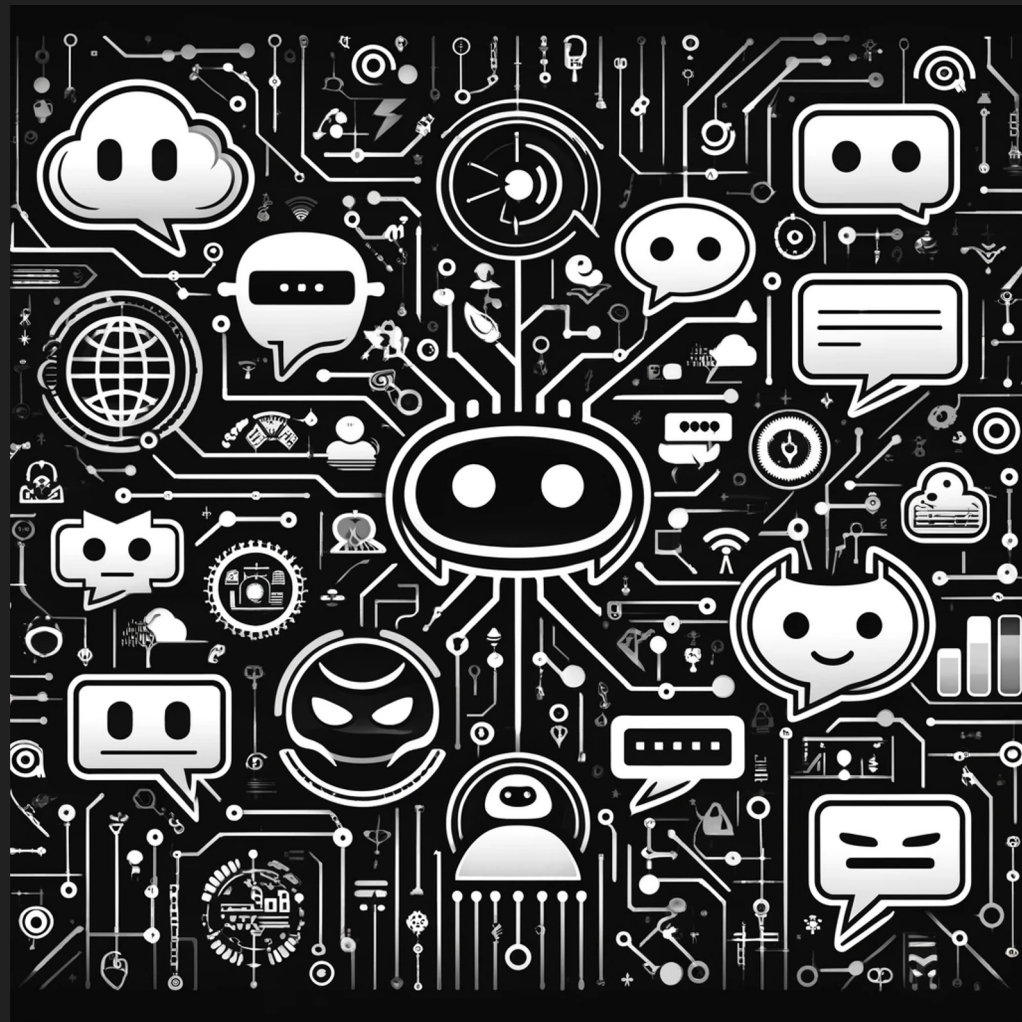
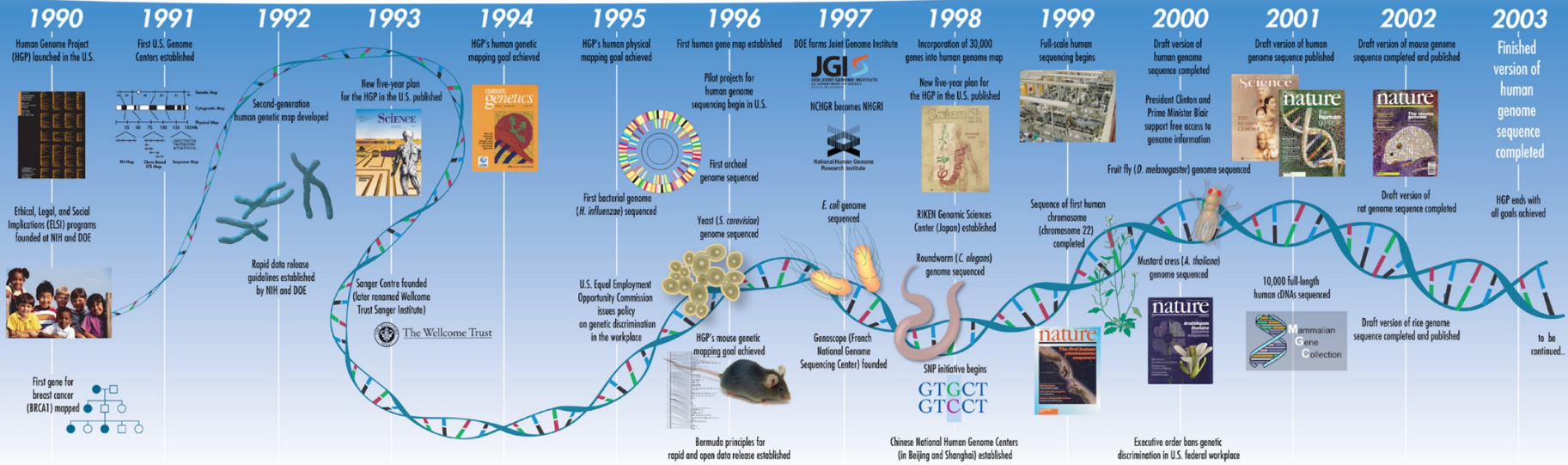
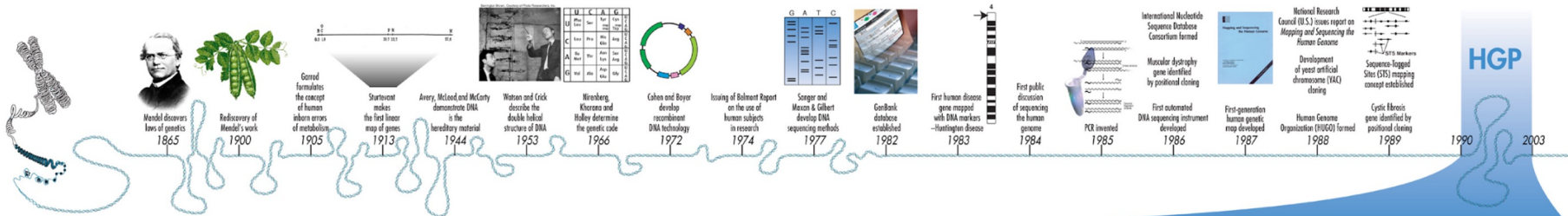




