording when it becomes available

Join our researcher community!

Please return by **2:25 pm ET** if you signed up for the *All of Us* Researcher Onboarding Walkthrough

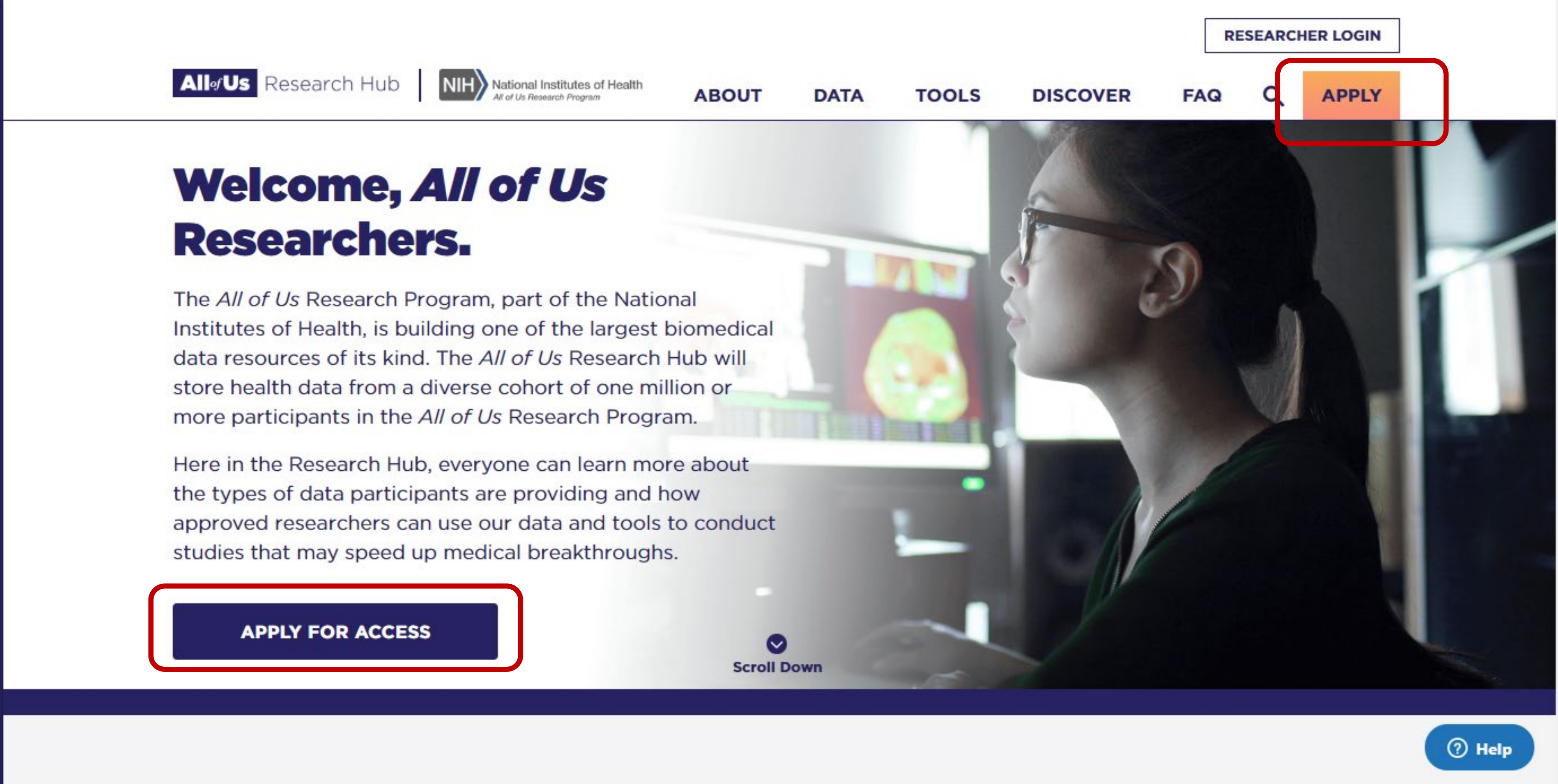
Please check your email for breakout session information

All of Us Researcher Onboarding Walkthrough

This breakout session will cover:

- What it means to be a beta researcher
- Current requirements for applying for access
- Institutional Agreements
  - Does your U.S.-based academic institution, healthcare institution, or nonprofit have one in place already?
  - How to get process started if your institution does not have one
- Applying for access through Research Hub website at researchallofus.org

researchallofus.org

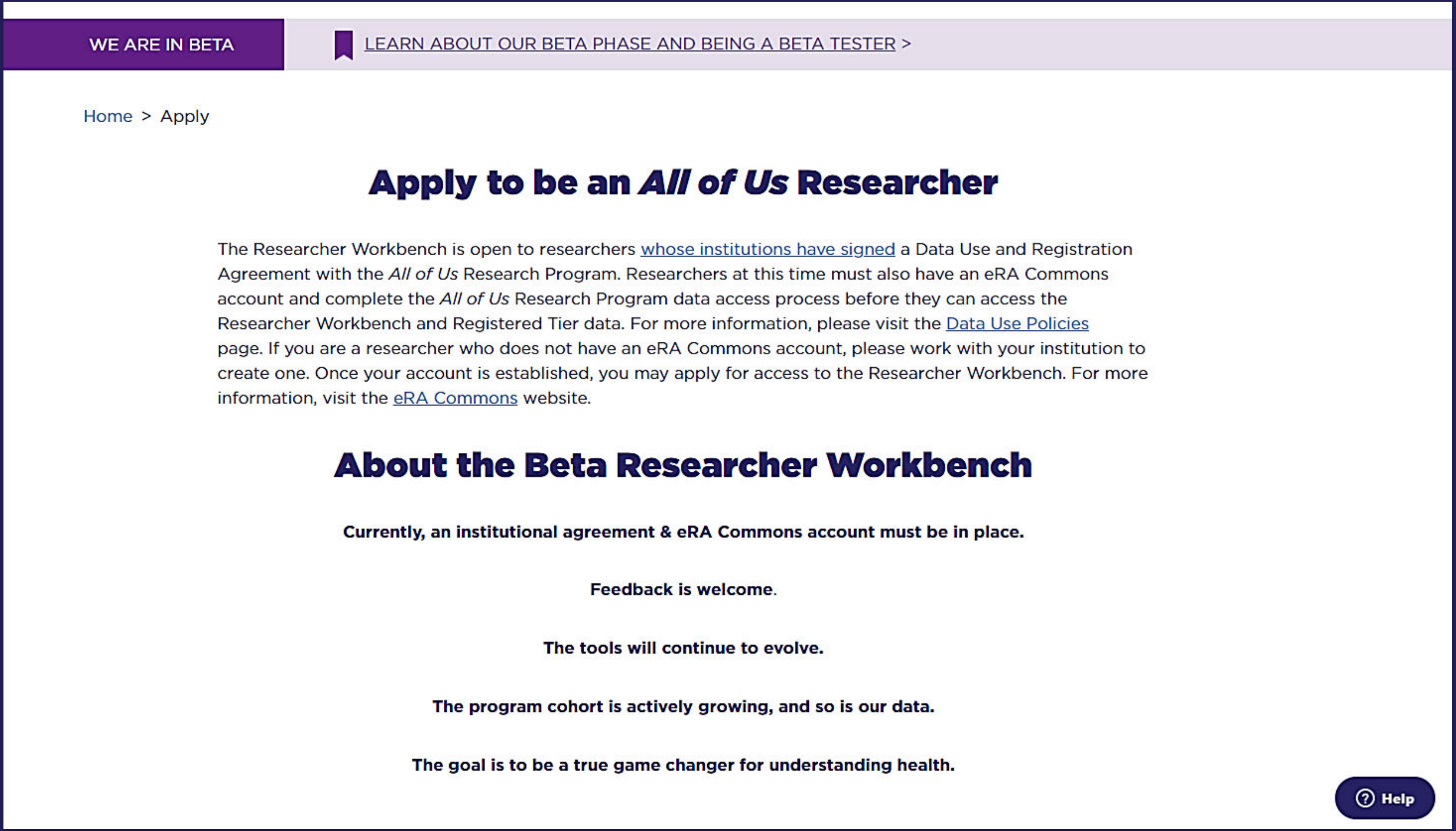


Welcome, *All of Us* Researchers.

The *All of Us* Research Program, part of the National Institutes of Health, is building one of the largest biomedical data resources of its kind. The *All of Us* Research Hub will store health data from a diverse cohort of one million or more participants in the *All of Us* Research Program.

Here in the Research Hub, everyone can learn more about the types of data participants are providing and how approved researchers can use our data and tools to conduct studies that may speed up medical breakthroughs.

researchallofus.org



Apply to be an *All of Us* Researcher

The Researcher Workbench is open to researchers whose institutions have signed a Data Use and Registration Agreement with the *All of Us* Research Program. Researchers at this time must also have an eRA Commons account and complete the *All of Us* Research Program data access process before they can access the Researcher Workbench and Registered Tier data. For more information, please visit the Data Use Policies page. If you are a researcher who does not have an eRA Commons account, please work with your institution to create one. Once your account is established, you may apply for access to the Researcher Workbench. For more information, visit the eRA Commons website.

About the Beta Researcher Workbench

Currently, an institutional agreement & eRA Commons account must be in place.

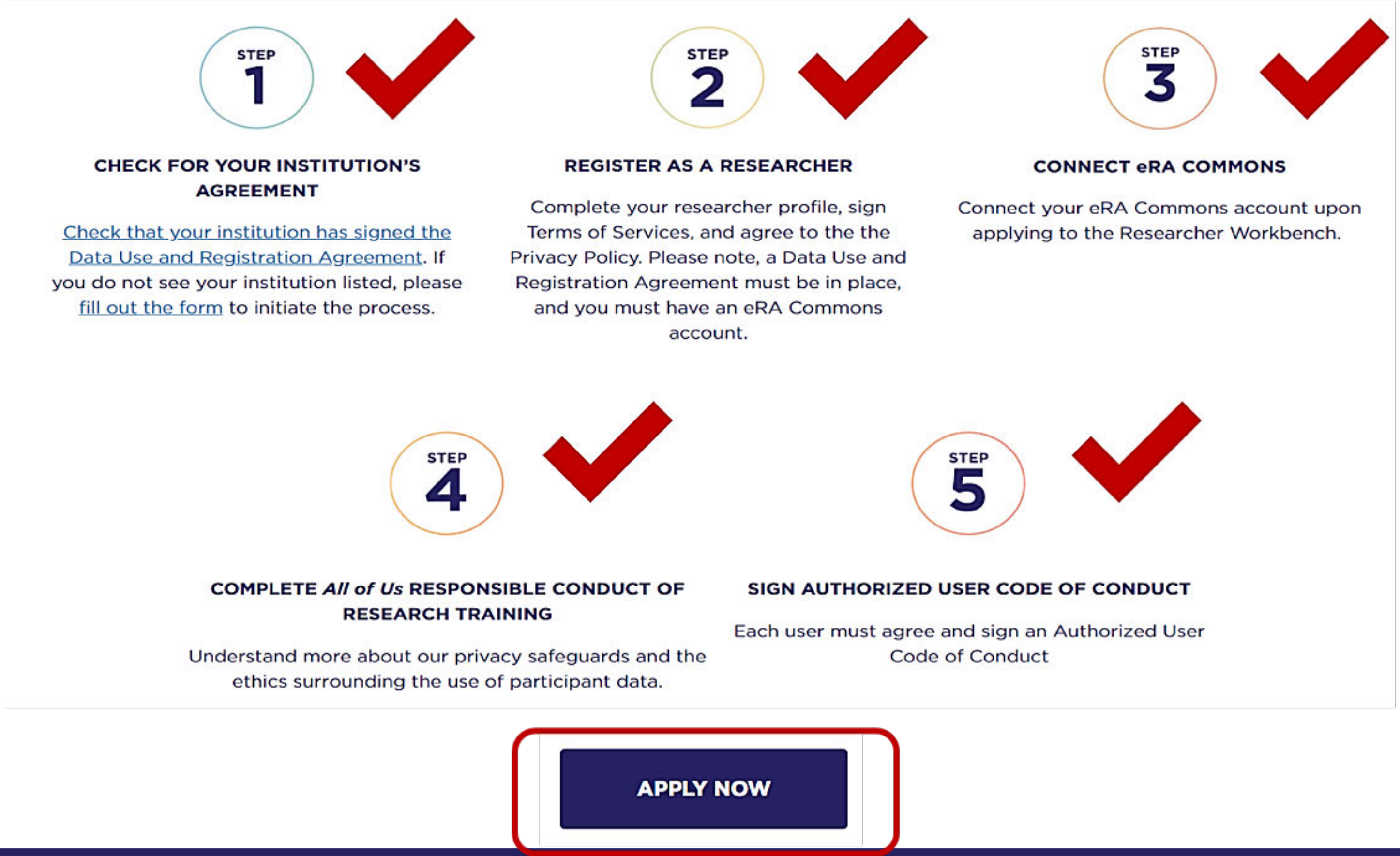
Feedback is welcome.

The tools will continue to evolve.

The program cohort is actively growing, and so is our data.

The goal is to be a true game changer for understanding health.

Complete the Application Process



STEP 1

CHECK FOR YOUR INSTITUTION'S AGREEMENT

Check that your institution has signed the Data Use and Registration Agreement. If you do not see your institution listed, please fill out the form to initiate the process.

STEP 2

REGISTER AS A RESEARCHER

Complete your researcher profile, sign Terms of Services, and agree to the the Privacy Policy. Please note, a Data Use and Registration Agreement must be in place, and you must have an eRA Commons account.

STEP 3

CONNECT eRA COMMONS

Connect your eRA Commons account upon applying to the Researcher Workbench.

STEP 4

COMPLETE *All of Us* RESPONSIBLE CONDUCT OF RESEARCH TRAINING

Understand more about our privacy safeguards and the ethics surrounding the use of participant data.

STEP 5

SIGN AUTHORIZED USER CODE OF CONDUCT

Each user must agree and sign an Authorized User Code of Conduct

Already have a Researcher Workbench account?

Don't have an account?

WARNING NOTICE

You are accessing a web site created by the *All of Us* Research Program, funded by the National Institutes of Health.

Unauthorized attempts to upload information, change information, or use of this web site may result in disciplinary action, civil, and/or criminal penalties. Unauthorized users of this website should have no expectation of privacy

All of Us Research Program

Researcher Workbench Terms of Use and Privacy Statement

The *All of Us* Research Program (the "***All of Us* Research Program**") is part of an effort to advance individualized healthcare by enrolling one million or more participants to contribute their health data and developing custom tools and software to help researchers analyze that data. The goal of the *All of Us* Research Program is to speed up health research discoveries, enabling new kinds of individualized health care. Among the resources developed in connection with the *All of Us* Research Program is the Researcher Workbench, which is a platform and collection of tools through which researchers can access registered tiered data, workspaces, and tools such as a cohort builder, and an interactive notebook (the "**Workbench**").

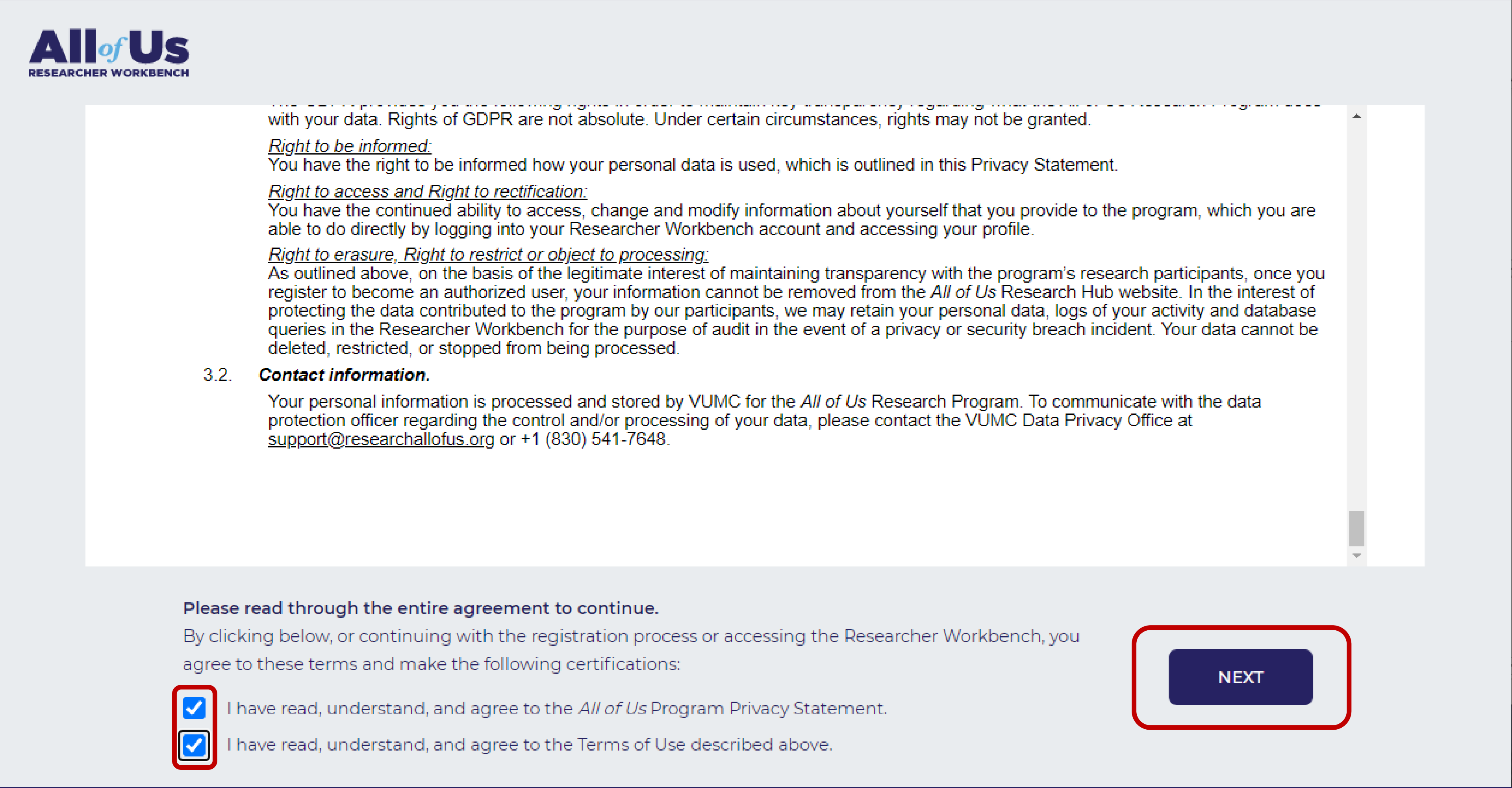
These Terms of Use ("**Terms**") describe the terms and conditions that govern how Vanderbilt University Medical Center ("**VUMC**," "**We**," or "**Us**") makes the Workbench available to users that meet the *All of Us* Research Program Workbench requirements, complete the registration process and are approved to access the Workbench ("**Users**"). These Terms, the *All of Us* Research Program Privacy Statement ("**Privacy Statement**") and any applicable Third Party Terms referenced below govern use of the Workbench by each User, including you ("**You**"). Additionally, the *All of Us* Research Program Data User Code of Conduct ("**Data User Code of Conduct**") describes how You can and cannot use the *All of Us* Research Program data, which includes any data that is made available through the Workbench.

Please read through the entire agreement to continue.

By clicking below, or continuing with the registration process or accessing the Researcher Workbench, you agree to these terms and make the following certifications:

☐ I have read, understand, and agree to the *All of Us* Program Privacy Statement.

☐ I have read, understand, and agree to the Terms of Use described above.



with your data. Rights of GDPR are not absolute. Under certain circumstances, rights may not be granted.

Right to be informed:
You have the right to be informed how your personal data is used, which is outlined in this Privacy Statement.

Right to access and Right to rectification:
You have the continued ability to access, change and modify information about yourself that you provide to the program, which you are able to do directly by logging into your Researcher Workbench account and accessing your profile.

Right to erasure, Right to restrict or object to processing:
As outlined above, on the basis of the legitimate interest of maintaining transparency with the program's research participants, once you register to become an authorized user, your information cannot be removed from the *All of Us* Research Hub website. In the interest of protecting the data contributed to the program by our participants, we may retain your personal data, logs of your activity and database queries in the Researcher Workbench for the purpose of audit in the event of a privacy or security breach incident. Your data cannot be deleted, restricted, or stopped from being processed.