21st Century







The Human Genome Project



**Sequenced
92% of the
human genome
(<400 gaps)***



**Global effort:
20 institutions
in 6 countries**



**\$3 Billion
over 13 years
(1990-2003)**



**12 anonymous
volunteers**

+ innovation to recruit,
acquire informed consent,
and collect samples

WANTED
20 Volunteers
to participate in the
Human Genome Project
a very large international scientific research effort.

The goal is to decode the human hereditary information (*human blueprint*) that determines all individual traits inherited from parents. The outcome of the project will have tremendous impact on future progress of medical science and lead to improved diagnosis and treatment of hereditary diseases.

Volunteers will receive information about the project from the Clinical Genetics Service at Roswell Park, and sign a consent form before participating.

No personal information will be maintained or transferred.

Volunteers will provide a one-time donation of a small blood specimen. A small monetary reimbursement will be provided to the participants for their time and effort.

Individuals must be at least 18 years of age.
Persons who have undergone chemotherapy are not eligible.

ROSSELL PARK
CANCER INSTITUTE

For more information please contact the
Clinical Genetics Service
845-5720 (9:00 am - 3:00 pm)
March 24 - 26, 1997

A 1997 Buffalo, New York newspaper advertisement recruiting volunteers to provide blood samples and DNA for the Human Genome Project. (NHGRI History of Genomics Program Archive)

*The first truly complete human genome sequence produced by the Telomere-to-Telomere (T2T) consortium in 2022

[illegible]

Factors of Risk in the Development of Coronary Heart Disease— Six-Year Follow-up Experience

The Framingham Study

WILLIAM B. KANNEL, M.D., THOMAS R. DAWBER, M.D., F.A.C.P.,
ABRAHAM KAGAN, M.D., F.A.C.P., NICHOLAS REVOTSKIE, M.D.,
AND JOSEPH STOKES, III, M.D.
Framingham, Massachusetts



Framingham Heart Study

INCREASINGLY RELIABLE ESTIMATES of the prevalence and incidence of coronary heart disease (CHD) emphasize the importance of this disease as a contemporary health hazard. Cardiovascular disease is now the leading cause of death in the United States. The incidence of coronary heart disease of all heart disease in the diagnosis of CHD has increased in the past decade, no doubt and is expected to continue. This slight increase in incidence which has been substantiated by many studies. Because of the fact that CHD is a "silent" infarct, mortality in hospital is still the best therapeutic preventive measure.

Received for publication from the Framingham Heart Study, National Institute of Health, U. S. Department of Health, Education and Welfare, Bethesda, Md. Presented at The American Heart Association, 1961, Baltimore. Requests for reprints to Thomas R. Dawber, Heart Disease and Stroke Prevention Branch, National Heart Institute, Bethesda, Md.

Since it has been established that coronary atherosclerosis is present for many years prior to the development of symptomatic CHD, it seems evident that efforts at prevention must begin many years before the onset of CHD. At present, the only

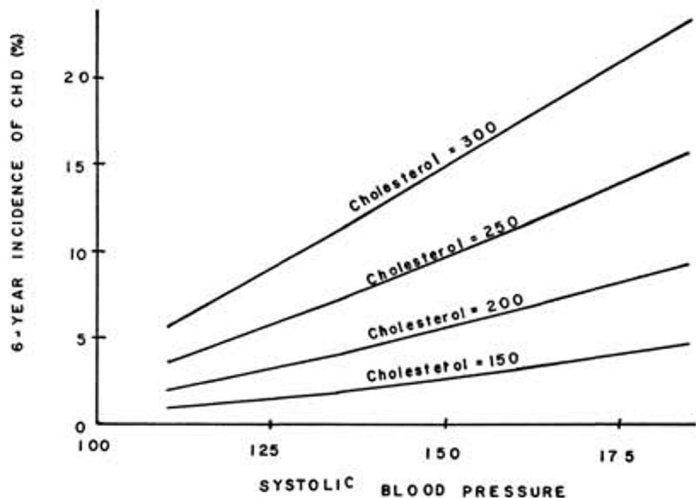


FIGURE 2. Six-year incidence of coronary heart disease according to level of systolic blood pressure at specified serum cholesterol levels (men 45 to 62 years). For explanation, see legends for Figure 1.

Enrolled 5,209 men and women in 1948

Some Framingham early discoveries:

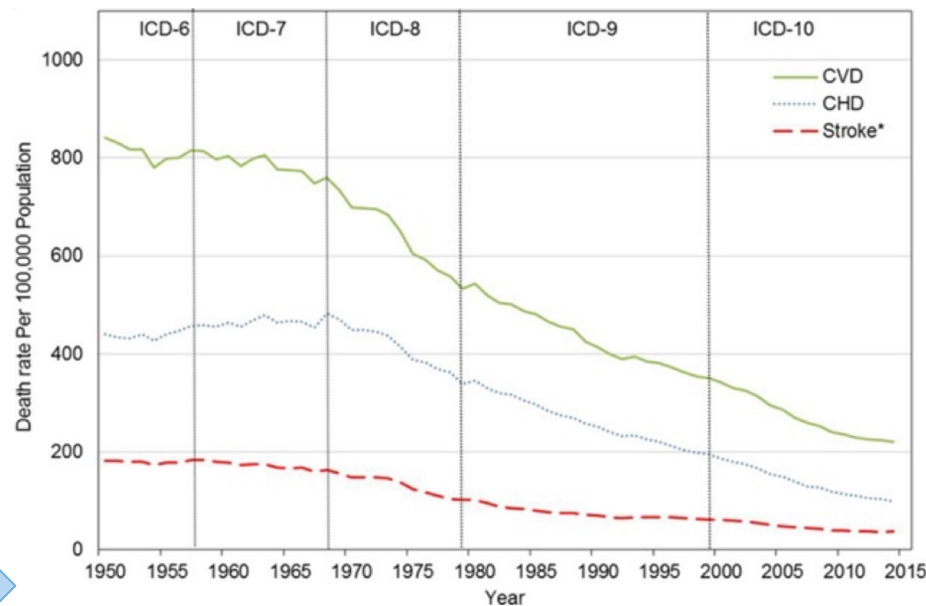
- 1960 – Cigarettes increase heart disease
- 1961 – cholesterol, blood pressure increase heart disease
- 1967 – exercise decreases risk of heart disease; obesity increases it
- 1970 – high blood pressure and atrial fibrillation cause stroke