3.2. ***Contact information.***

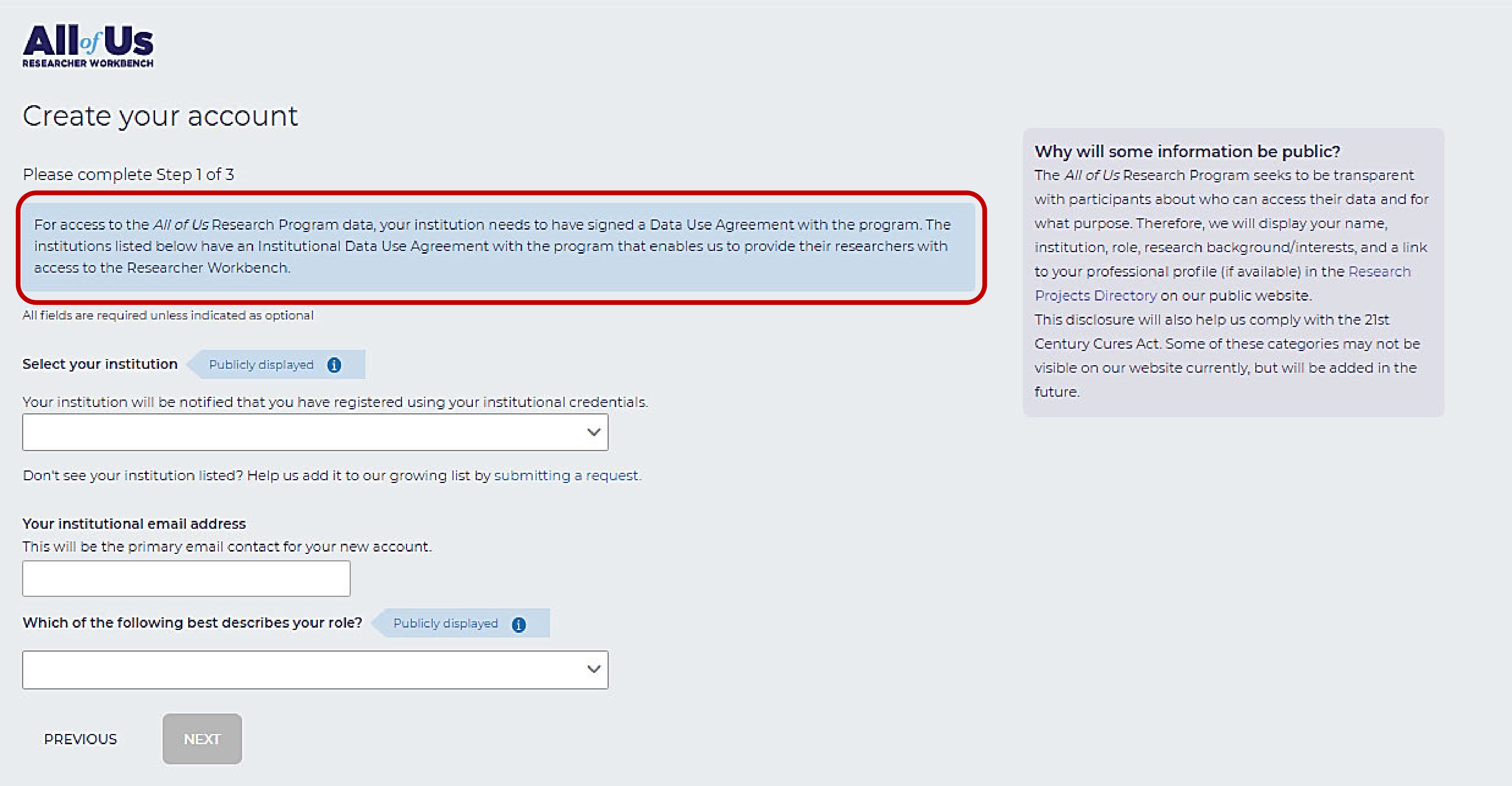
Your personal information is processed and stored by VUMC for the *All of Us* Research Program. To communicate with the data protection officer regarding the control and/or processing of your data, please contact the VUMC Data Privacy Office at support@researchallofus.org or +1 (830) 541-7648.

Please read through the entire agreement to continue.

By clicking below, or continuing with the registration process or accessing the Researcher Workbench, you agree to these terms and make the following certifications:

☑ I have read, understand, and agree to the *All of Us* Program Privacy Statement.

☑ I have read, understand, and agree to the Terms of Use described above.



Create your account

Please complete Step 1 of 3

For access to the *All of Us* Research Program data, your institution needs to have signed a Data Use Agreement with the program. The institutions listed below have an Institutional Data Use Agreement with the program that enables us to provide their researchers with access to the Researcher Workbench.

All fields are required unless indicated as optional

Select your institution

Your institution will be notified that you have registered using your institutional credentials.

Don't see your institution listed? Help us add it to our growing list by submitting a request.

Your institutional email address

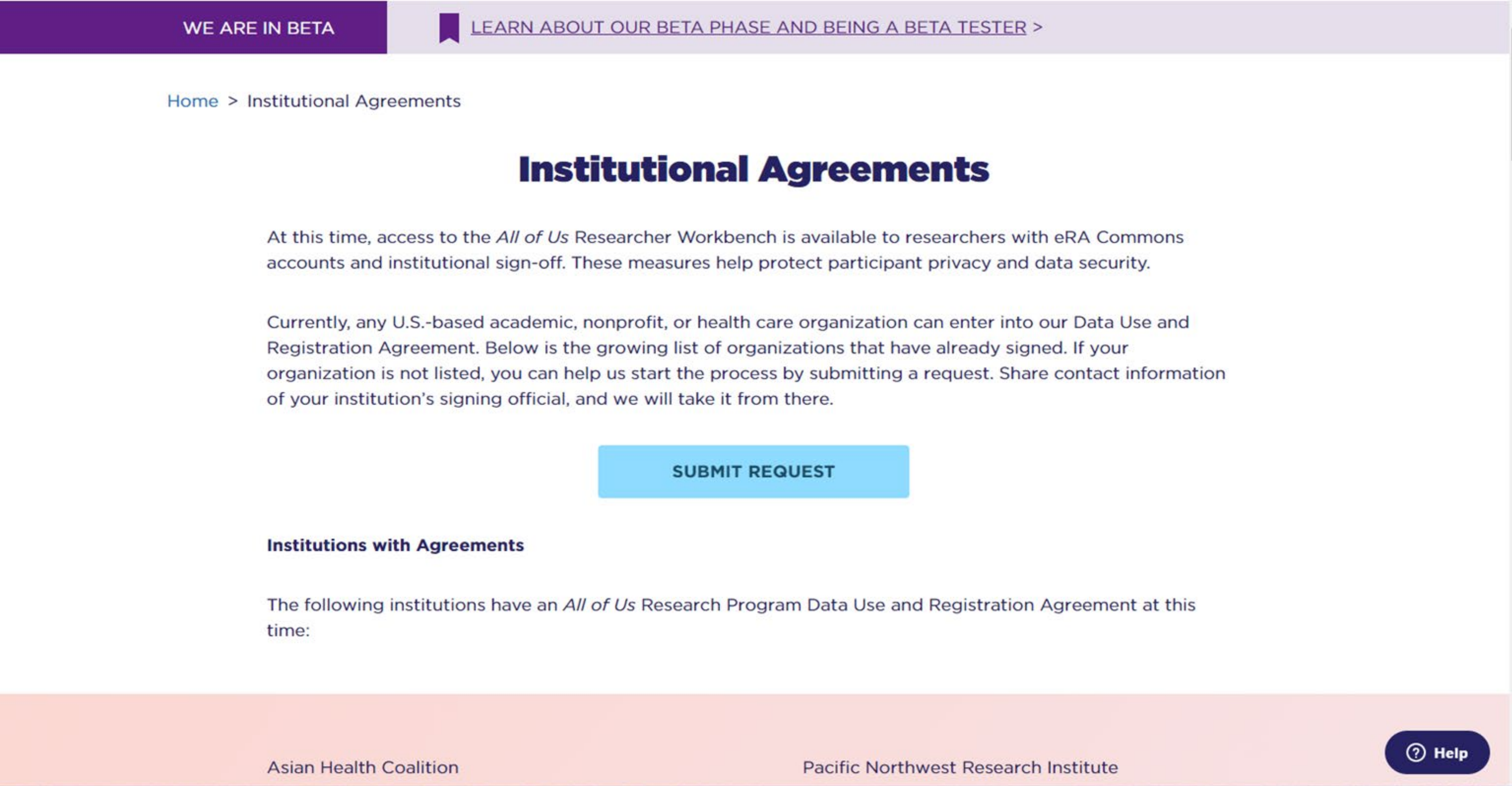
This will be the primary email contact for your new account.

Which of the following best describes your role?

Why will some information be public?

The *All of Us* Research Program seeks to be transparent with participants about who can access their data and for what purpose. Therefore, we will display your name, institution, role, research background/interests, and a link to your professional profile (if available) in the Research Projects Directory on our public website.

This disclosure will also help us comply with the 21st Century Cures Act. Some of these categories may not be visible on our website currently, but will be added in the future.



Institutional Agreements

At this time, access to the *All of Us* Researcher Workbench is available to researchers with eRA Commons accounts and institutional sign-off. These measures help protect participant privacy and data security.

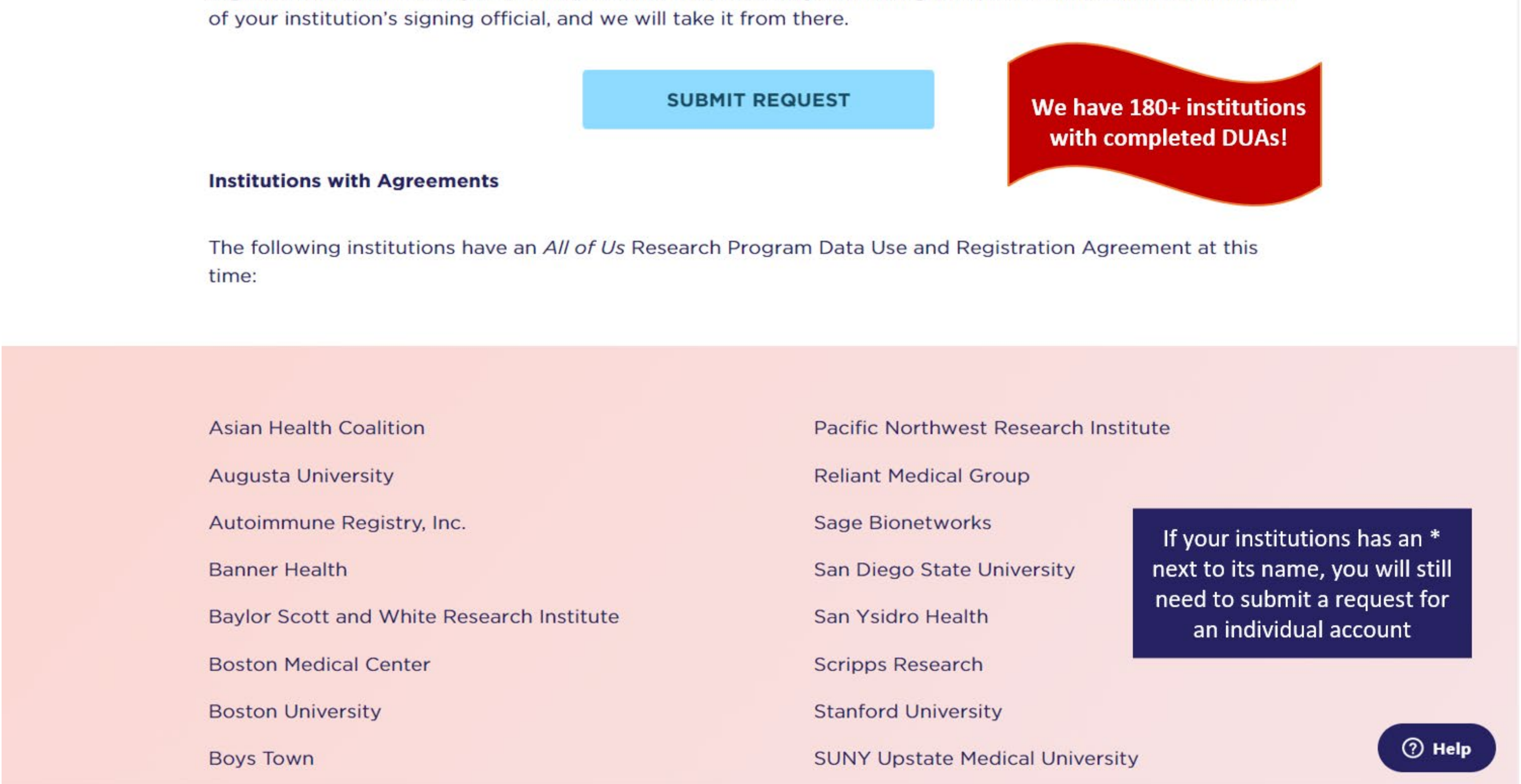
Currently, any U.S.-based academic, nonprofit, or health care organization can enter into our Data Use and Registration Agreement. Below is the growing list of organizations that have already signed. If your organization is not listed, you can help us start the process by submitting a request. Share contact information of your institution's signing official, and we will take it from there.

Institutions with Agreements

The following institutions have an *All of Us* Research Program Data Use and Registration Agreement at this time:

Asian Health Coalition

Pacific Northwest Research Institute



organization is not listed, you can help us start the process by submitting a request. Share contact information of your institution's signing official, and we will take it from there.

We have 180+ institutions with completed DUAs!

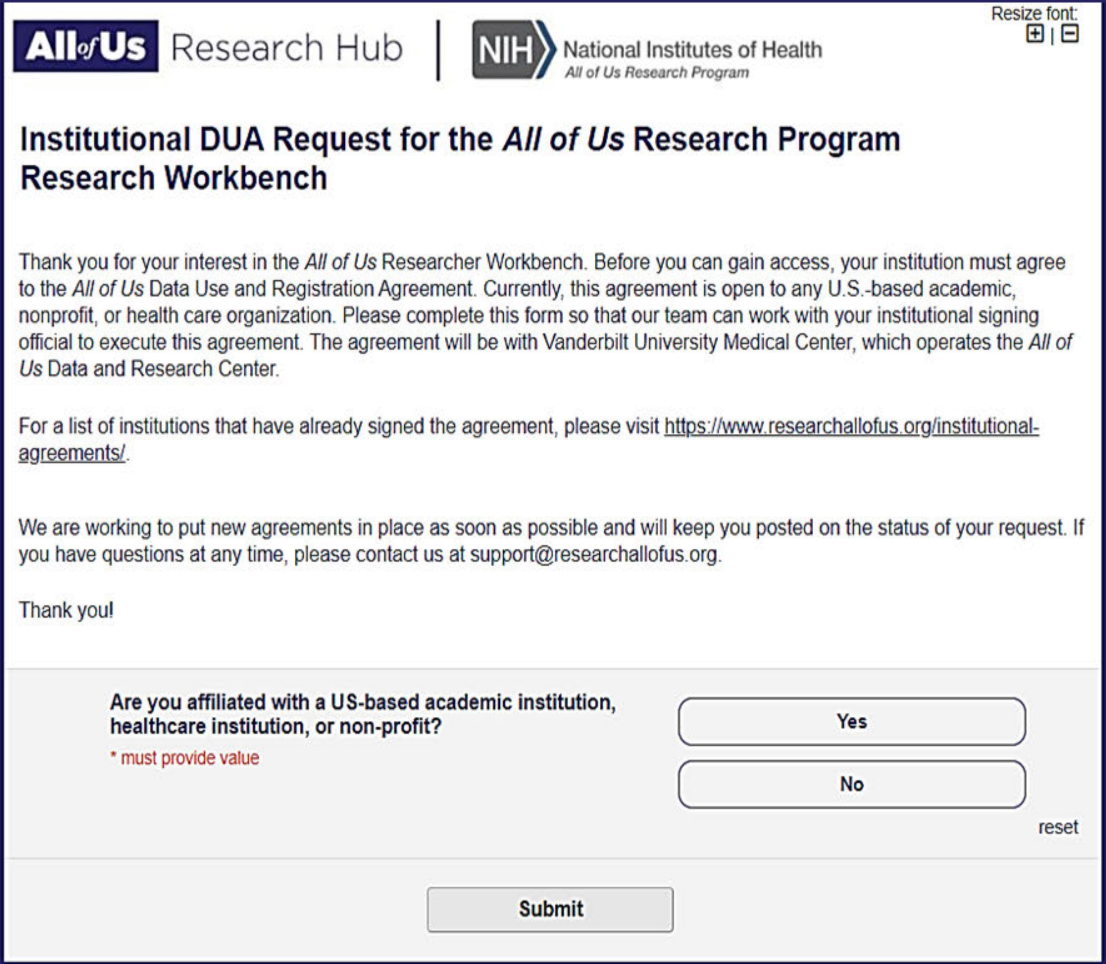
Institutions with Agreements

The following institutions have an *All of Us* Research Program Data Use and Registration Agreement at this time:

Asian Health Coalition
Augusta University
Autoimmune Registry, Inc.
Banner Health
Baylor Scott and White Research Institute
Boston Medical Center
Boston University
Boys Town

Pacific Northwest Research Institute
Reliant Medical Group
Sage Bionetworks
San Diego State University
San Ysidro Health
Scripps Research
Stanford University
SUNY Upstate Medical University

If your institutions has an * next to its name, you will still need to submit a request for an individual account



Institutional DUA Request for the *All of Us* Research Program Research Workbench

Thank you for your interest in the *All of Us* Researcher Workbench. Before you can gain access, your institution must agree to the *All of Us* Data Use and Registration Agreement. Currently, this agreement is open to any U.S.-based academic, nonprofit, or health care organization. Please complete this form so that our team can work with your institutional signing official to execute this agreement. The agreement will be with Vanderbilt University Medical Center, which operates the *All of Us* Data and Research Center.

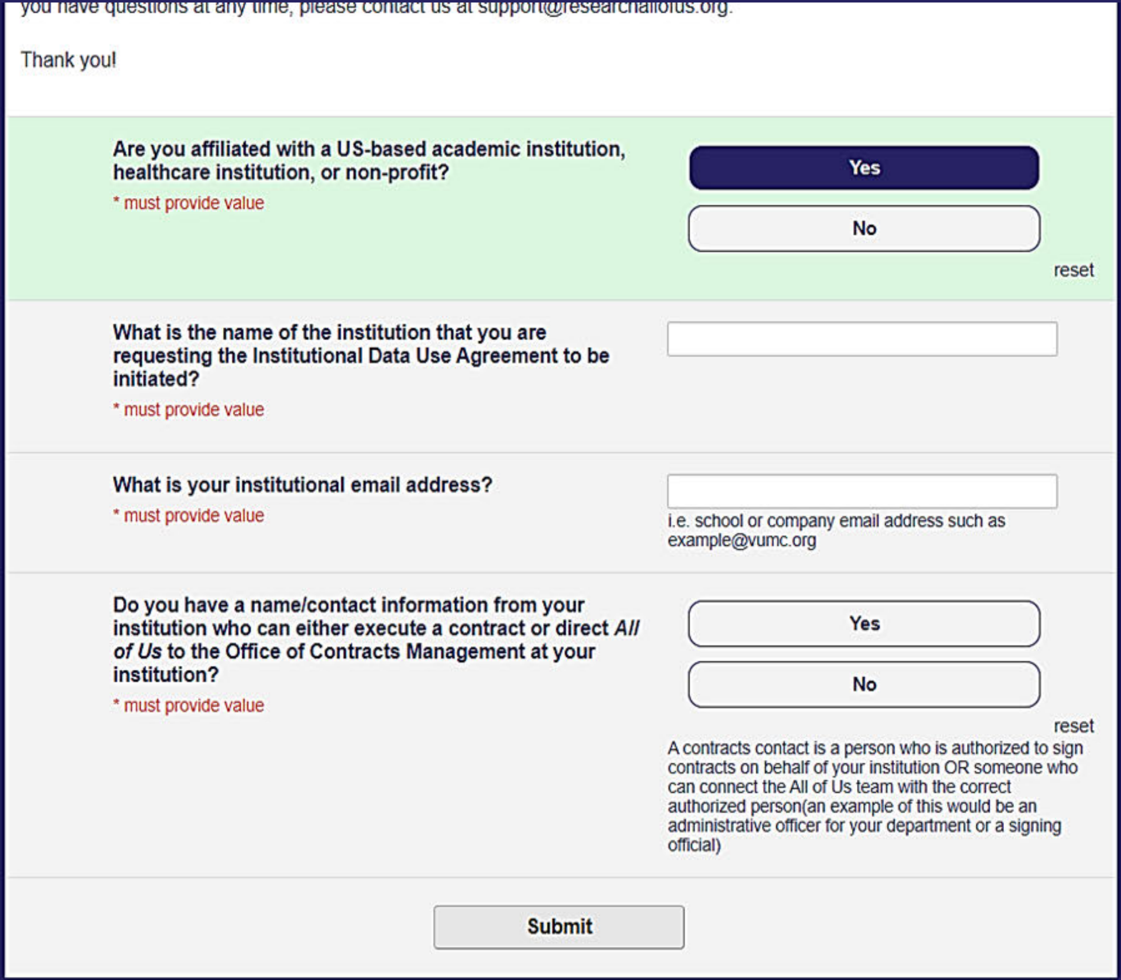
For a list of institutions that have already signed the agreement, please visit https://www.researchallofus.org/institutional-agreements/.

We are working to put new agreements in place as soon as possible and will keep you posted on the status of your request. If you have questions at any time, please contact us at support@researchallofus.org.

Thank you!

Are you affiliated with a US-based academic institution, healthcare institution, or non-profit?

* must provide value



you have questions at any time, please contact us at support@researchallofus.org.

Thank you!

Are you affiliated with a US-based academic institution, healthcare institution, or non-profit?

* must provide value

[✓ Yes] No

What is the name of the institution that you are requesting the Institutional Data Use Agreement to be initiated?

* must provide value

What is your institutional email address?