The Strength of Large Cohort Studies

The impact of Framingham (and similar cohorts) has been dramatic

From 1950 - 1996: Heart disease mortality fell 56%, stroke rates fell by 70%

Since 1990, heart disease mortality has continued to fall by 22%



Mensah et al. *Circ Res.* 2017

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

All of Us Research Program

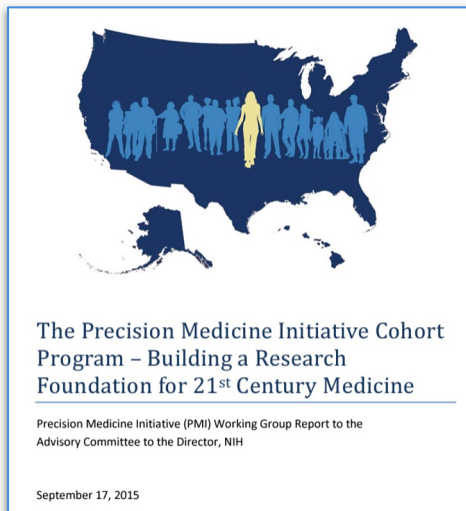
May 17, 2024

Chris Lunt
Chief Technology Officer



What is the NIH *All of Us* Research Program?

The *All of Us* Research Program is a historic, longitudinal effort to **gather data from one million or more people** living in the United States **to accelerate research and improve health**. By taking into account individual differences **in lifestyle, socioeconomic, environment, and biology**, we hope that researchers will uncover paths toward delivering **precision medicine – or individualized prevention, treatment, and care – for all of us**.