* must provide value

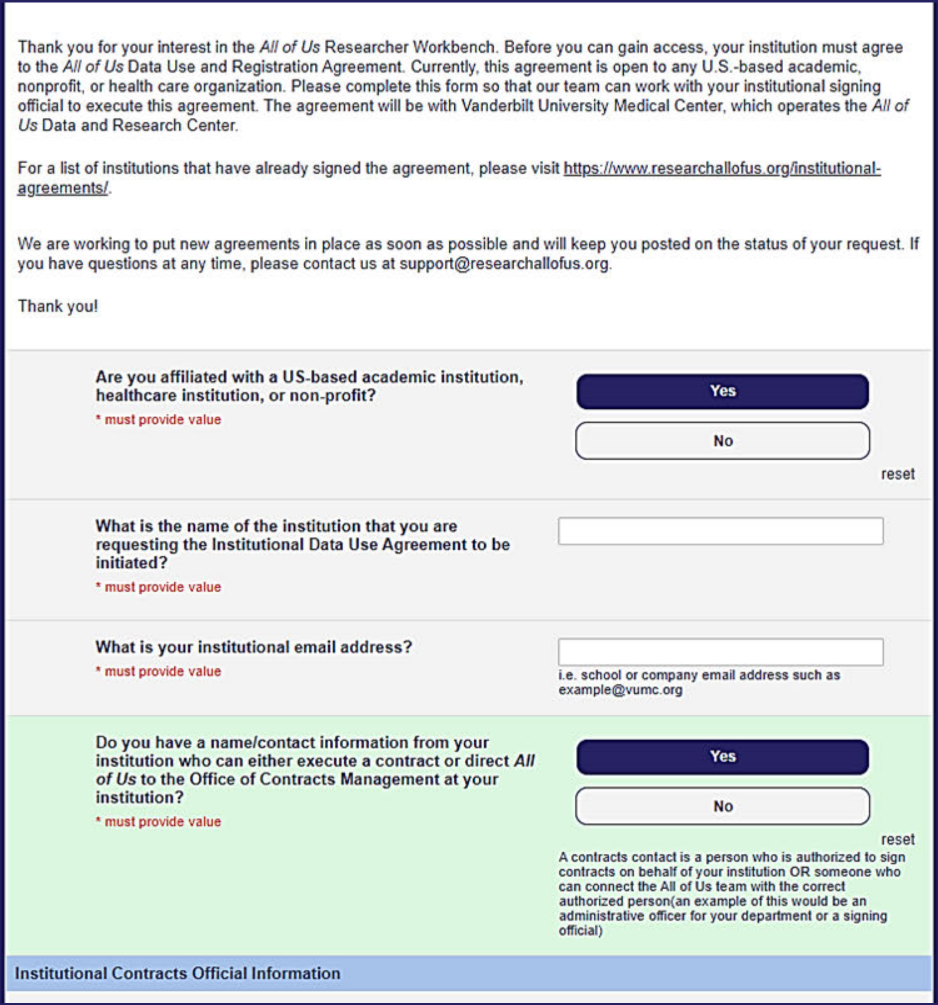
i.e. school or company email address such as example@vumc.org

Do you have a name/contact information from your institution who can either execute a contract or direct *All of Us* to the Office of Contracts Management at your institution?

* must provide value

Yes No

A contracts contact is a person who is authorized to sign contracts on behalf of your institution OR someone who can connect the All of Us team with the correct authorized person(an example of this would be an administrative officer for your department or a signing official)



Thank you for your interest in the *All of Us* Researcher Workbench. Before you can gain access, your institution must agree to the *All of Us* Data Use and Registration Agreement. Currently, this agreement is open to any U.S.-based academic, nonprofit, or health care organization. Please complete this form so that our team can work with your institutional signing official to execute this agreement. The agreement will be with Vanderbilt University Medical Center, which operates the *All of Us* Data and Research Center.

For a list of institutions that have already signed the agreement, please visit https://www.researchallofus.org/institutional-agreements/.

We are working to put new agreements in place as soon as possible and will keep you posted on the status of your request. If you have questions at any time, please contact us at support@researchallofus.org.

Thank you!

Are you affiliated with a US-based academic institution, healthcare institution, or non-profit?

* must provide value

What is the name of the institution that you are requesting the Institutional Data Use Agreement to be initiated?

* must provide value

What is your institutional email address?

* must provide value

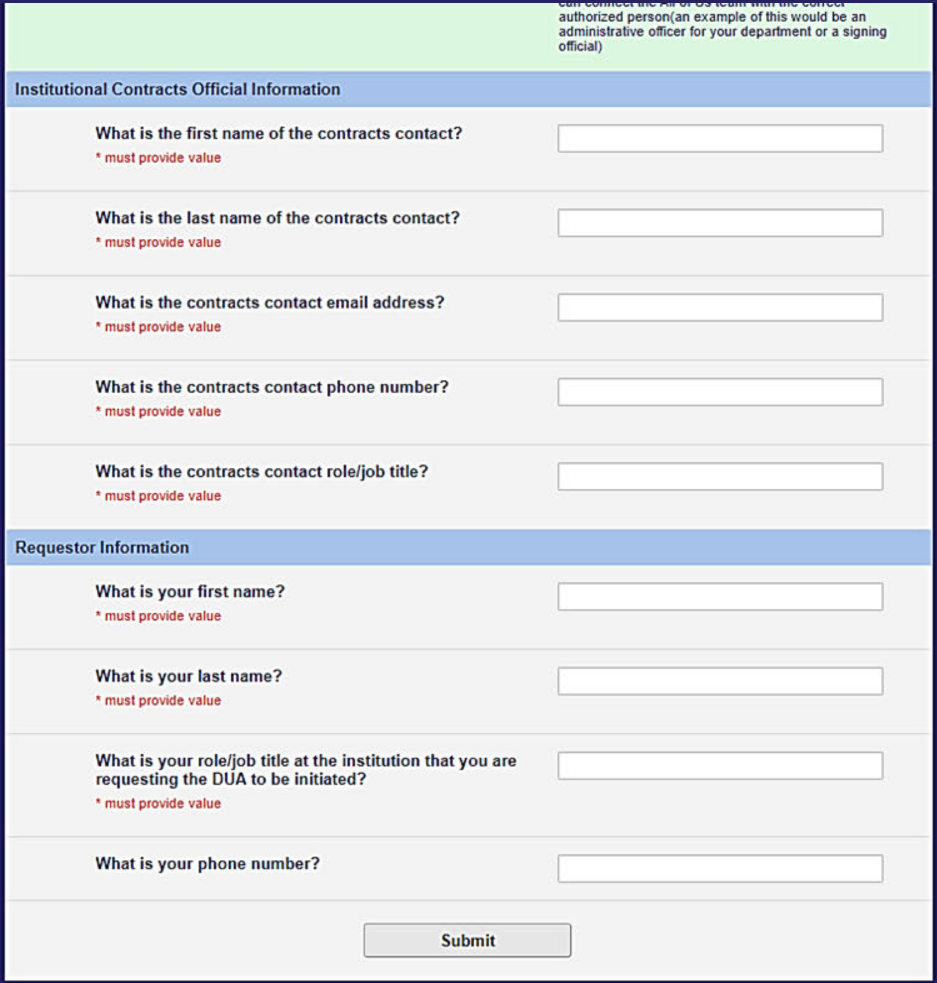
i.e. school or company email address such as example@vumc.org

Do you have a name/contact information from your institution who can either execute a contract or direct *All of Us* to the Office of Contracts Management at your institution?

* must provide value

A contracts contact is a person who is authorized to sign contracts on behalf of your institution OR someone who can connect the All of Us team with the correct authorized person(an example of this would be an administrative officer for your department or a signing official)

Institutional Contracts Official Information



authorized person(an example of this would be an administrative officer for your department or a signing official)

Institutional Contracts Official Information

What is the first name of the contracts contact?
* must provide value

What is the last name of the contracts contact?
* must provide value

What is the contracts contact email address?
* must provide value

What is the contracts contact phone number?
* must provide value

What is the contracts contact role/job title?
* must provide value

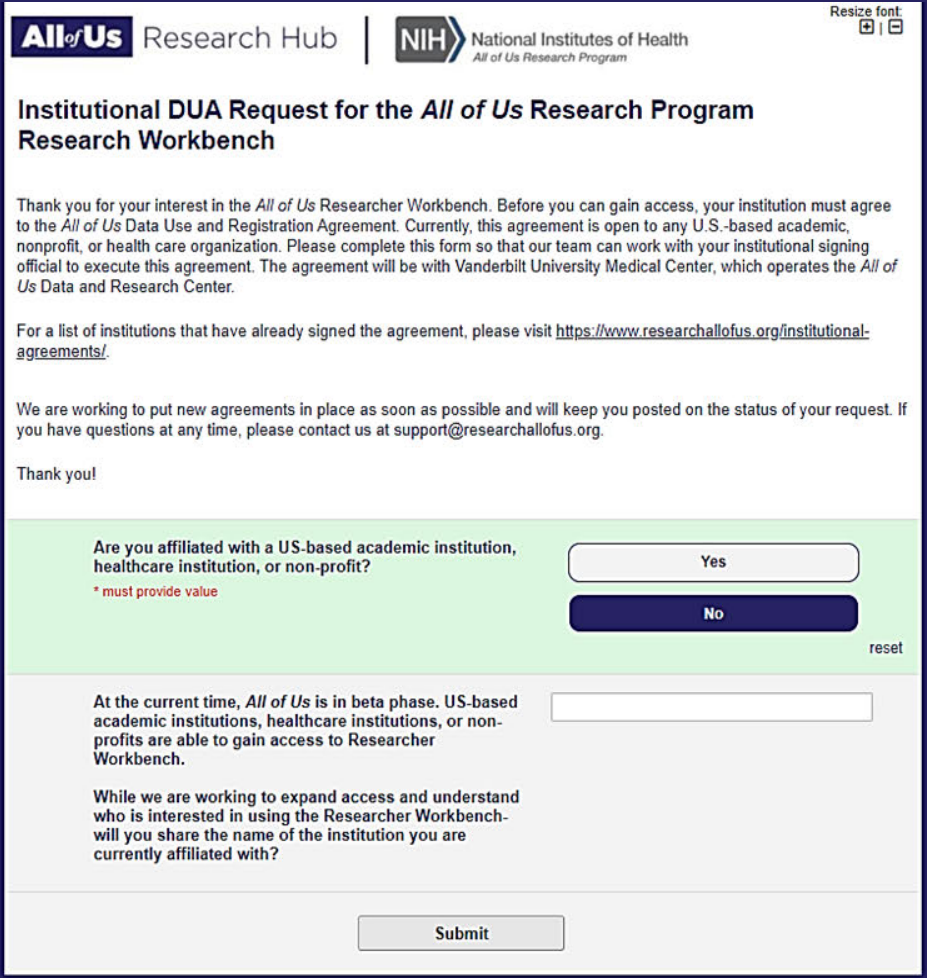
Requestor Information

What is your first name?
* must provide value

What is your last name?
* must provide value

What is your role/job title at the institution that you are requesting the DUA to be initiated?
* must provide value

What is your phone number?



Institutional DUA Request for the *All of Us* Research Program Research Workbench

Thank you for your interest in the *All of Us* Researcher Workbench. Before you can gain access, your institution must agree to the *All of Us* Data Use and Registration Agreement. Currently, this agreement is open to any U.S.-based academic, nonprofit, or health care organization. Please complete this form so that our team can work with your institutional signing official to execute this agreement. The agreement will be with Vanderbilt University Medical Center, which operates the *All of Us* Data and Research Center.

For a list of institutions that have already signed the agreement, please visit https://www.researchallofus.org/institutional-agreements/.

We are working to put new agreements in place as soon as possible and will keep you posted on the status of your request. If you have questions at any time, please contact us at support@researchallofus.org.

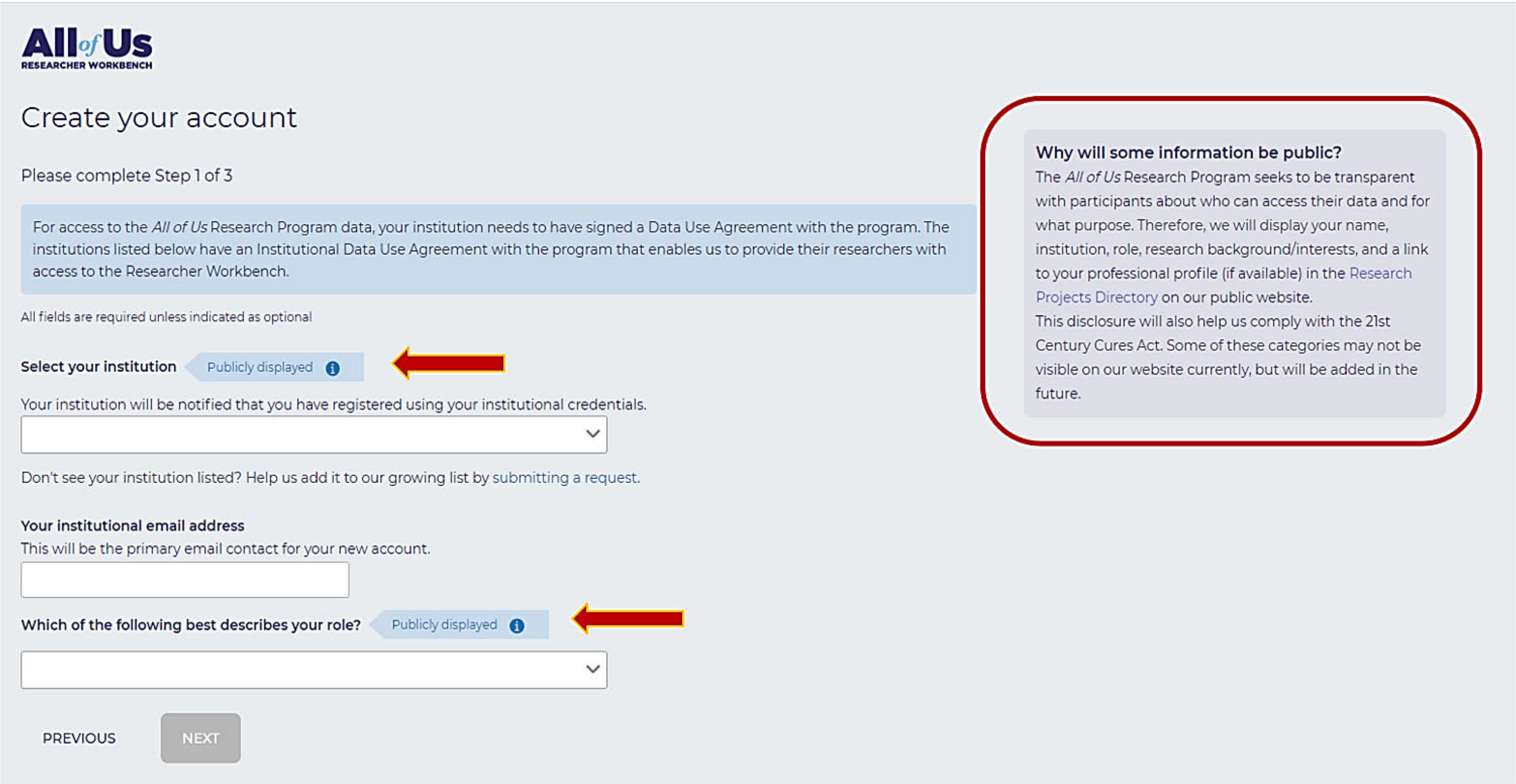
Thank you!

Are you affiliated with a US-based academic institution, healthcare institution, or non-profit?
* must provide value

Yes [✓ No]

At the current time, *All of Us* is in beta phase. US-based academic institutions, healthcare institutions, or non-profits are able to gain access to Researcher Workbench.

While we are working to expand access and understand who is interested in using the Researcher Workbench- will you share the name of the institution you are currently affiliated with?



Create your account

Please complete Step 1 of 3

For access to the *All of Us* Research Program data, your institution needs to have signed a Data Use Agreement with the program. The institutions listed below have an Institutional Data Use Agreement with the program that enables us to provide their researchers with access to the Researcher Workbench.

All fields are required unless indicated as optional

Select your institution

Your institution will be notified that you have registered using your institutional credentials.

Don't see your institution listed? Help us add it to our growing list by submitting a request.

Your institutional email address

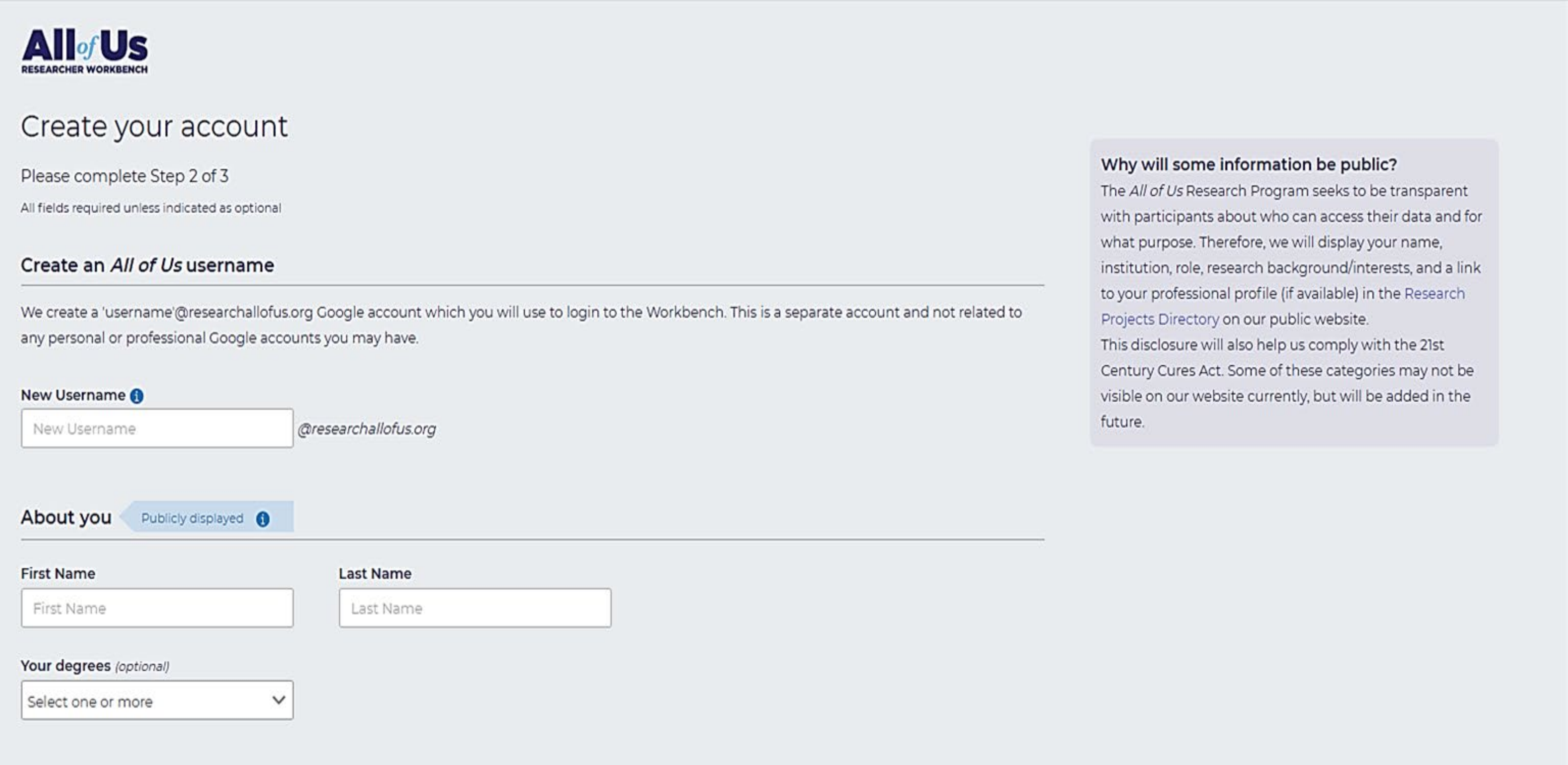
This will be the primary email contact for your new account.

Which of the following best describes your role?

Why will some information be public?

The *All of Us* Research Program seeks to be transparent with participants about who can access their data and for what purpose. Therefore, we will display your name, institution, role, research background/interests, and a link to your professional profile (if available) in the Research Projects Directory on our public website.

This disclosure will also help us comply with the 21st Century Cures Act. Some of these categories may not be visible on our website currently, but will be added in the future.



Create your account

Please complete Step 2 of 3

All fields required unless indicated as optional

Create an *All of Us* username

We create a 'username'@researchallofus.org Google account which you will use to login to the Workbench. This is a separate account and not related to any personal or professional Google accounts you may have.