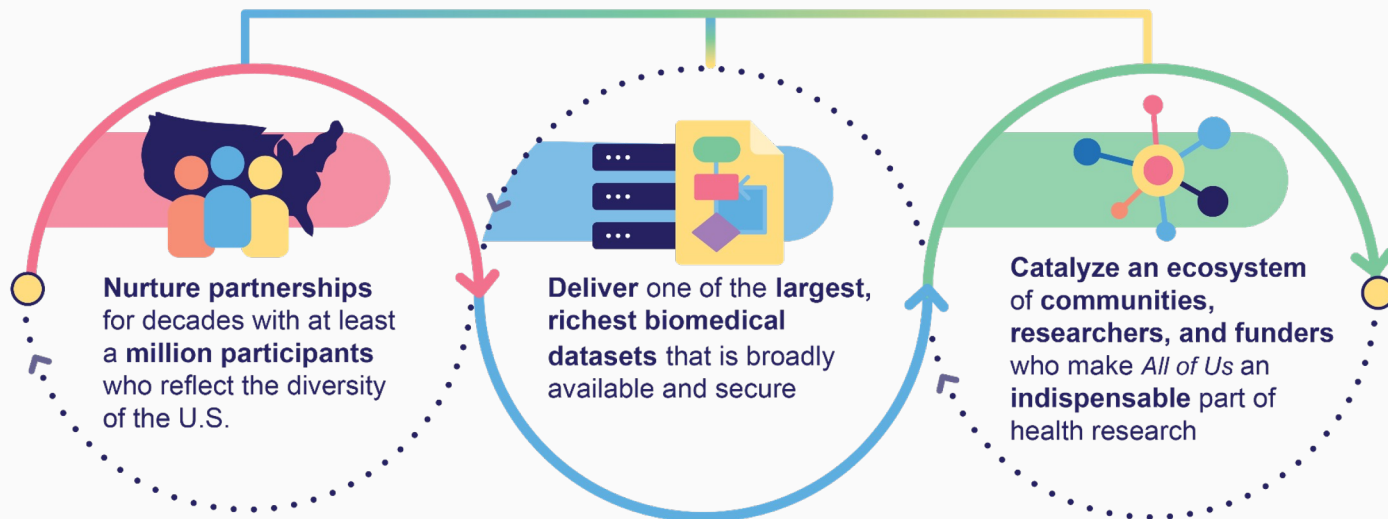


“All of Us is among the most ambitious research efforts that our nation has undertaken!”

NIH Director Francis Collins, M.D., Ph.D.