New Username ⓘ

New Username @researchallofus.org

About you

First Name

First Name

Last Name

Last Name

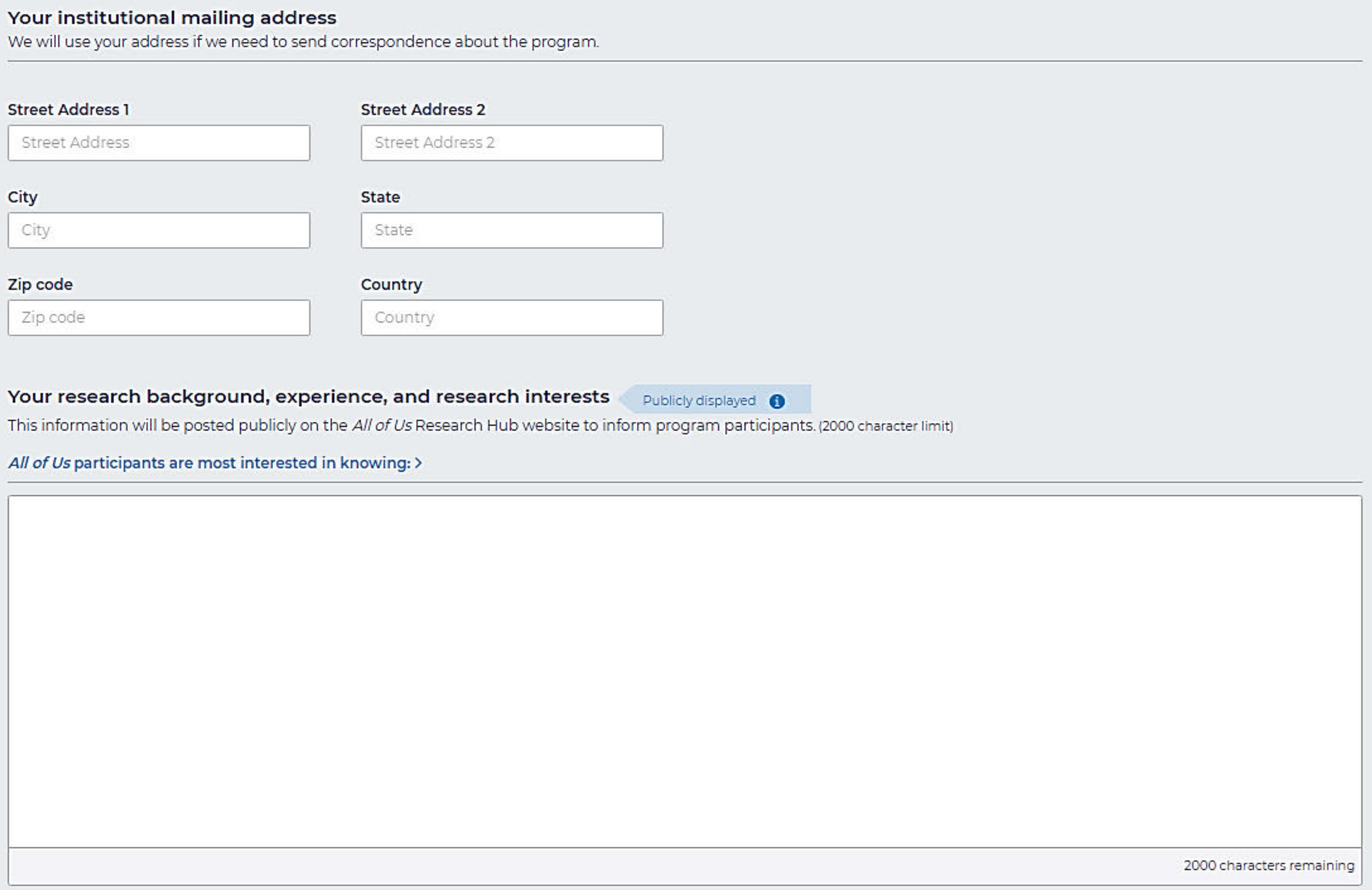
Your degrees *(optional)*

Select one or more ⌄

Why will some information be public?

The *All of Us* Research Program seeks to be transparent with participants about who can access their data and for what purpose. Therefore, we will display your name, institution, role, research background/interests, and a link to your professional profile (if available) in the Research Projects Directory on our public website.

This disclosure will also help us comply with the 21st Century Cures Act. Some of these categories may not be visible on our website currently, but will be added in the future.



Your institutional mailing address

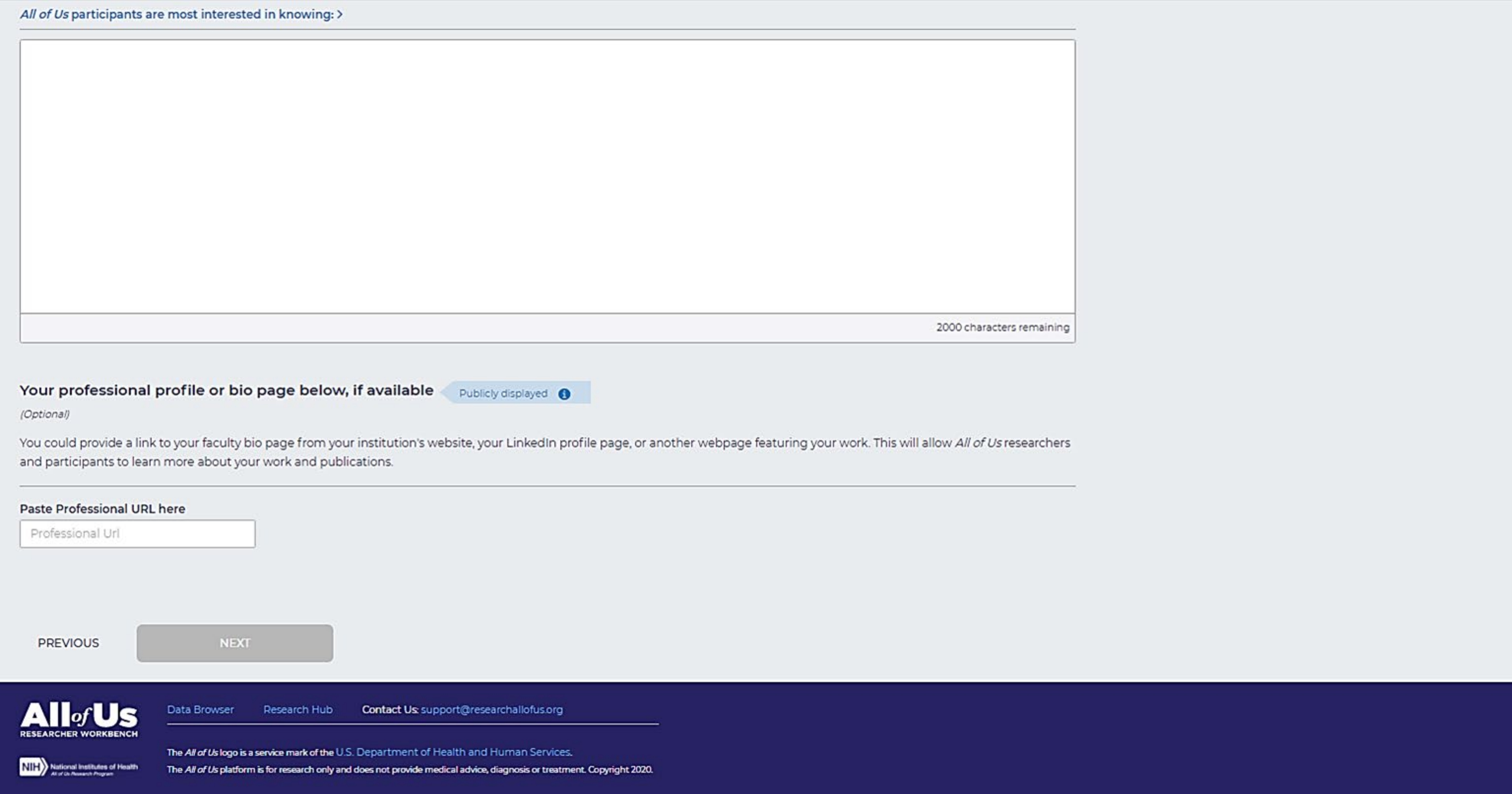
We will use your address if we need to send correspondence about the program.

Street Address 1	**Street Address 2**
Street Address	Street Address 2
City	**State**
City	State
Zip code	**Country**
Zip code	Country

Your research background, experience, and research interests

This information will be posted publicly on the *All of Us* Research Hub website to inform program participants. (2000 character limit)

2000 characters remaining



All of Us participants are most interested in knowing: >

2000 characters remaining

Your professional profile or bio page below, if available

(Optional)

You could provide a link to your faculty bio page from your institution's website, your LinkedIn profile page, or another webpage featuring your work. This will allow *All of Us* researchers and participants to learn more about your work and publications.

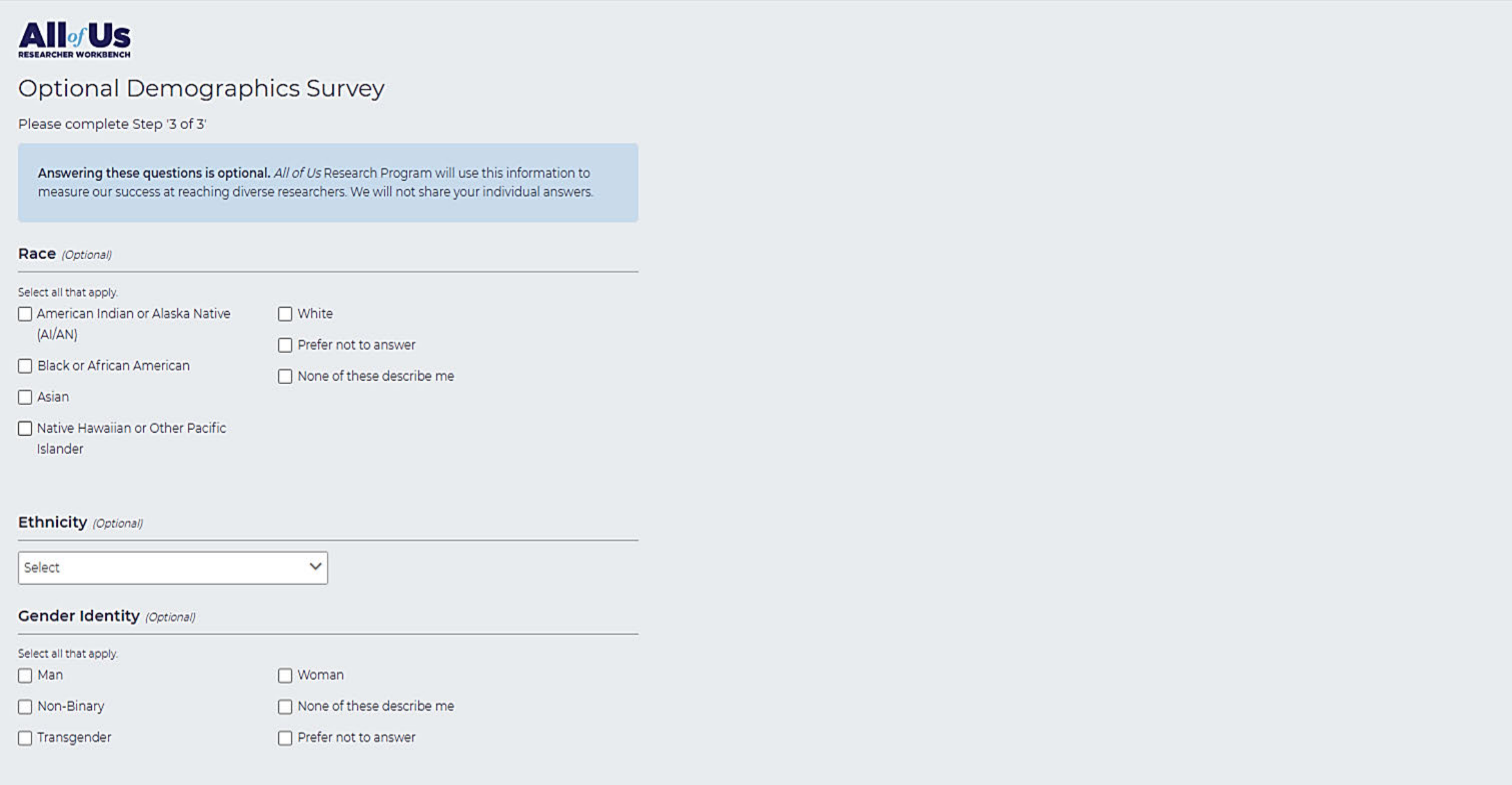
Paste Professional URL here

Professional Url

Contact Us: support@researchallofus.org

The *All of Us* logo is a service mark of the U.S. Department of Health and Human Services.

The *All of Us* platform is for research only and does not provide medical advice, diagnosis or treatment. Copyright 2020.



Optional Demographics Survey

Please complete Step '3 of 3'

Answering these questions is optional. *All of Us* Research Program will use this information to measure our success at reaching diverse researchers. We will not share your individual answers.

Race *(Optional)*

Select all that apply.

☐ American Indian or Alaska Native (AI/AN)
☐ Black or African American
☐ Asian
☐ Native Hawaiian or Other Pacific Islander
☐ White
☐ Prefer not to answer
☐ None of these describe me

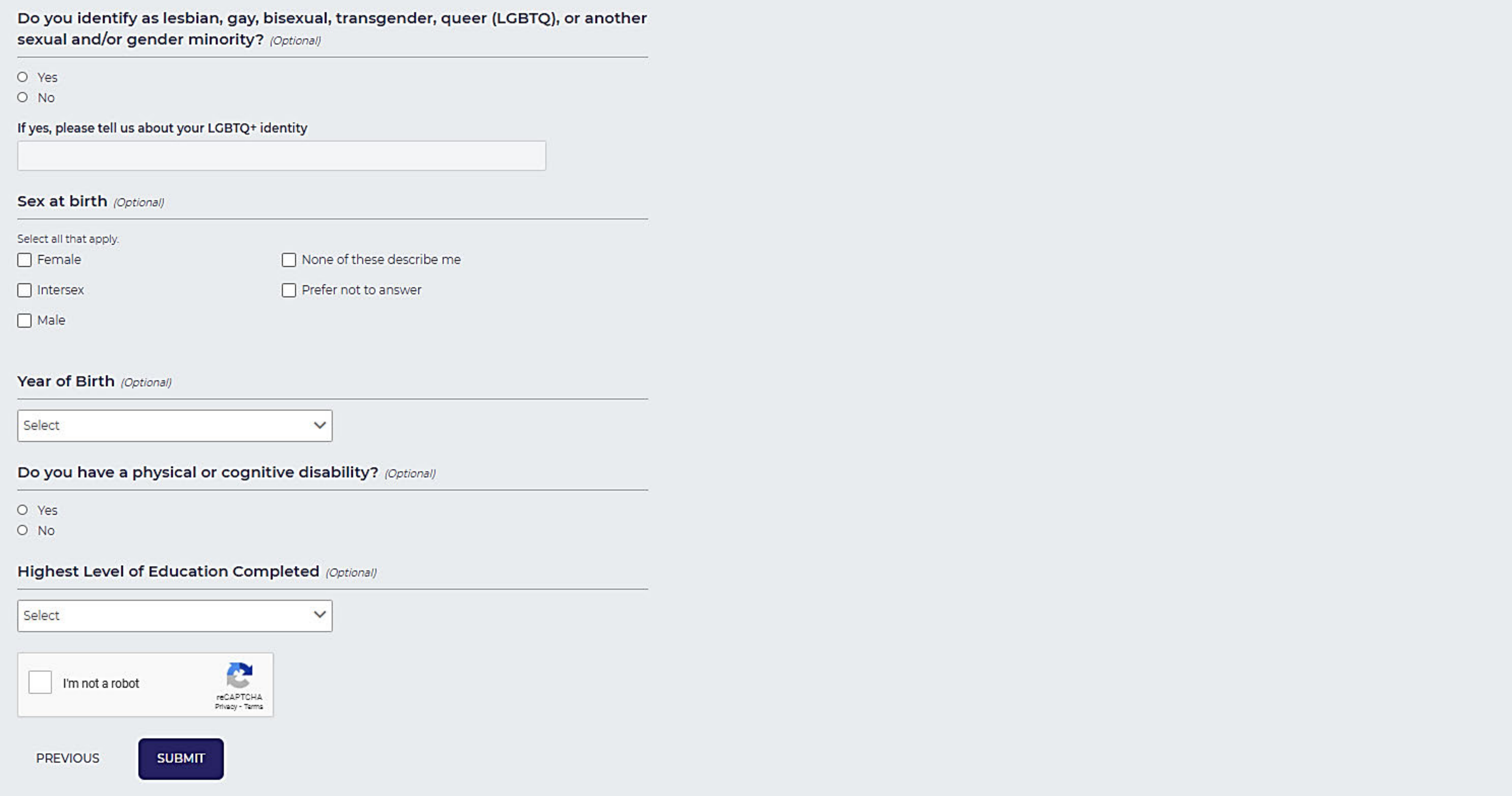
Ethnicity *(Optional)*

Select

Gender Identity *(Optional)*

Select all that apply.

☐ Man
☐ Non-Binary
☐ Transgender
☐ Woman
☐ None of these describe me
☐ Prefer not to answer



Do you identify as lesbian, gay, bisexual, transgender, queer (LGBTQ), or another sexual and/or gender minority? *(Optional)*

○ Yes
○ No

If yes, please tell us about your LGBTQ+ identity

Sex at birth *(Optional)*

Select all that apply.

☐ Female
☐ Intersex
☐ Male
☐ None of these describe me
☐ Prefer not to answer

Year of Birth *(Optional)*

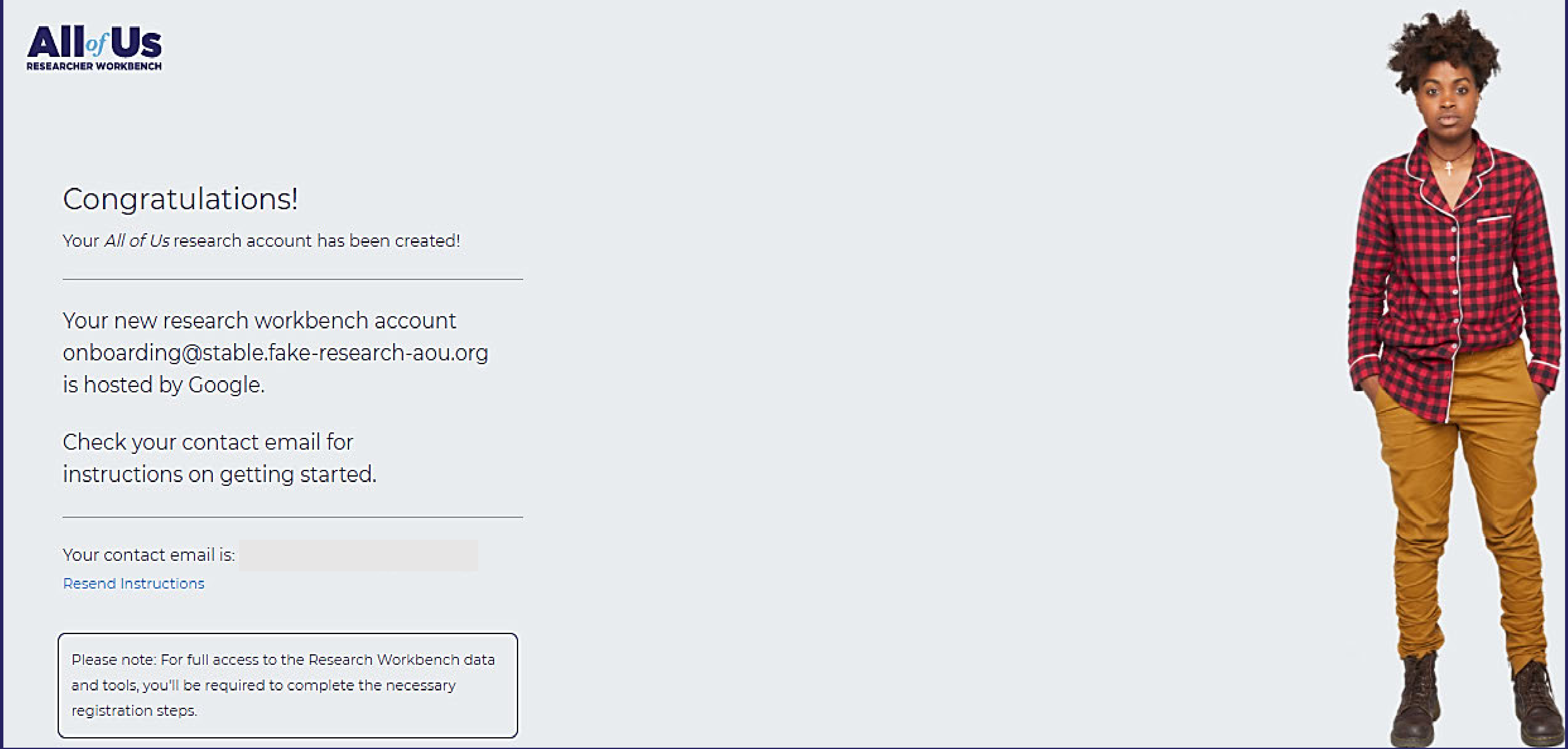
Select

Do you have a physical or cognitive disability? *(Optional)*

○ Yes
○ No

Highest Level of Education Completed *(Optional)*

Select



Congratulations!

Your *All of Us* research account has been created!

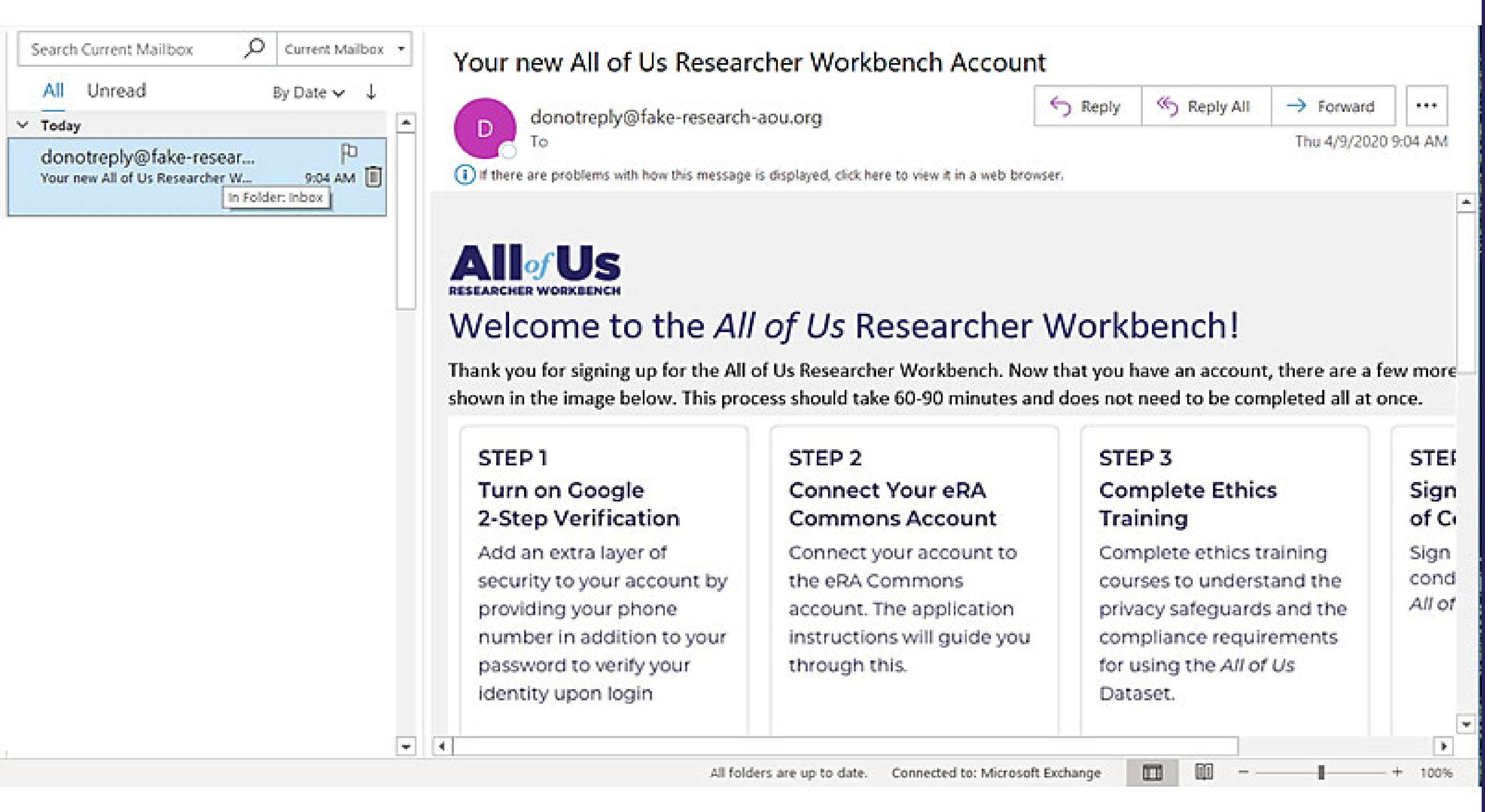
Your new research workbench account onboarding@stable.fake-research-aou.org is hosted by Google.

Check your contact email for instructions on getting started.

Your contact email is: [redacted]

Resend Instructions

Please note: For full access to the Research Workbench data and tools, you'll be required to complete the necessary registration steps.



Your new All of Us Researcher Workbench Account

donotreply@fake-research-aou.org
To
Thu 4/9/2020 9:04 AM

If there are problems with how this message is displayed, click here to view it in a web browser.

All of Us
RESEARCHER WORKBENCH

Welcome to the *All of Us* Researcher Workbench!

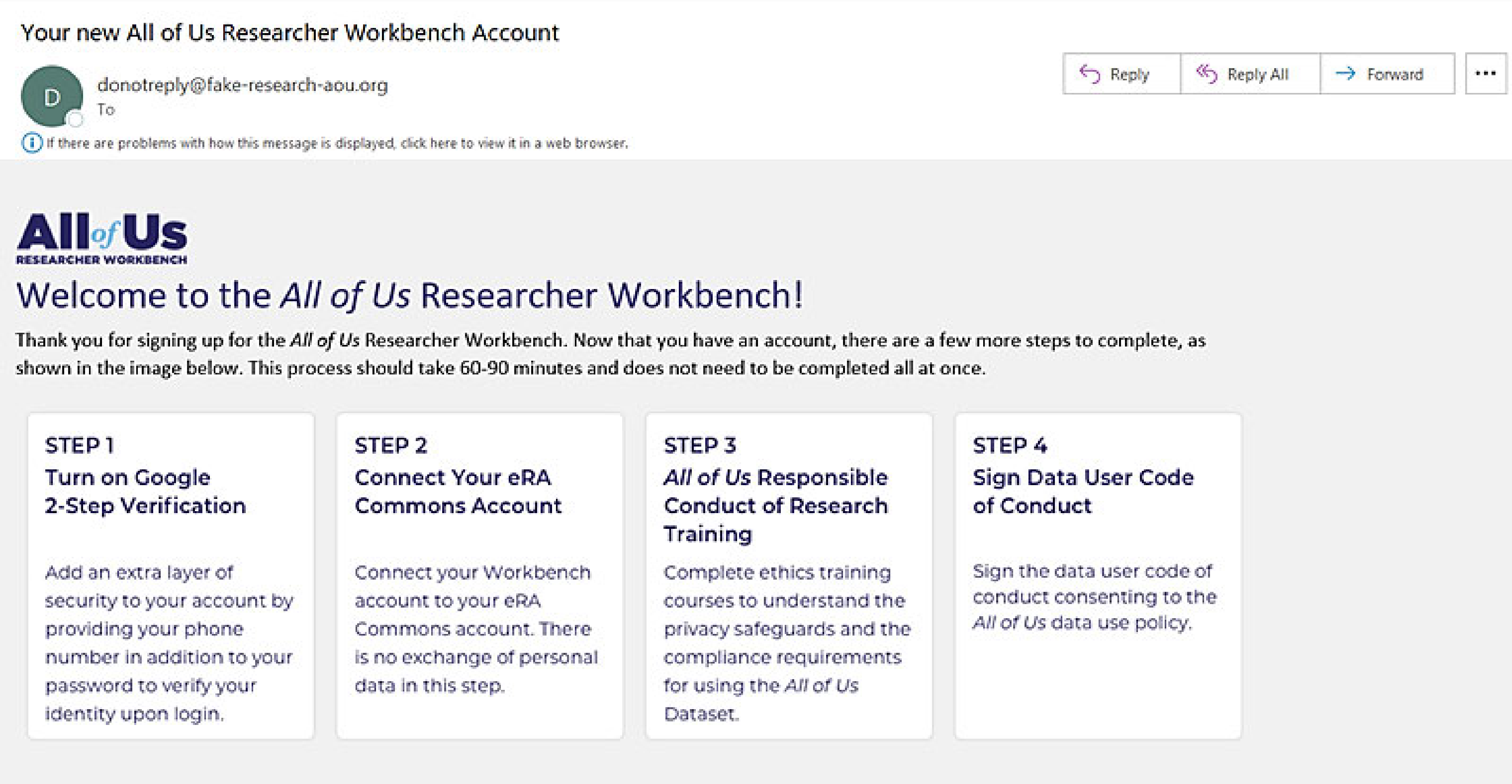
Thank you for signing up for the All of Us Researcher Workbench. Now that you have an account, there are a few more
shown in the image below. This process should take 60-90 minutes and does not need to be completed all at once.

STEP 1
Turn on Google 2-Step Verification
Add an extra layer of security to your account by providing your phone number in addition to your password to verify your identity upon login

STEP 2
Connect Your eRA Commons Account
Connect your account to the eRA Commons account. The application instructions will guide you through this.

STEP 3
Complete Ethics Training
Complete ethics training courses to understand the privacy safeguards and the compliance requirements for using the *All of Us* Dataset.

STEP
Sign
of C
Sign
cond
All of



Your new All of Us Researcher Workbench Account

donotreply@fake-research-aou.org
To

If there are problems with how this message is displayed, click here to view it in a web browser.

All of Us
RESEARCHER WORKBENCH

Welcome to the *All of Us* Researcher Workbench!

Thank you for signing up for the *All of Us* Researcher Workbench. Now that you have an account, there are a few more steps to complete, as shown in the image below. This process should take 60-90 minutes and does not need to be completed all at once.

STEP 1
Turn on Google 2-Step Verification

Add an extra layer of security to your account by providing your phone number in addition to your password to verify your identity upon login.

STEP 2
Connect Your eRA Commons Account

Connect your Workbench account to your eRA Commons account. There is no exchange of personal data in this step.

STEP 3
***All of Us* Responsible Conduct of Research Training**

Complete ethics training courses to understand the privacy safeguards and the compliance requirements for using the *All of Us* Dataset.

STEP 4
Sign Data User Code of Conduct

Sign the data user code of conduct consenting to the *All of Us* data use policy.

About your Researcher Workbench account

Your Researcher Workbench account is hosted by Google. This is different and isolated from your personal and professional Google accounts. There are two places during the registration process where you will interface directly with Google:

- Google Terms of Service agreement: Because the workbench account is hosted by Google, you will be asked to agree to Google's Terms of Service agreement.
- 2-step verification: Setting up 2 step verification on your Workbench account adds an extra level of security to your account.

Helpful tips before you get started:

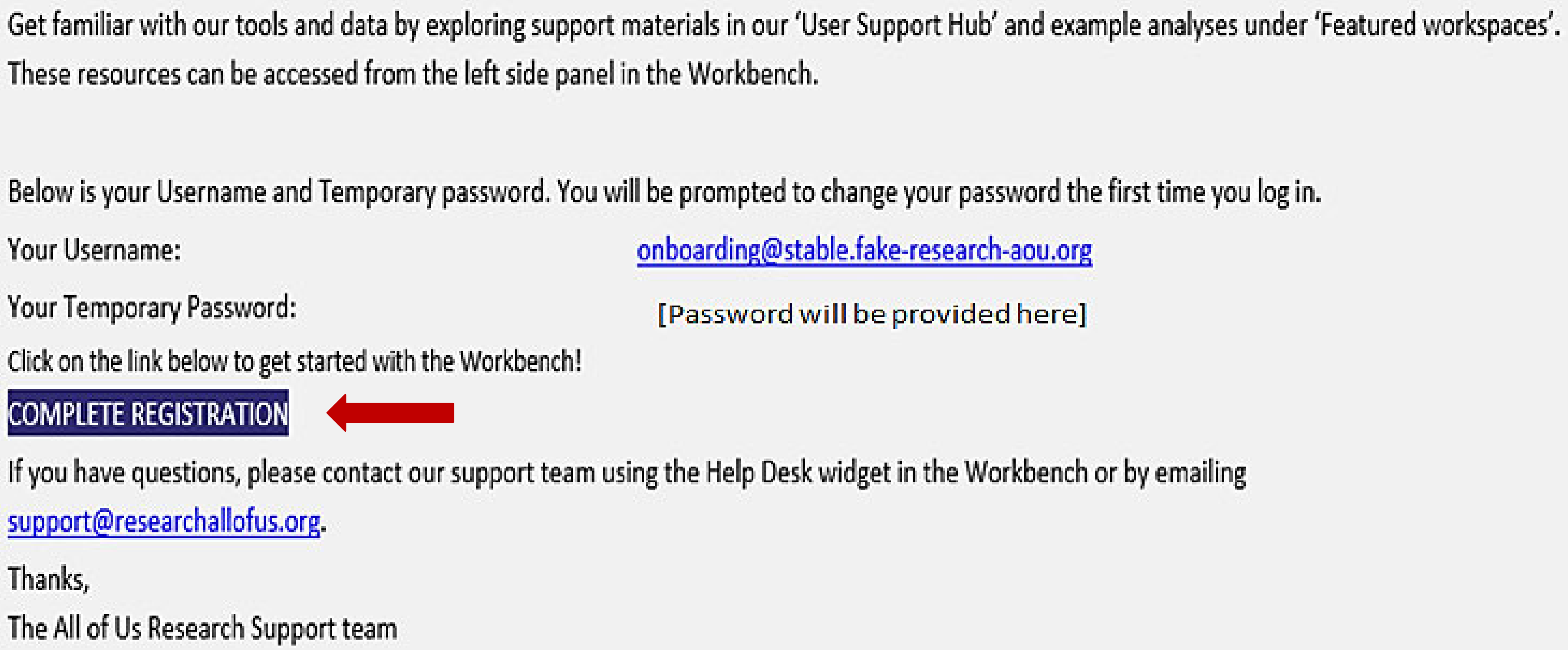
1

Chrome is the primary browser we currently support for the Researcher Workbench. We highly recommend using Chrome for the best user experience.

2

We recommend creating a Chrome Profile with your *All of Us* Researcher Workbench Google account. This will keep your Workbench browser sessions isolated if you use other Google accounts.

3



Get familiar with our tools and data by exploring support materials in our 'User Support Hub' and example analyses under 'Featured workspaces'. These resources can be accessed from the left side panel in the Workbench.

Below is your Username and Temporary password. You will be prompted to change your password the first time you log in.

Your Username: onboarding@stable.fake-research-aou.org

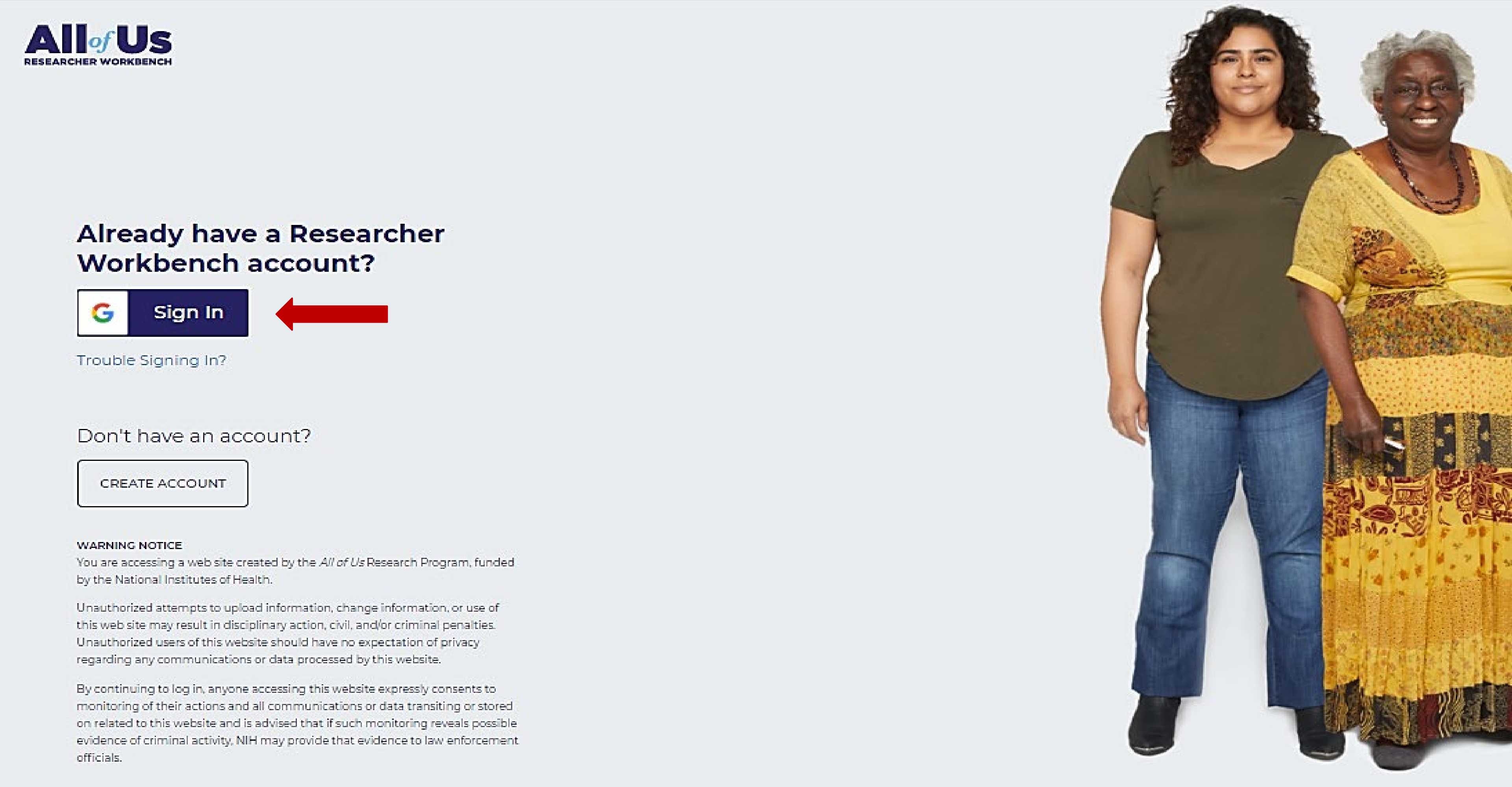
Your Temporary Password: [Password will be provided here]

Click on the link below to get started with the Workbench!

If you have questions, please contact our support team using the Help Desk widget in the Workbench or by emailing support@researchallofus.org.

Thanks,

The All of Us Research Support team



Already have a Researcher Workbench account?

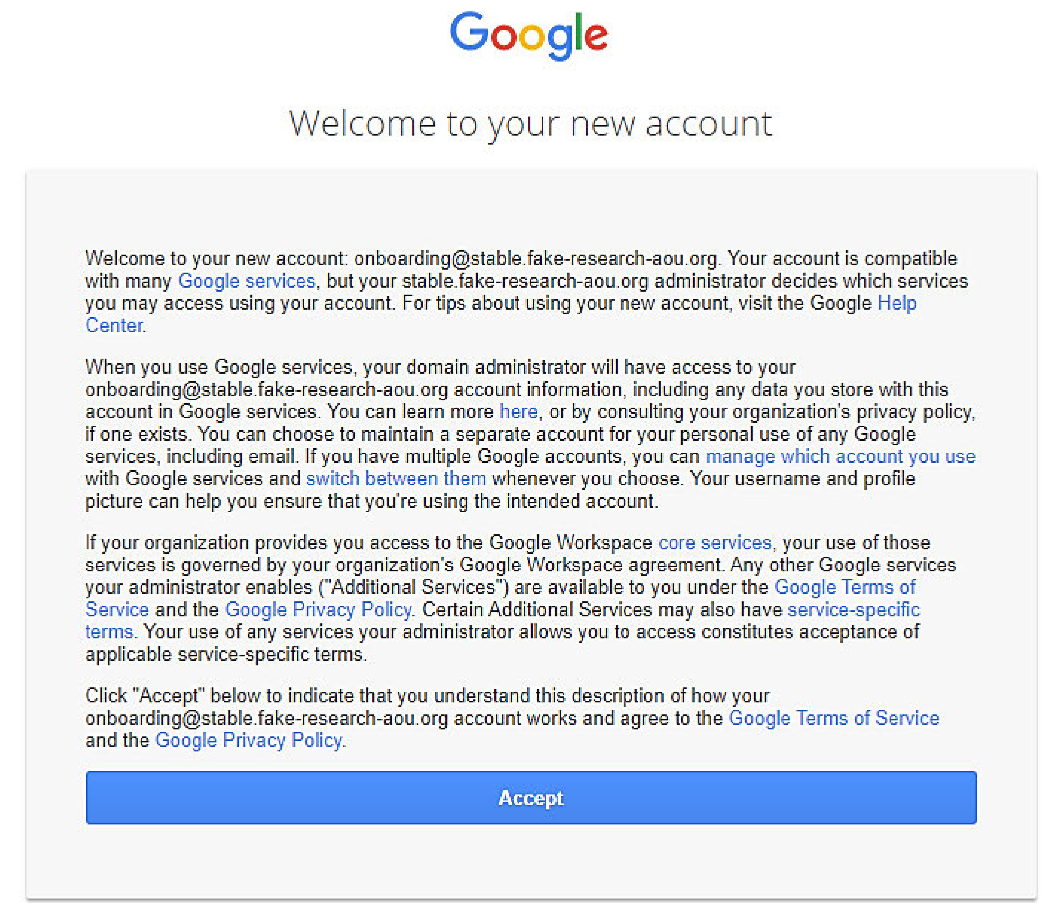
Don't have an account?

WARNING NOTICE

You are accessing a web site created by the *All of Us* Research Program, funded by the National Institutes of Health.

Unauthorized attempts to upload information, change information, or use of this web site may result in disciplinary action, civil, and/or criminal penalties. Unauthorized users of this website should have no expectation of privacy regarding any communications or data processed by this website.

By continuing to log in, anyone accessing this website expressly consents to monitoring of their actions and all communications or data transiting or stored on related to this website and is advised that if such monitoring reveals possible evidence of criminal activity, NIH may provide that evidence to law enforcement officials.