The *All of Us* Research Program Mission

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



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

Participants

Enrolled 793K+ Participants With Continued Growth

Participant Enrollment

793,000+

Participants

436,000+

Electronic Health
Records

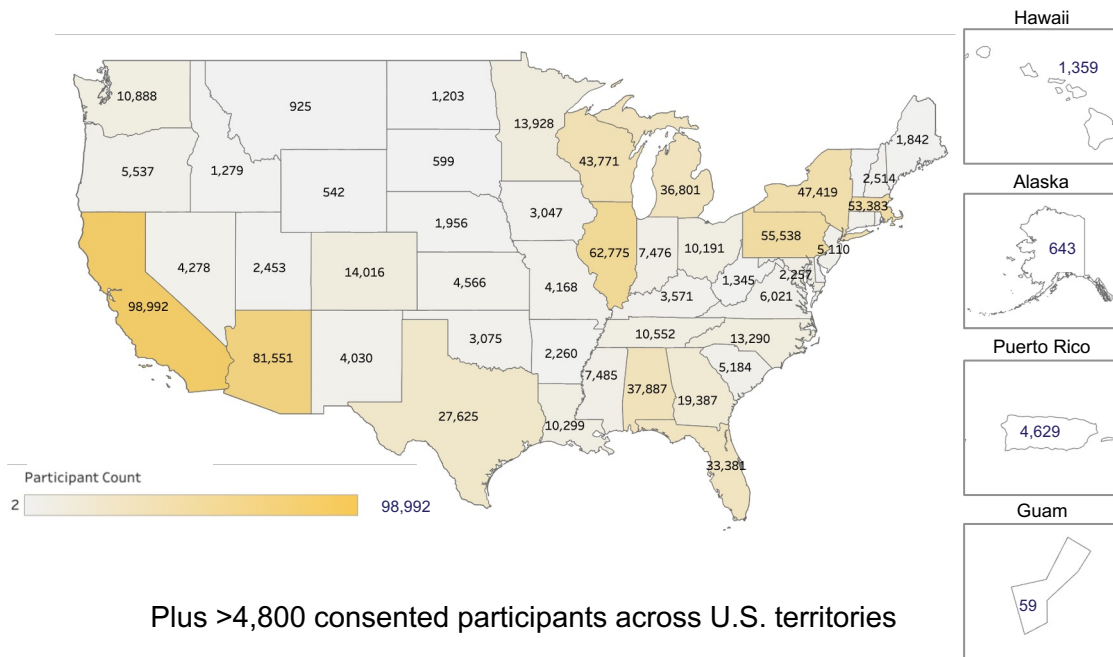
545,000+

Participants who have
completed initial steps of the
program

563,000+

Biosamples

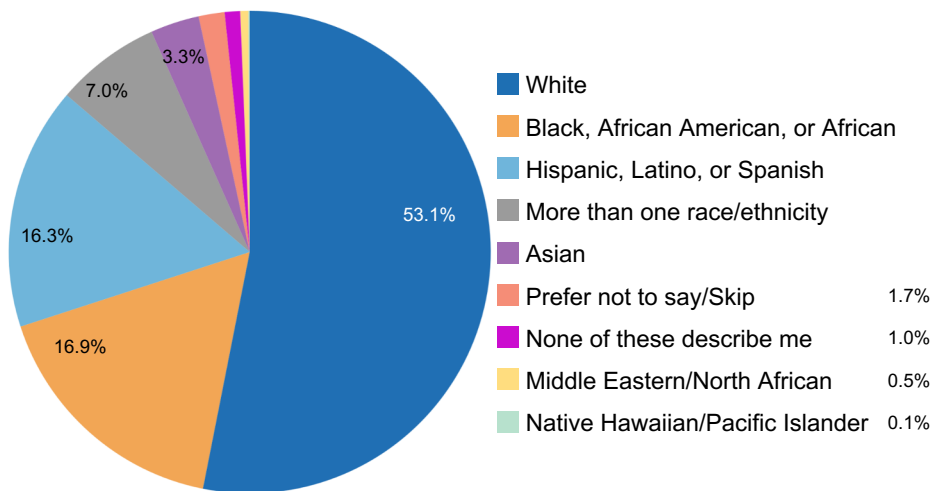
Map of Consented Participants



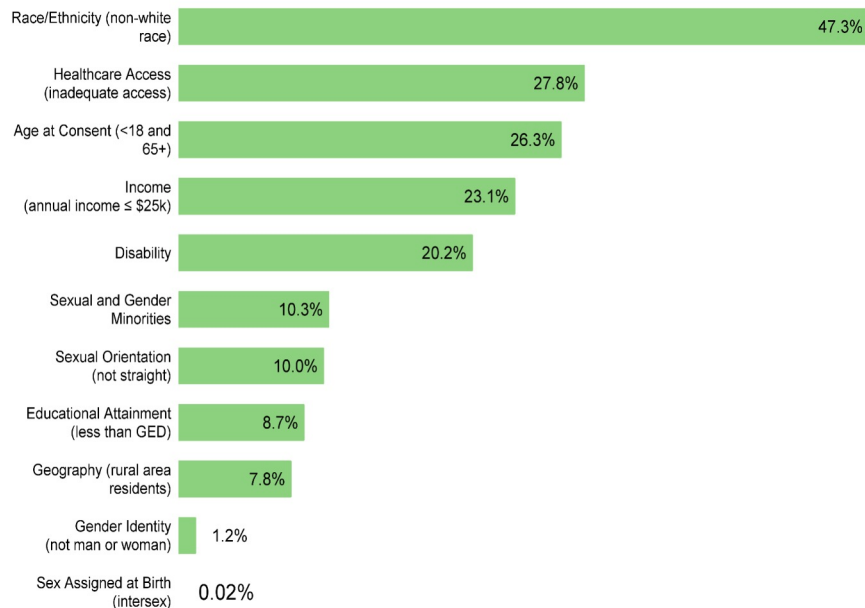
Numbers current as of May 1, 2024

Participant Diversity

Race & Ethnicity of Participants



UBR Category



Over 87% of *All of Us* participants are underrepresented in biomedical research

Numbers current as of May 1, 2024

Data

Data Types Collected from *All of Us* Participants