Google

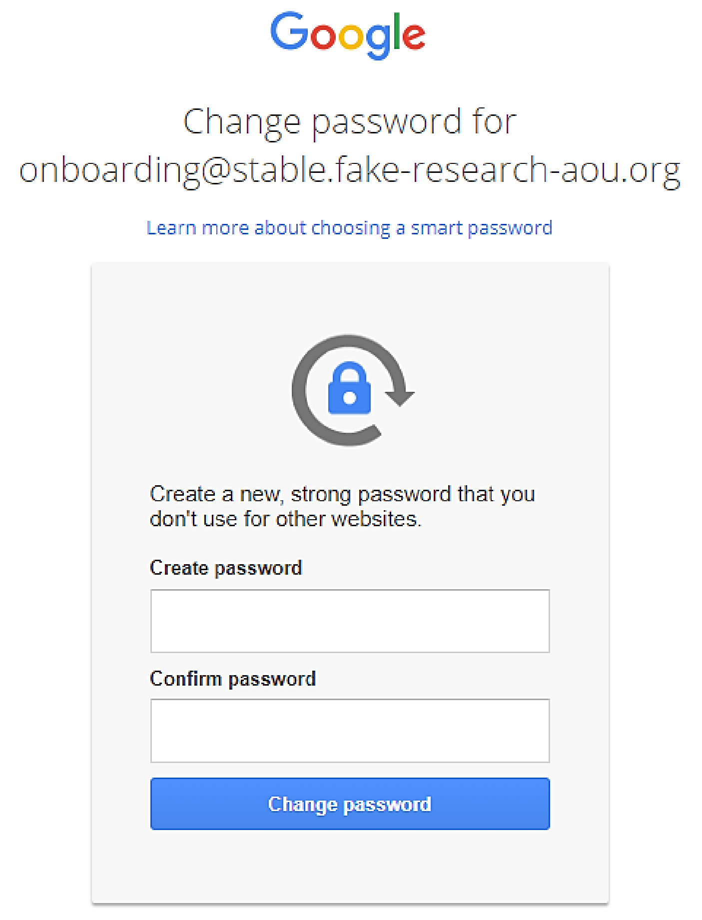
Welcome to your new account

Welcome to your new account: onboarding@stable.fake-research-aou.org. Your account is compatible with many Google services, but your stable.fake-research-aou.org administrator decides which services you may access using your account. For tips about using your new account, visit the Google Help Center.

When you use Google services, your domain administrator will have access to your onboarding@stable.fake-research-aou.org account information, including any data you store with this account in Google services. You can learn more here, or by consulting your organization's privacy policy, if one exists. You can choose to maintain a separate account for your personal use of any Google services, including email. If you have multiple Google accounts, you can manage which account you use with Google services and switch between them whenever you choose. Your username and profile picture can help you ensure that you're using the intended account.

If your organization provides you access to the Google Workspace core services, your use of those services is governed by your organization's Google Workspace agreement. Any other Google services your administrator enables ("Additional Services") are available to you under the Google Terms of Service and the Google Privacy Policy. Certain Additional Services may also have service-specific terms. Your use of any services your administrator allows you to access constitutes acceptance of applicable service-specific terms.

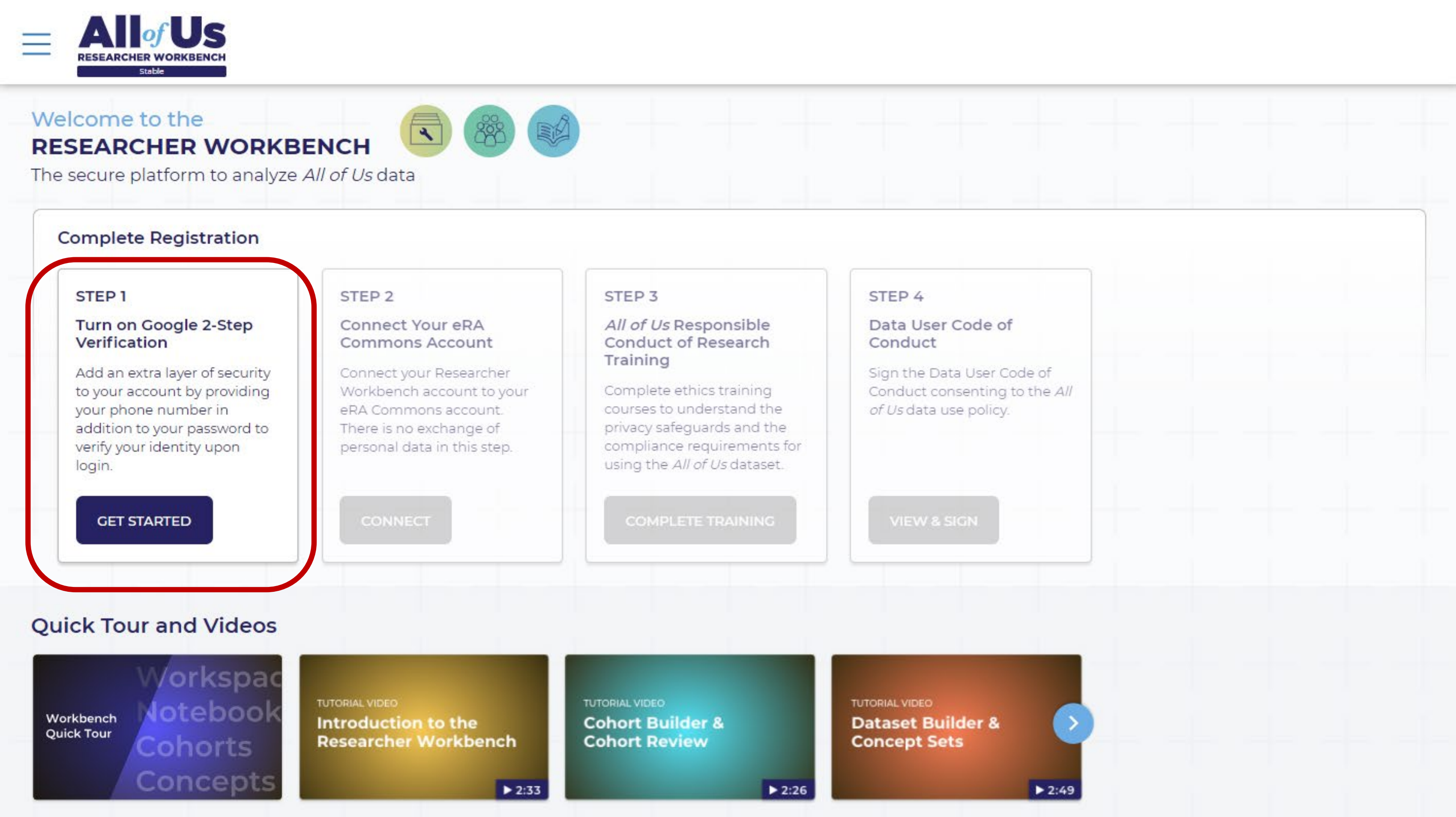
Click "Accept" below to indicate that you understand this description of how your onboarding@stable.fake-research-aou.org account works and agree to the Google Terms of Service and the Google Privacy Policy.



Change password for onboarding@stable.fake-research-aou.org

Learn more about choosing a smart password

Create a new, strong password that you don't use for other websites.



Welcome to the

RESEARCHER WORKBENCH

The secure platform to analyze *All of Us* data

Complete Registration

STEP 1

Turn on Google 2-Step Verification

Add an extra layer of security to your account by providing your phone number in addition to your password to verify your identity upon login.

GET STARTED

STEP 2

Connect Your eRA Commons Account

Connect your Researcher Workbench account to your eRA Commons account. There is no exchange of personal data in this step.

CONNECT

STEP 3

***All of Us* Responsible Conduct of Research Training**

Complete ethics training courses to understand the privacy safeguards and the compliance requirements for using the *All of Us* dataset.

COMPLETE TRAINING

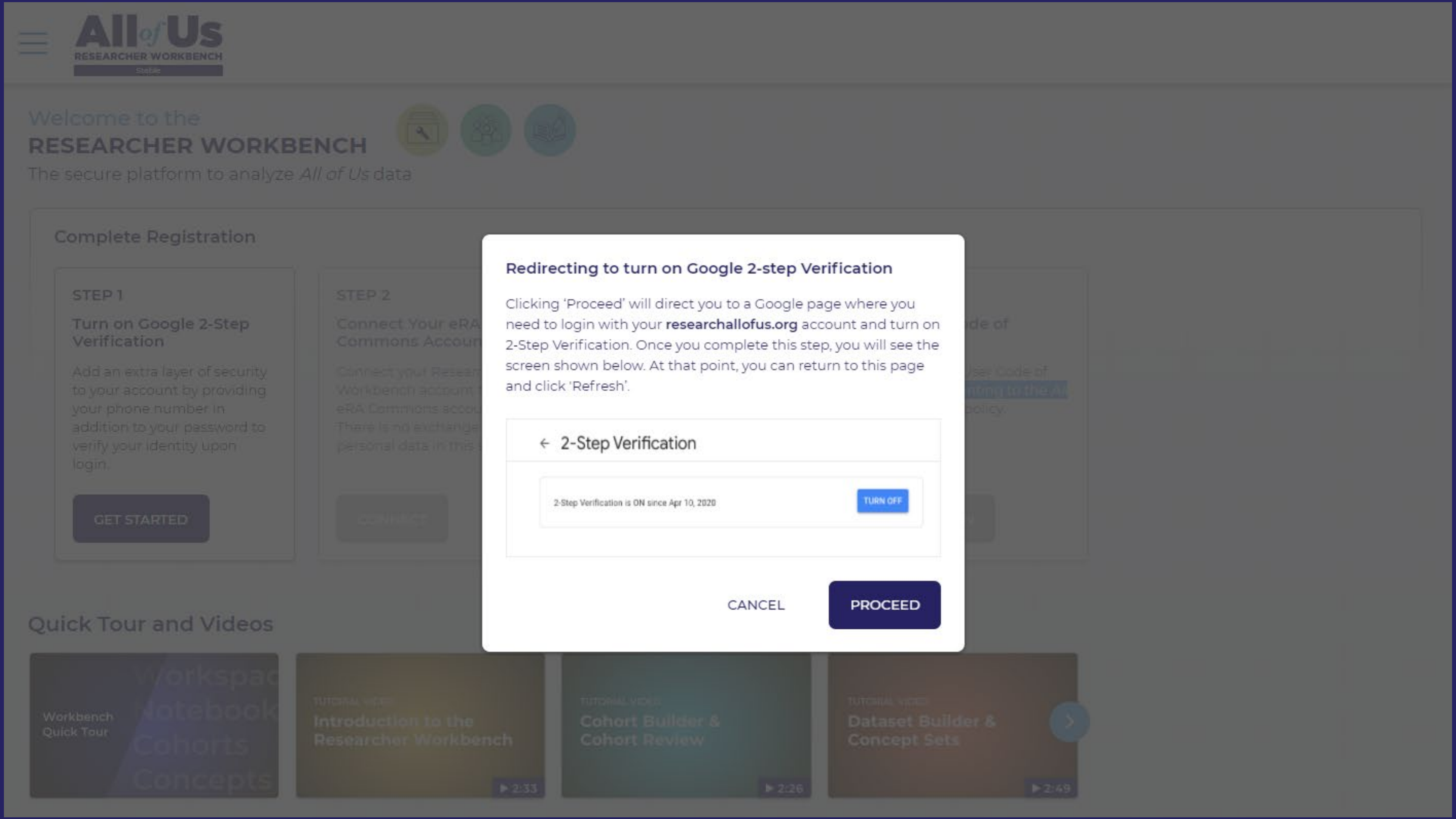
STEP 4

Data User Code of Conduct

Sign the Data User Code of Conduct consenting to the *All of Us* data use policy.

VIEW & SIGN

Quick Tour and Videos



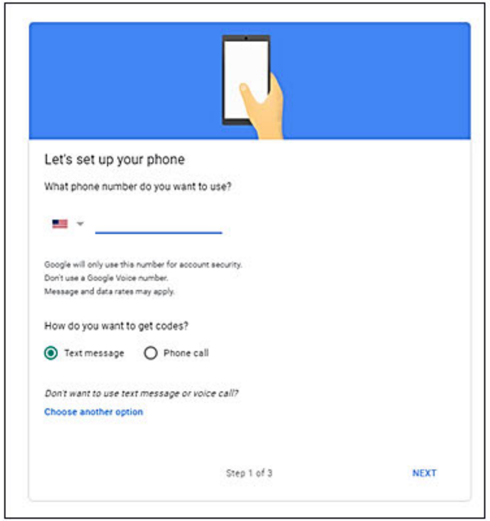
Redirecting to turn on Google 2-step Verification

Clicking 'Proceed' will direct you to a Google page where you need to login with your **researchallofus.org** account and turn on 2-Step Verification. Once you complete this step, you will see the screen shown below. At that point, you can return to this page and click 'Refresh'.

← 2-Step Verification

2-Step Verification is ON since Apr 10, 2020

1 → 2 → 3



Let's set up your phone

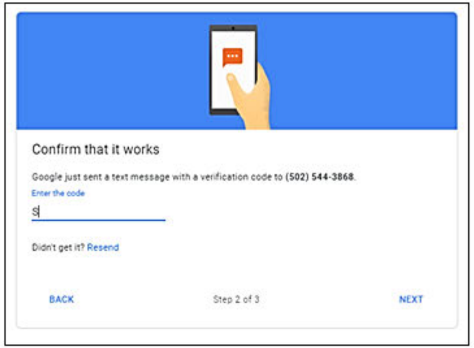
What phone number do you want to use?

Google will only use this number for account security.
Don't use a Google Voice number.
Message and data rates may apply.

How do you want to get codes?

● Text message ○ Phone call

Don't want to use text message or voice call?

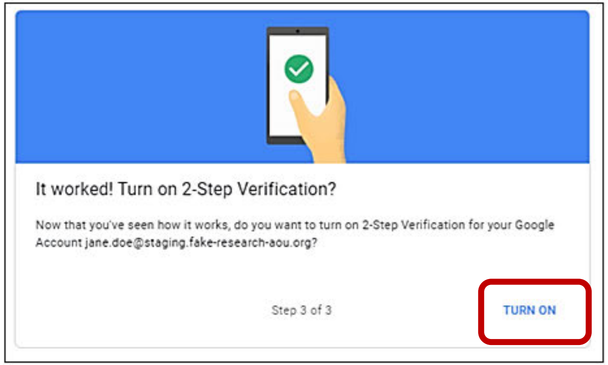


Confirm that it works

Google just sent a text message with a verification code to **(502) 544-3868**.

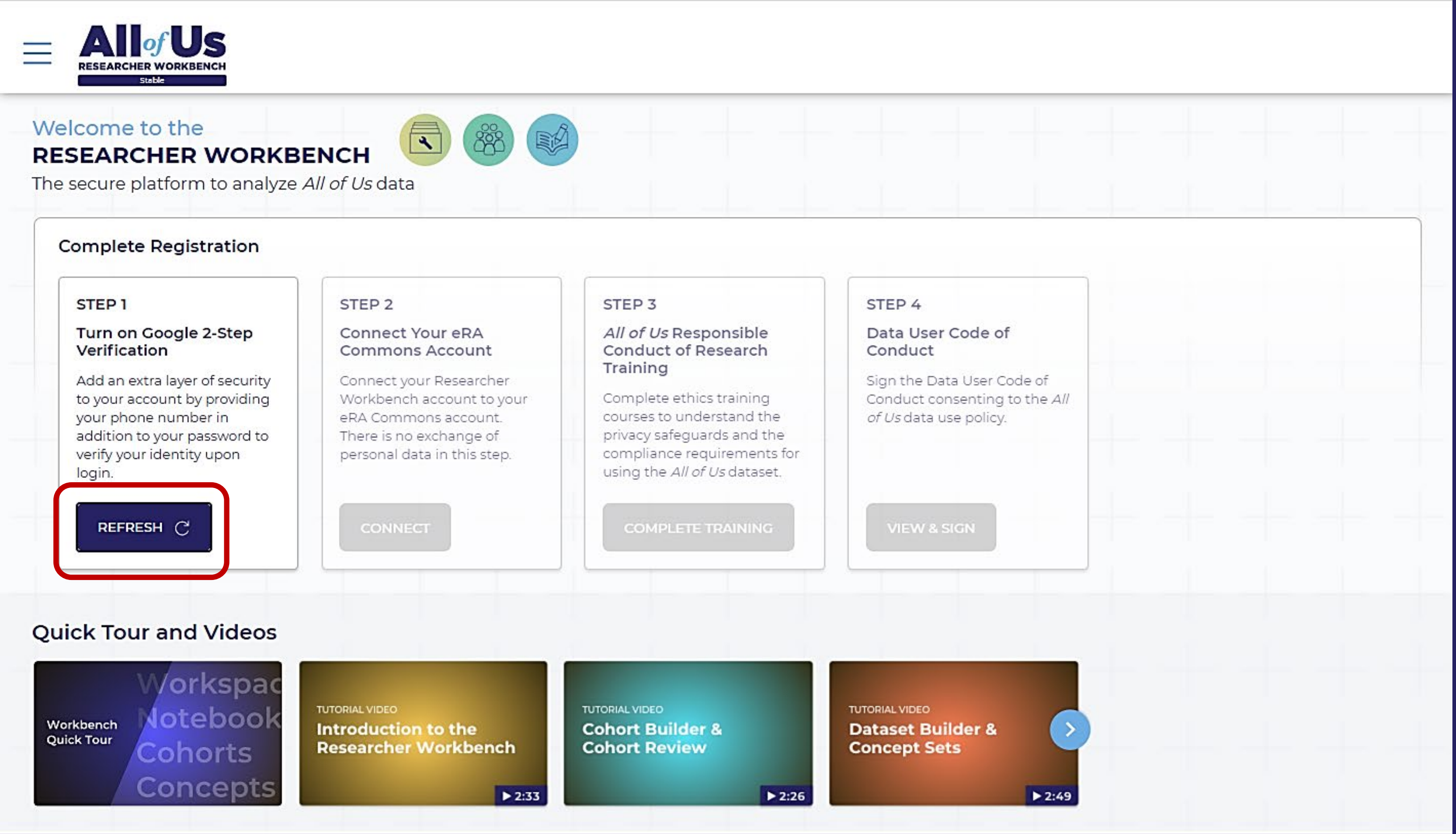
Enter the code: S

Didn't get it? Resend



It worked! Turn on 2-Step Verification?

Now that you've seen how it works, do you want to turn on 2-Step Verification for your Google Account jane.doe@staging.fake-research-aou.org?



Welcome to the

RESEARCHER WORKBENCH

The secure platform to analyze *All of Us* data

Complete Registration

STEP 1

Turn on Google 2-Step Verification

Add an extra layer of security to your account by providing your phone number in addition to your password to verify your identity upon login.

STEP 2

Connect Your eRA Commons Account

Connect your Researcher Workbench account to your eRA Commons account. There is no exchange of personal data in this step.

STEP 3

***All of Us* Responsible Conduct of Research Training**

Complete ethics training courses to understand the privacy safeguards and the compliance requirements for using the *All of Us* dataset.

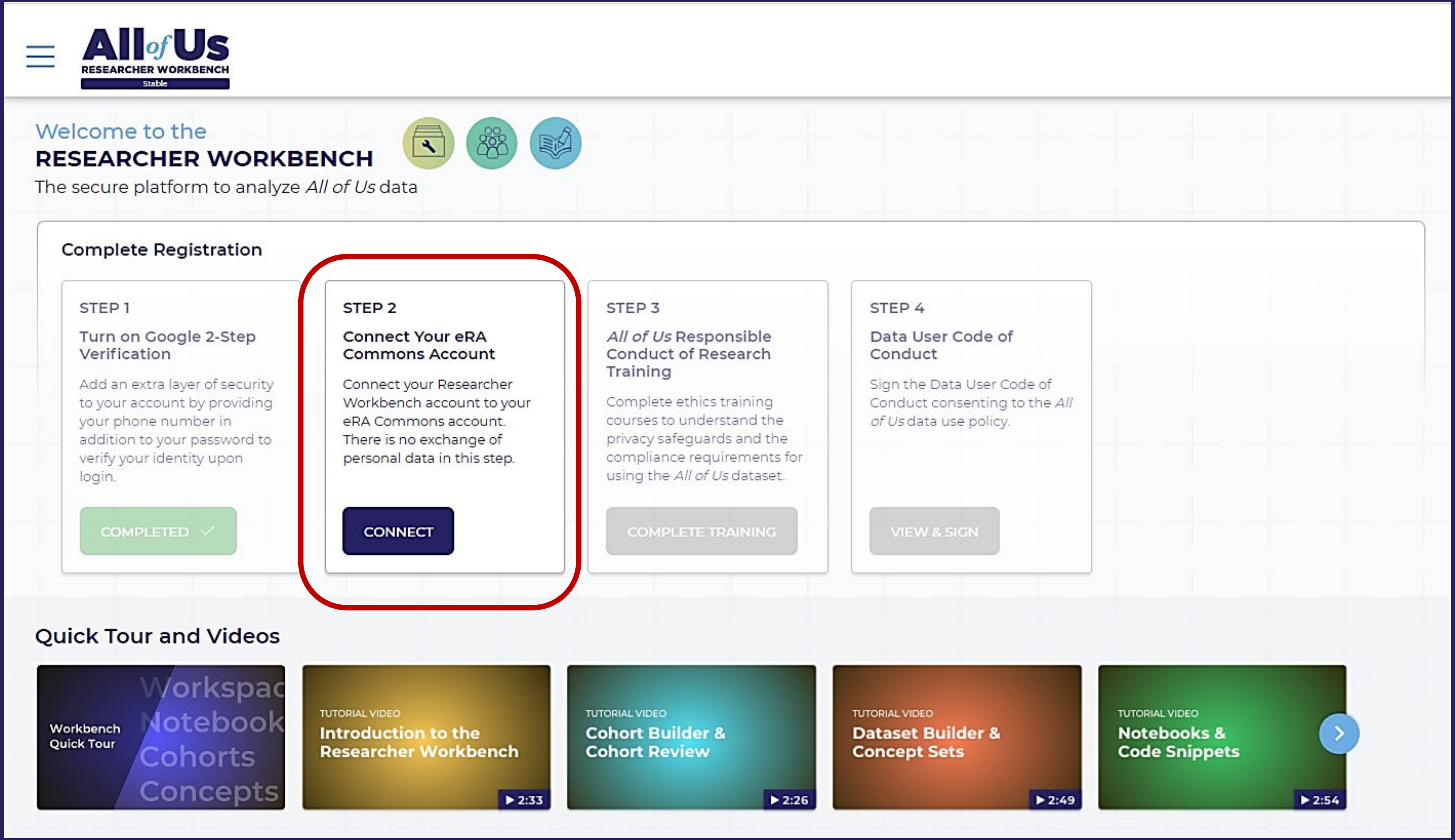
STEP 4

Data User Code of Conduct

Sign the Data User Code of Conduct consenting to the *All of Us* data use policy.