Electronic Health Records

Data types collected from EHR include:

- Demographics
- Vital signs
- Diagnoses
- Procedures
- Medications
- Doctor and Laboratory Visits



Participant Surveys

The Basics Health Care Access & Utilization Mental Health and Well-Being
Overall Health Personal and Family Medical History
Lifestyle Social Determinants of Health



Physical Measurements

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



Biosamples

- Blood
- Saliva
- Urine



Wearable Data

Fitbit data, including:

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

Data Collected and Return of Value to *All of Us* Participants



Consent and Electronic Health Records



Participant Surveys



Physical Measurements



Biosamples

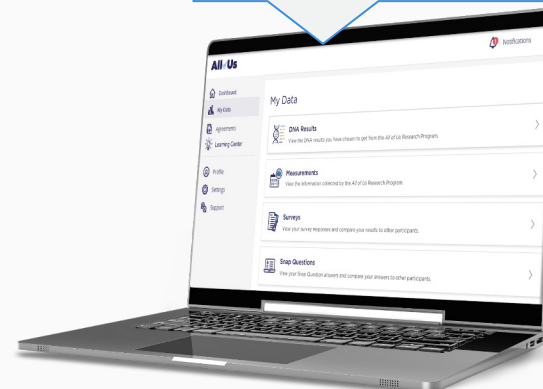
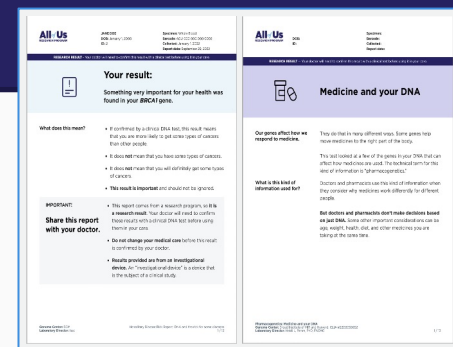