Quick Tour and Videos

All of Us Researcher Onramp



Welcome to the

RESEARCHER WORKBENCH

The secure platform to analyze *All of Us* data

Complete Registration

STEP 1

Turn on Google 2-Step Verification

Add an extra layer of security to your account by providing your phone number in addition to your password to verify your identity upon login.

STEP 2

Connect Your eRA Commons Account

Connect your Researcher Workbench account to your eRA Commons account. There is no exchange of personal data in this step.

STEP 3

All of Us Responsible Conduct of Research Training

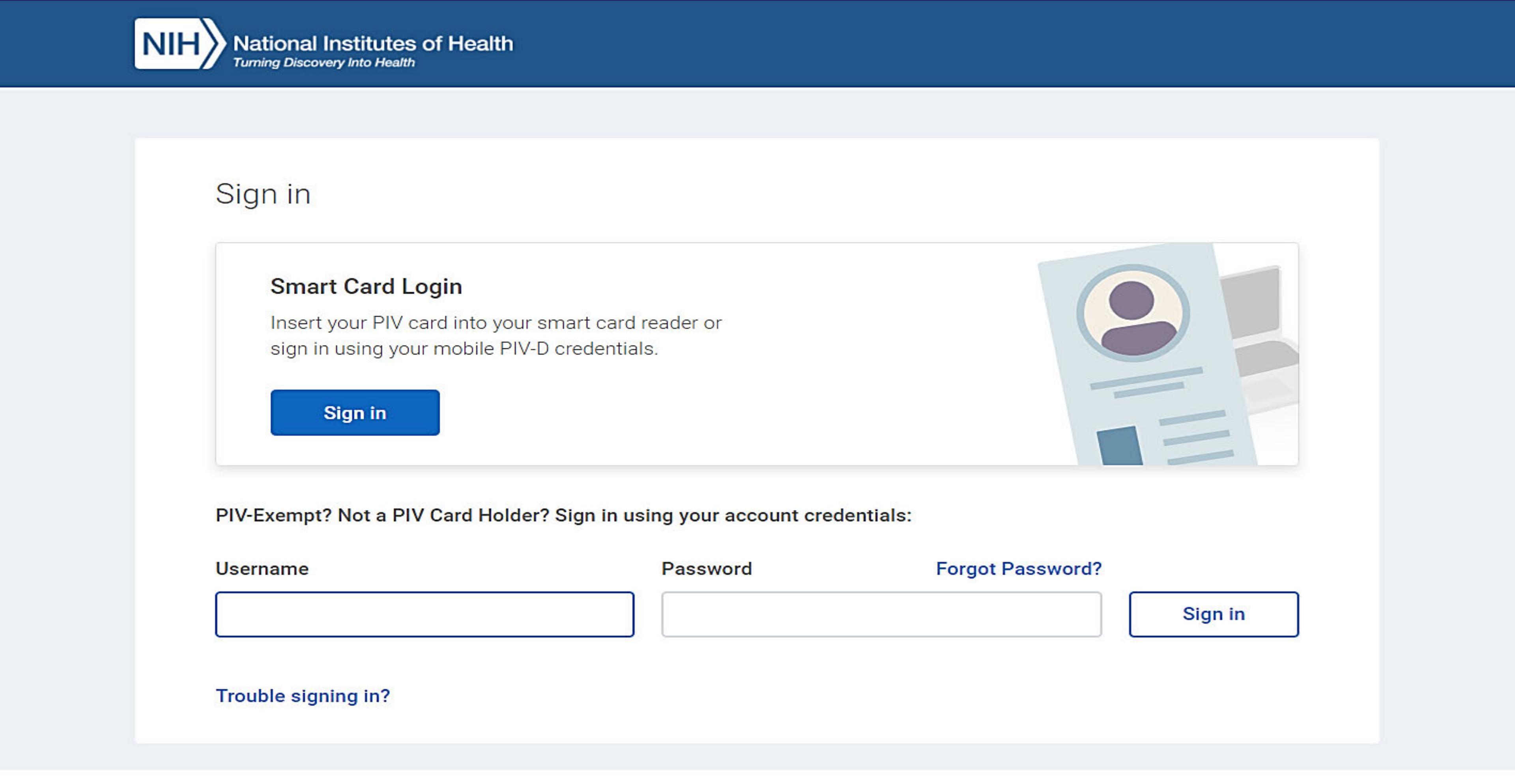
Complete ethics training courses to understand the privacy safeguards and the compliance requirements for using the *All of Us* dataset.

STEP 4

Data User Code of Conduct

Sign the Data User Code of Conduct consenting to the *All of Us* data use policy.

Quick Tour and Videos



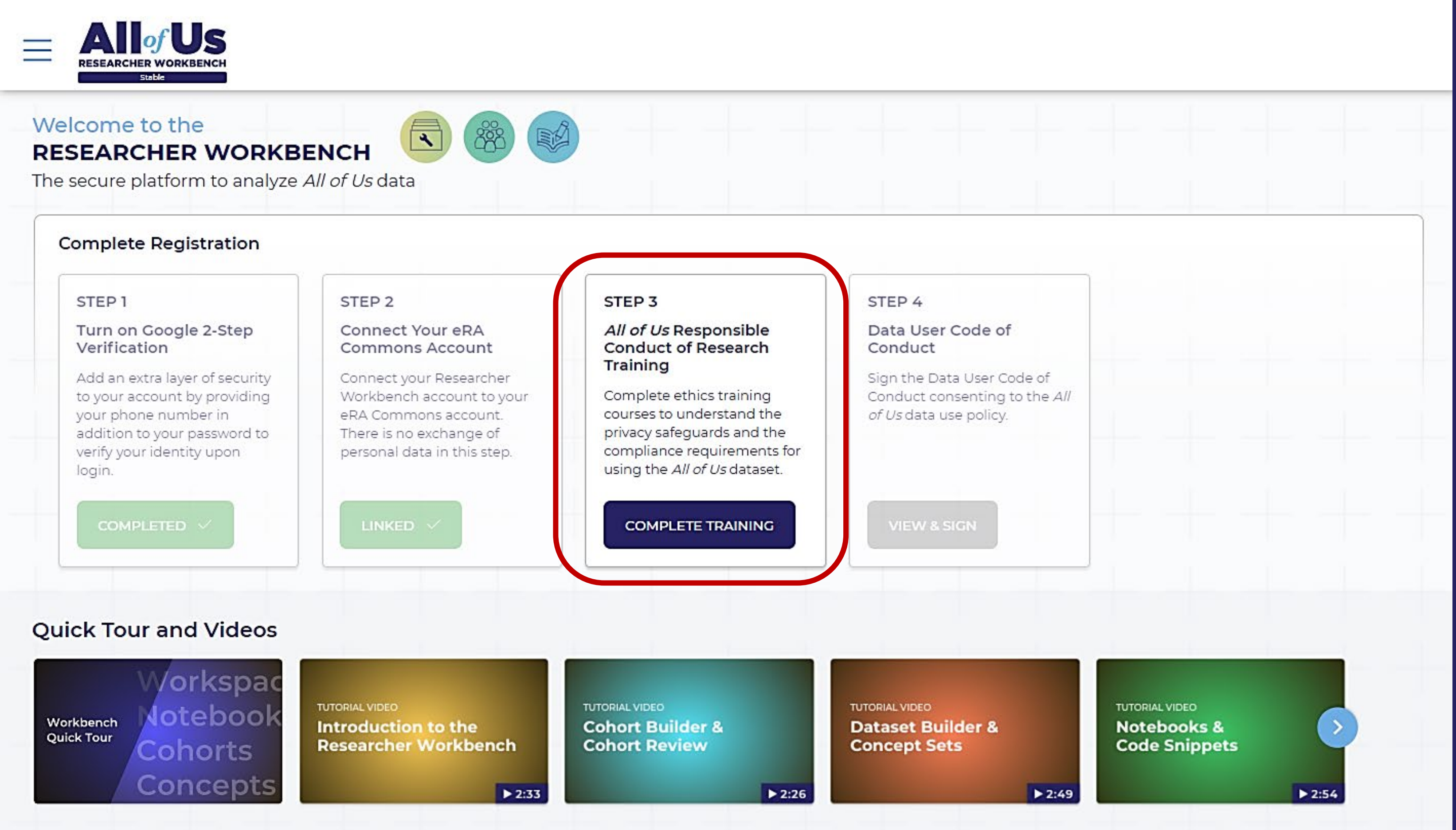
NIH National Institutes of Health
Turning Discovery Into Health

Sign in

Smart Card Login

Insert your PIV card into your smart card reader or sign in using your mobile PIV-D credentials.

PIV-Exempt? Not a PIV Card Holder? Sign in using your account credentials:



Welcome to the

RESEARCHER WORKBENCH

The secure platform to analyze *All of Us* data

Complete Registration

STEP 1

Turn on Google 2-Step Verification

Add an extra layer of security to your account by providing your phone number in addition to your password to verify your identity upon login.

STEP 2

Connect Your eRA Commons Account

Connect your Researcher Workbench account to your eRA Commons account. There is no exchange of personal data in this step.

STEP 3

***All of Us* Responsible Conduct of Research Training**

Complete ethics training courses to understand the privacy safeguards and the compliance requirements for using the *All of Us* dataset.

STEP 4

Data User Code of Conduct

Sign the Data User Code of Conduct consenting to the *All of Us* data use policy.

Quick Tour and Videos

Workbench Training

Dashboard / Workbench Training

Data Research Training

[Your name here] from *All of Us* Research Workbench

Courses needed for the *All of Us* Workbench Badge

Highlighted courses have been completed!

✖ [Test] All of Us Researcher Ethics Training 2020

All of Us Researcher Ethics Training 2020

Dashboard / My courses / AoU Researcher Ethics 2020

General

Complete Course

Topic 1

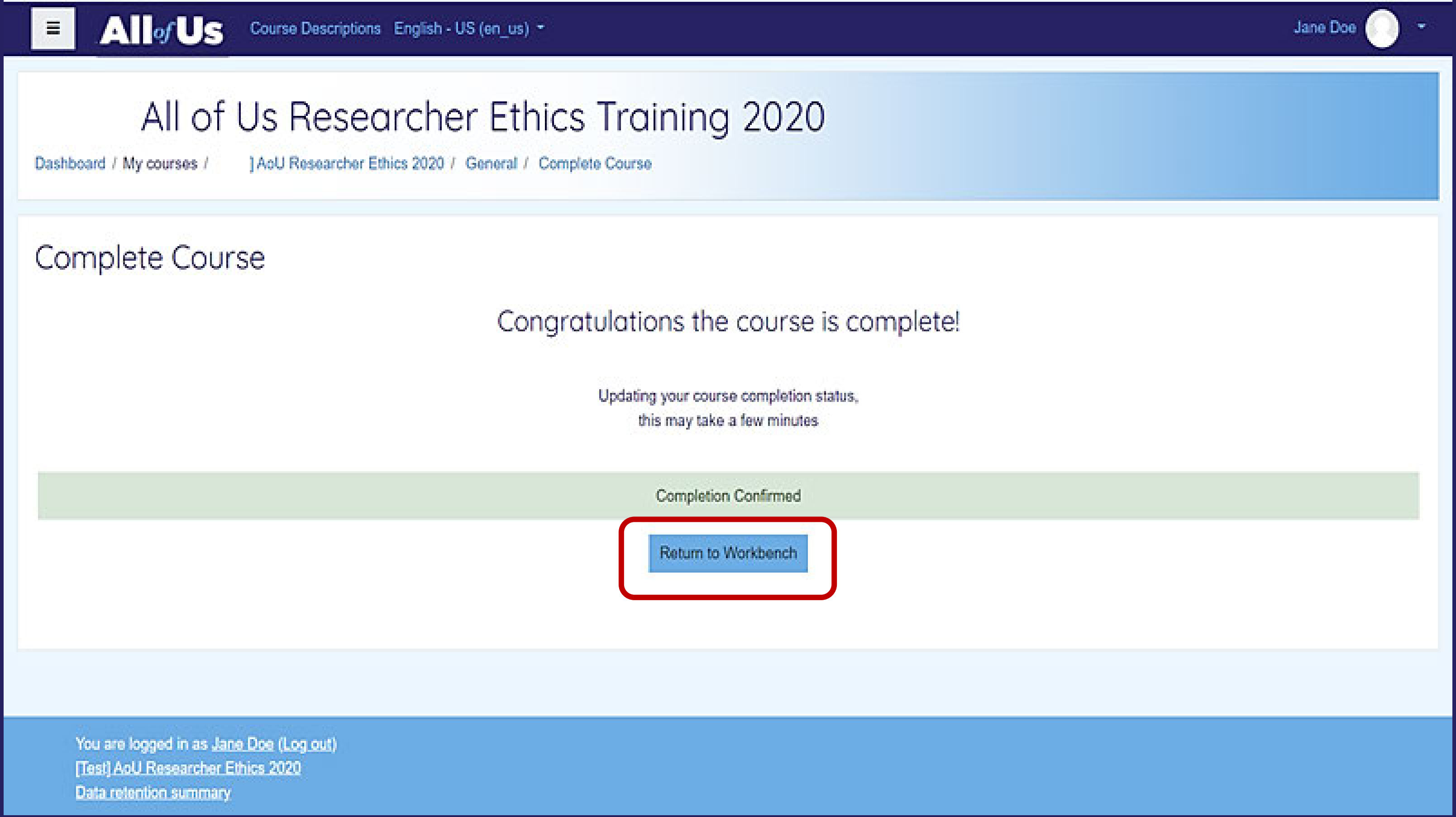
Course completion status

Status: *Not yet started*

Any criteria below are required:

Required criteria	**Status**
Manager	No
Activity completion	0 of 1

More details



All of Us Researcher Ethics Training 2020

Dashboard / My courses /] AoU Researcher Ethics 2020 / General / Complete Course

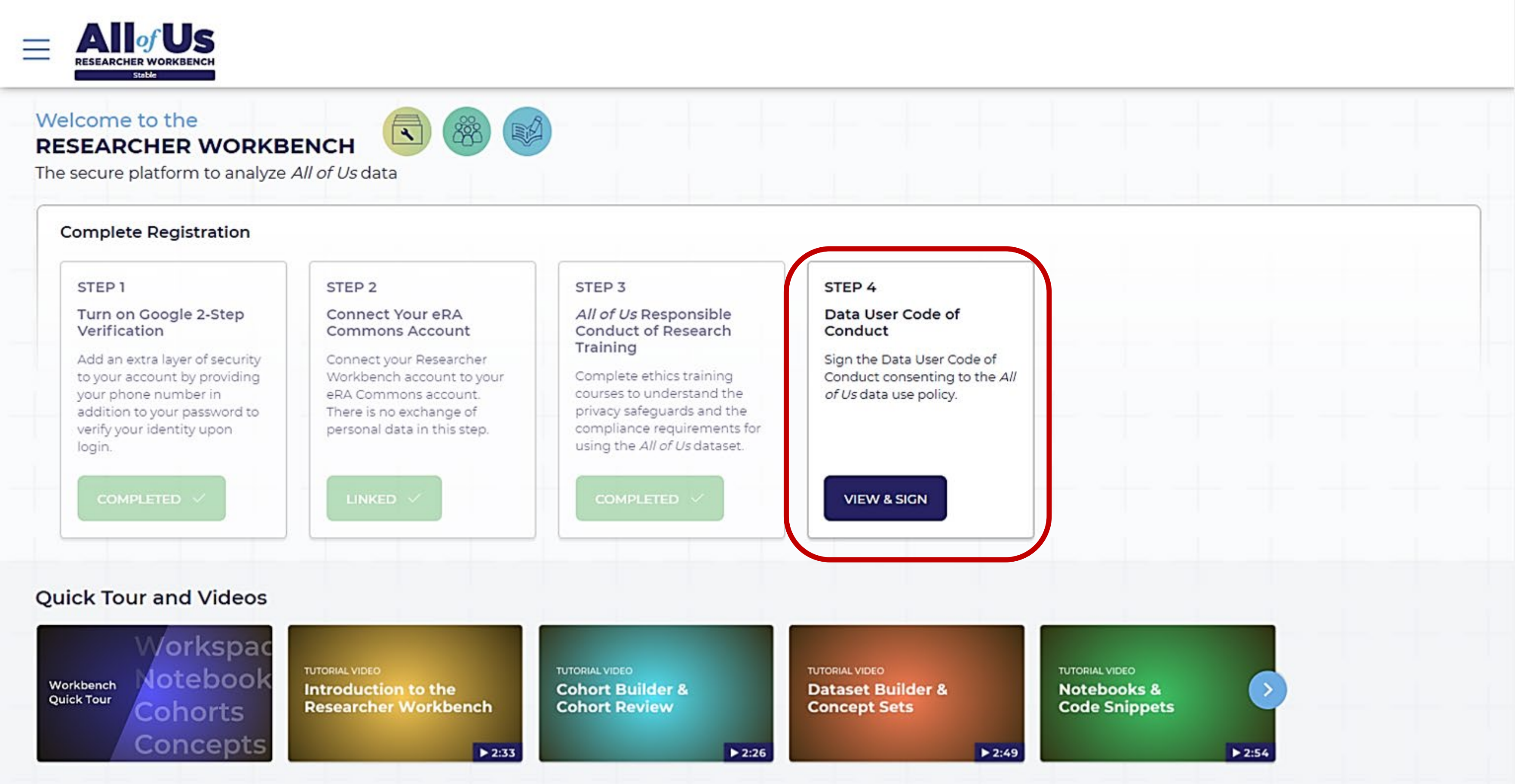
Complete Course

Congratulations the course is complete!

Updating your course completion status,
this may take a few minutes

Completion Confirmed

You are logged in as Jane Doe (Log out)
[Test] AoU Researcher Ethics 2020
Data retention summary



Welcome to the
RESEARCHER WORKBENCH
The secure platform to analyze *All of Us* data

Complete Registration

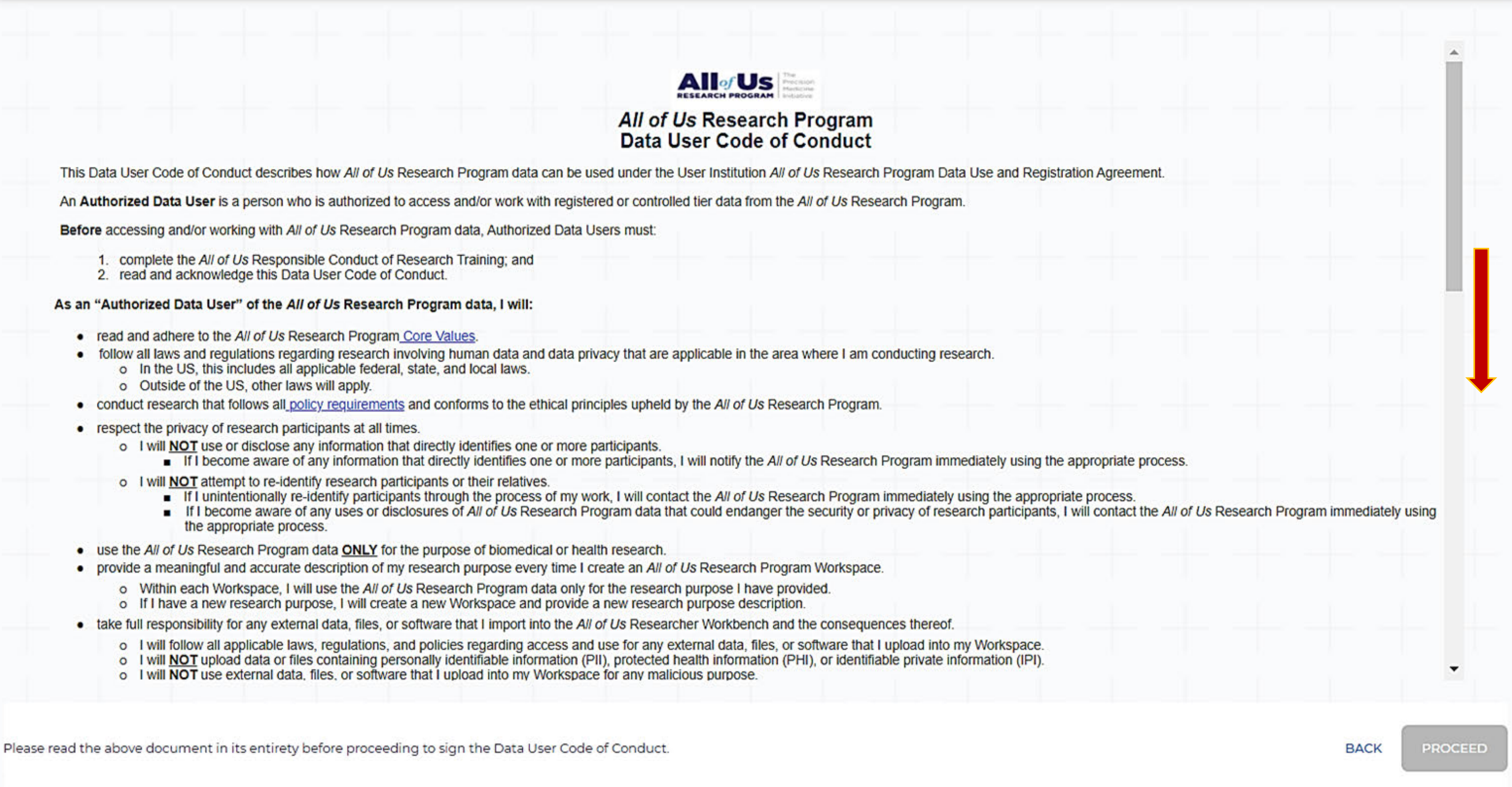
STEP 1
Turn on Google 2-Step Verification
Add an extra layer of security to your account by providing your phone number in addition to your password to verify your identity upon login.

STEP 2
Connect Your eRA Commons Account
Connect your Researcher Workbench account to your eRA Commons account. There is no exchange of personal data in this step.

STEP 3
***All of Us* Responsible Conduct of Research Training**
Complete ethics training courses to understand the privacy safeguards and the compliance requirements for using the *All of Us* dataset.

STEP 4
Data User Code of Conduct
Sign the Data User Code of Conduct consenting to the *All of Us* data use policy.

Quick Tour and Videos



All of Us Research Program
Data User Code of Conduct

This Data User Code of Conduct describes how *All of Us* Research Program data can be used under the User Institution *All of Us* Research Program Data Use and Registration Agreement.

An **Authorized Data User** is a person who is authorized to access and/or work with registered or controlled tier data from the *All of Us* Research Program.

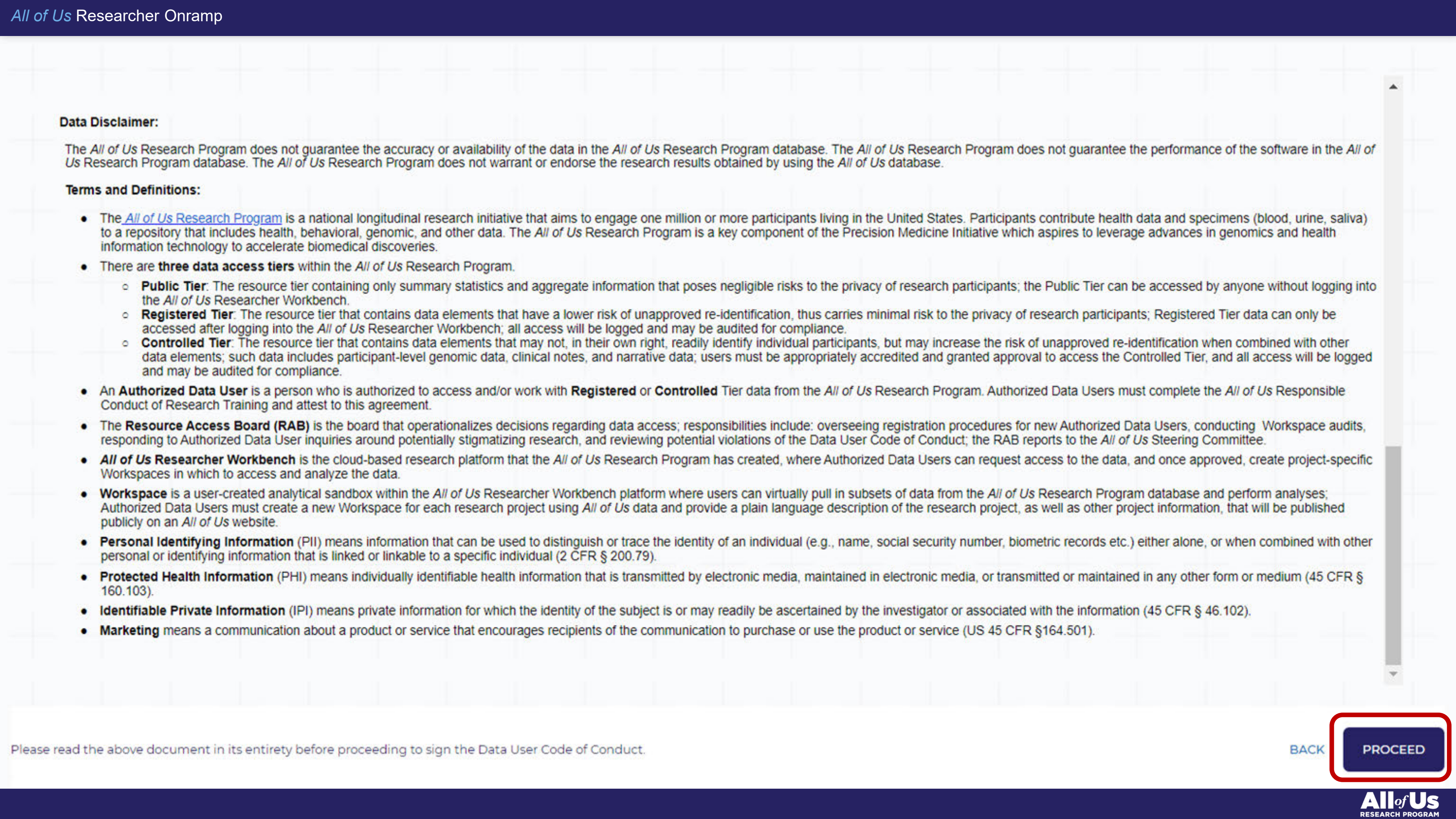
Before accessing and/or working with *All of Us* Research Program data, Authorized Data Users must:

1. complete the *All of Us* Responsible Conduct of Research Training; and
2. read and acknowledge this Data User Code of Conduct.

As an "Authorized Data User" of the *All of Us* Research Program data, I will:

- read and adhere to the *All of Us* Research Program Core Values.
- follow all laws and regulations regarding research involving human data and data privacy that are applicable in the area where I am conducting research.
    - In the US, this includes all applicable federal, state, and local laws.
    - Outside of the US, other laws will apply.
- conduct research that follows all policy requirements and conforms to the ethical principles upheld by the *All of Us* Research Program.
- respect the privacy of research participants at all times.
    - I will **NOT** use or disclose any information that directly identifies one or more participants.
        - If I become aware of any information that directly identifies one or more participants, I will notify the *All of Us* Research Program immediately using the appropriate process.
    - I will **NOT** attempt to re-identify research participants or their relatives.
        - If I unintentionally re-identify participants through the process of my work, I will contact the *All of Us* Research Program immediately using the appropriate process.
        - If I become aware of any uses or disclosures of *All of Us* Research Program data that could endanger the security or privacy of research participants, I will contact the *All of Us* Research Program immediately using the appropriate process.
- use the *All of Us* Research Program data **ONLY** for the purpose of biomedical or health research.
- provide a meaningful and accurate description of my research purpose every time I create an *All of Us* Research Program Workspace.
    - Within each Workspace, I will use the *All of Us* Research Program data only for the research purpose I have provided.
    - If I have a new research purpose, I will create a new Workspace and provide a new research purpose description.
- take full responsibility for any external data, files, or software that I import into the *All of Us* Researcher Workbench and the consequences thereof.
    - I will follow all applicable laws, regulations, and policies regarding access and use for any external data, files, or software that I upload into my Workspace.
    - I will **NOT** upload data or files containing personally identifiable information (PII), protected health information (PHI), or identifiable private information (IPI).
    - I will **NOT** use external data, files, or software that I upload into my Workspace for any malicious purpose.

Please read the above document in its entirety before proceeding to sign the Data User Code of Conduct.



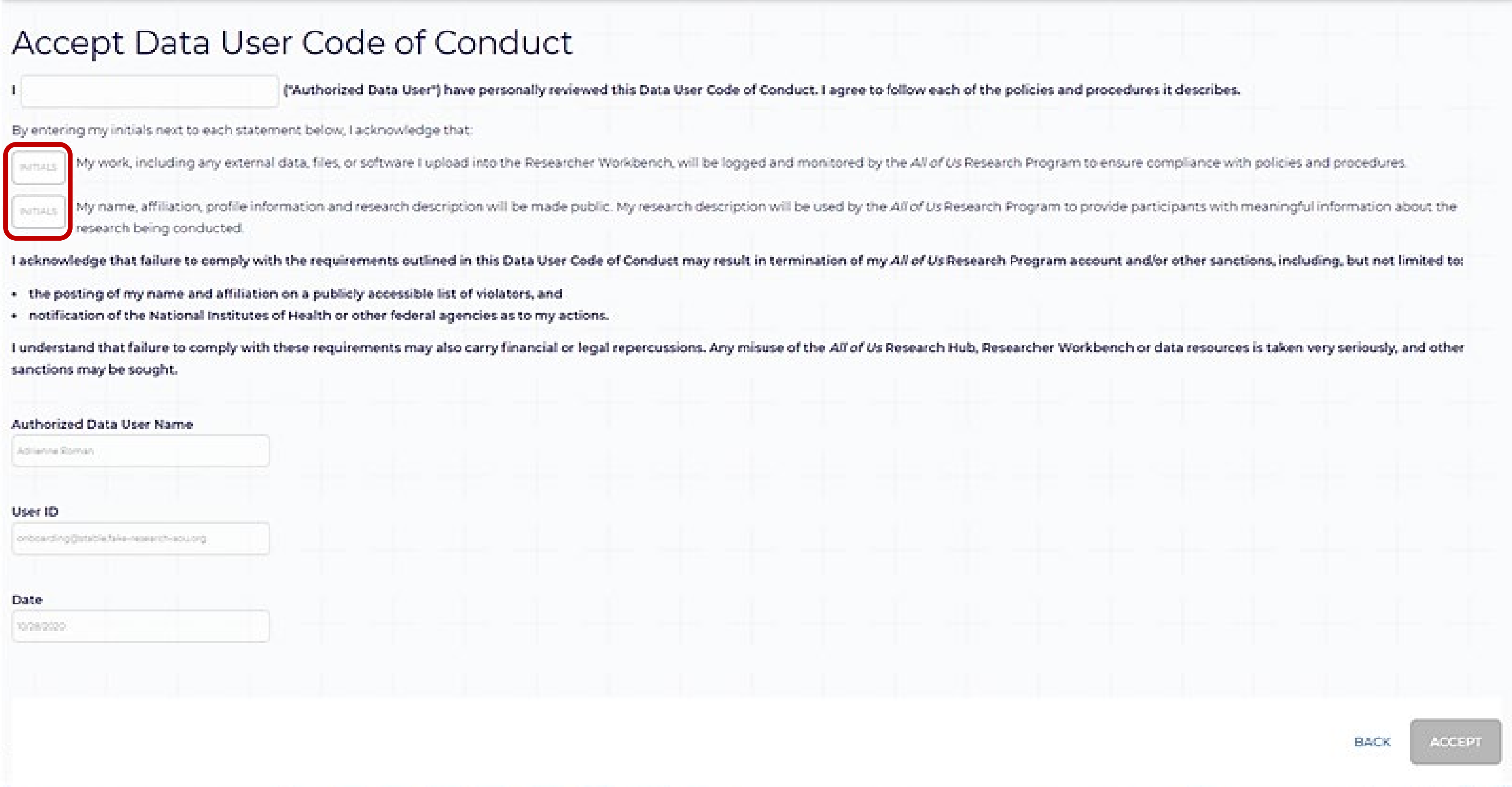
Data Disclaimer:

The *All of Us* Research Program does not guarantee the accuracy or availability of the data in the *All of Us* Research Program database. The *All of Us* Research Program does not guarantee the performance of the software in the *All of Us* Research Program database. The *All of Us* Research Program does not warrant or endorse the research results obtained by using the *All of Us* database.

Terms and Definitions:

- The *All of Us* Research Program is a national longitudinal research initiative that aims to engage one million or more participants living in the United States. Participants contribute health data and specimens (blood, urine, saliva) to a repository that includes health, behavioral, genomic, and other data. The *All of Us* Research Program is a key component of the Precision Medicine Initiative which aspires to leverage advances in genomics and health information technology to accelerate biomedical discoveries.
- There are **three data access tiers** within the *All of Us* Research Program.
    - **Public Tier**: The resource tier containing only summary statistics and aggregate information that poses negligible risks to the privacy of research participants; the Public Tier can be accessed by anyone without logging into the *All of Us* Researcher Workbench.
    - **Registered Tier**: The resource tier that contains data elements that have a lower risk of unapproved re-identification, thus carries minimal risk to the privacy of research participants; Registered Tier data can only be accessed after logging into the *All of Us* Researcher Workbench; all access will be logged and may be audited for compliance.
    - **Controlled Tier**: The resource tier that contains data elements that may not, in their own right, readily identify individual participants, but may increase the risk of unapproved re-identification when combined with other data elements; such data includes participant-level genomic data, clinical notes, and narrative data; users must be appropriately accredited and granted approval to access the Controlled Tier, and all access will be logged and may be audited for compliance.
- An **Authorized Data User** is a person who is authorized to access and/or work with **Registered** or **Controlled** Tier data from the *All of Us* Research Program. Authorized Data Users must complete the *All of Us* Responsible Conduct of Research Training and attest to this agreement.
- The **Resource Access Board (RAB)** is the board that operationalizes decisions regarding data access; responsibilities include: overseeing registration procedures for new Authorized Data Users, conducting Workspace audits, responding to Authorized Data User inquiries around potentially stigmatizing research, and reviewing potential violations of the Data User Code of Conduct; the RAB reports to the *All of Us* Steering Committee.
- ***All of Us* Researcher Workbench** is the cloud-based research platform that the *All of Us* Research Program has created, where Authorized Data Users can request access to the data, and once approved, create project-specific Workspaces in which to access and analyze the data.
- **Workspace** is a user-created analytical sandbox within the *All of Us* Researcher Workbench platform where users can virtually pull in subsets of data from the *All of Us* Research Program database and perform analyses; Authorized Data Users must create a new Workspace for each research project using *All of Us* data and provide a plain language description of the research project, as well as other project information, that will be published publicly on an *All of Us* website.
- **Personal Identifying Information** (PII) means information that can be used to distinguish or trace the identity of an individual (e.g., name, social security number, biometric records etc.) either alone, or when combined with other personal or identifying information that is linked or linkable to a specific individual (2 CFR § 200.79).
- **Protected Health Information** (PHI) means individually identifiable health information that is transmitted by electronic media, maintained in electronic media, or transmitted or maintained in any other form or medium (45 CFR § 160.103).
- **Identifiable Private Information** (IPI) means private information for which the identity of the subject is or may readily be ascertained by the investigator or associated with the information (45 CFR § 46.102).
- **Marketing** means a communication about a product or service that encourages recipients of the communication to purchase or use the product or service (US 45 CFR §164.501).

Please read the above document in its entirety before proceeding to sign the Data User Code of Conduct.



Accept Data User Code of Conduct

I ______ ("Authorized Data User") have personally reviewed this Data User Code of Conduct. I agree to follow each of the policies and procedures it describes.

By entering my initials next to each statement below, I acknowledge that:

My work, including any external data, files, or software I upload into the Researcher Workbench, will be logged and monitored by the *All of Us* Research Program to ensure compliance with policies and procedures.

My name, affiliation, profile information and research description will be made public. My research description will be used by the *All of Us* Research Program to provide participants with meaningful information about the research being conducted.

I acknowledge that failure to comply with the requirements outlined in this Data User Code of Conduct may result in termination of my *All of Us* Research Program account and/or other sanctions, including, but not limited to:

- **the posting of my name and affiliation on a publicly accessible list of violators, and**
- **notification of the National Institutes of Health or other federal agencies as to my actions.**

I understand that failure to comply with these requirements may also carry financial or legal repercussions. Any misuse of the *All of Us* Research Hub, Researcher Workbench or data resources is taken very seriously, and other sanctions may be sought.

Authorized Data User Name

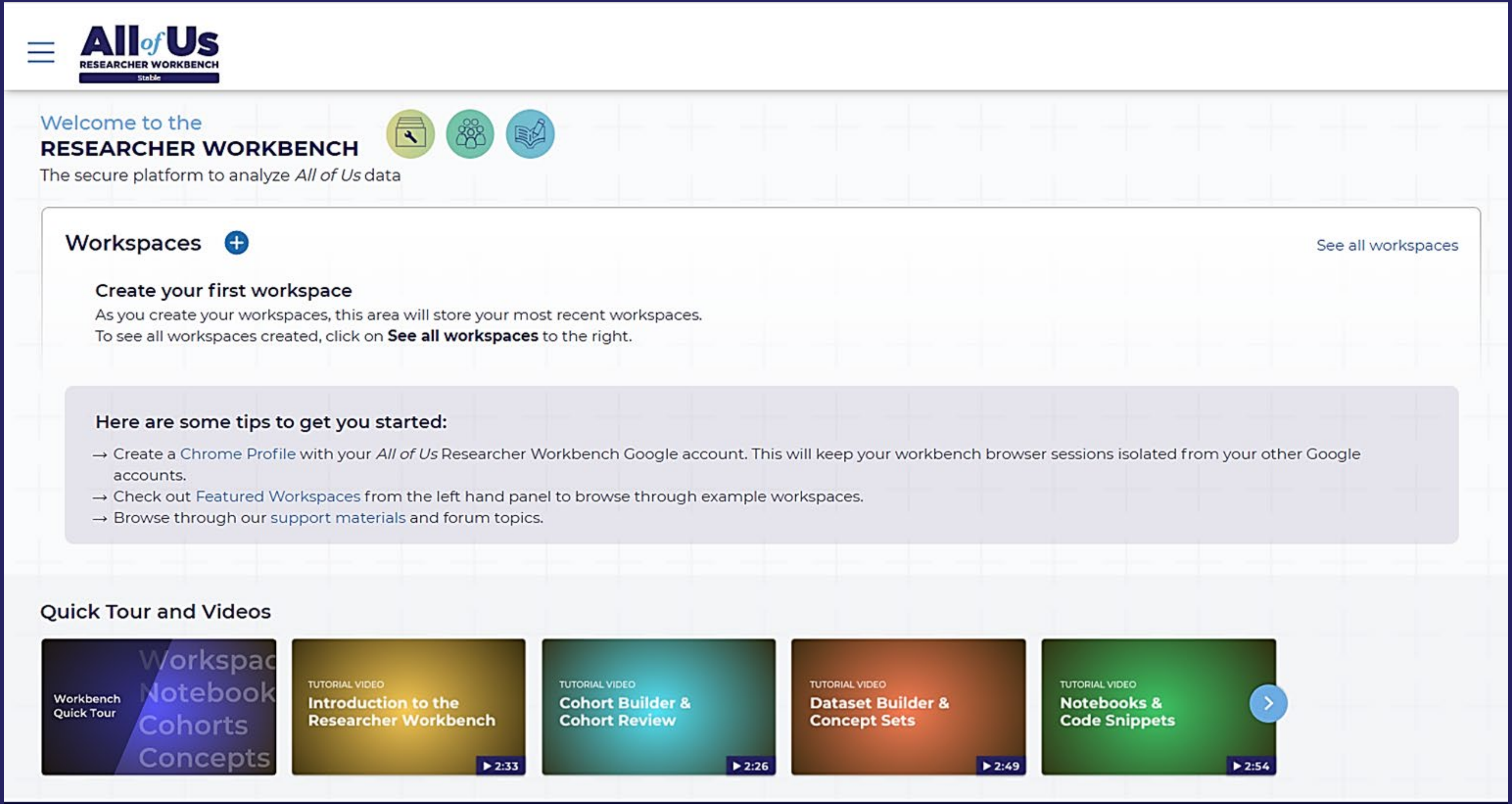
Adrienne Roman

User ID

onboarding@stable.fake-research-aou.org

Date

10/28/2020



Welcome to the

RESEARCHER WORKBENCH

The secure platform to analyze *All of Us* data

Workspaces

Create your first workspace

As you create your workspaces, this area will store your most recent workspaces.
To see all workspaces created, click on **See all workspaces** to the right.

Here are some tips to get you started:

→ Create a Chrome Profile with your *All of Us* Researcher Workbench Google account. This will keep your workbench browser sessions isolated from your other Google accounts.
→ Check out Featured Workspaces from the left hand panel to browse through example workspaces.
→ Browse through our support materials and forum topics.

Quick Tour and Videos

Thank you for attending the *All of Us* Researcher Onboarding Walkthrough breakout session!

Please email support@researchallofus.org with additional questions about the registration process