Mobile/Wearable Tech

Return of Value for Participants

Participants may receive:

- **Genetic information**
- Survey data (comparative)
- EHR and claims data
- Ongoing study updates
- Aggregate results
- Scientific findings
- Opportunities to be contacted for other research opportunities



v7 of the Curated Data repository



400K+

Survey
Responses



245K

Whole Genome
Sequences



300K+

Genotyping
Arrays



1K

Long-Read Sequences

NEW! In 2023



300K+

Physical
Measurements



11K

Structural Variants

NEW! In 2023



280K+

Electronic Health
Records



16K

Fitbit
Records

NEW! Sleep Data

Genomic Data Is Paired With Rich Phenotypic Data



206,100+

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



245,100+

Have Whole Genome Sequences + Physical Measurements +
Survey Responses



206,150+

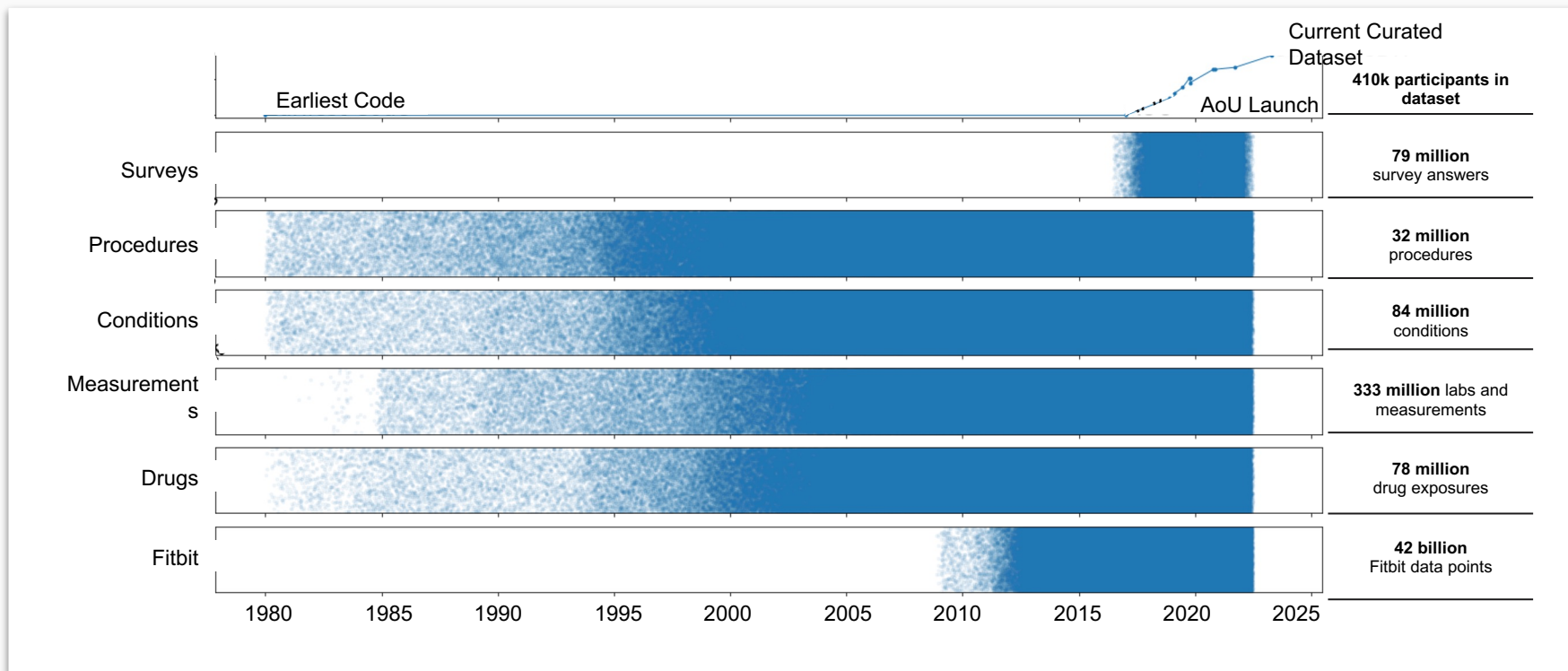
Have Whole Genome Sequences + Electronic Health Records



8,800+

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

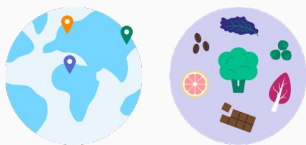
EHR Data on the Researcher Workbench is Diverse and Longitudinal



Participants Can Receive Four Types of Genetic Research Results

Engaging insights

Genetic ancestry and traits results



7 regions (20 subregions) and 4 traits

- Sub-Saharan Africa
- Europe
- Oceania
- Southern Asia
- Eastern and northern Asia
- The Middle East and North Africa
- The Americas
- Ear wax
- Bitter taste perception
- Cilantro preference
- Lactose intolerance

Health results

Medicine and your DNA Report



7 pharmacogenomics (PGx) genes and 50+ medications

- *CYP2C19*
- *DPYD*
- *G6PD*
- *SLCO1B1*
- *NUDT15*
- *TPMT*
- *UGT1A1*

Hereditary Disease Risk (HDR) Report



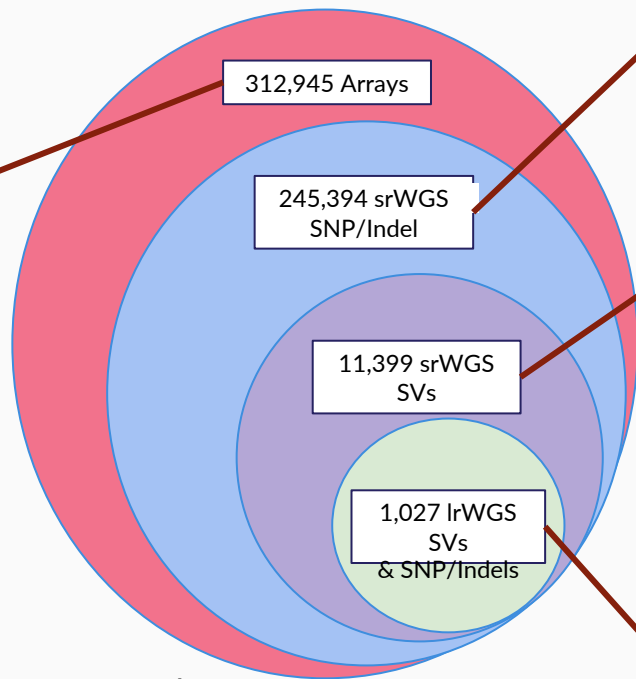
59 genes (SNVs + indels, e.g. no SVs)

- Breast cancer
- Ovarian cancer
- Uterine cancer
- Colorectal cancer
- Prostate cancer
- Melanoma
- Brain cancer
- Pancreatic cancer
- Stomach cancer
- Familial hypercholesterolemia
- Cardiomyopathies
- Arrhythmias
- Arteriopathies

Genomics



Illumina *Infinum*™
Global Diversity Array



srWGS = short read whole genome sequence

SNP = single nucleotide polymorphism

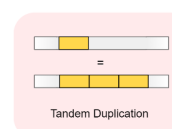
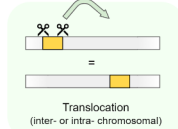
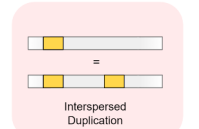
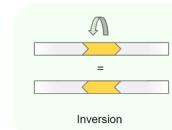
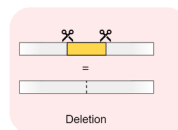
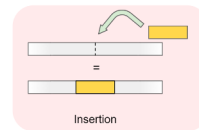
Indel = short inserts and/or deletes

SVs = structural variations **new!**

SV Types

Destructive (non-balanced)

Non-destructive (balanced)



New! Hail VDS file format for srWGS data along with VCFs, Hail Matrix Tables, and PLINK files for smaller callsets.

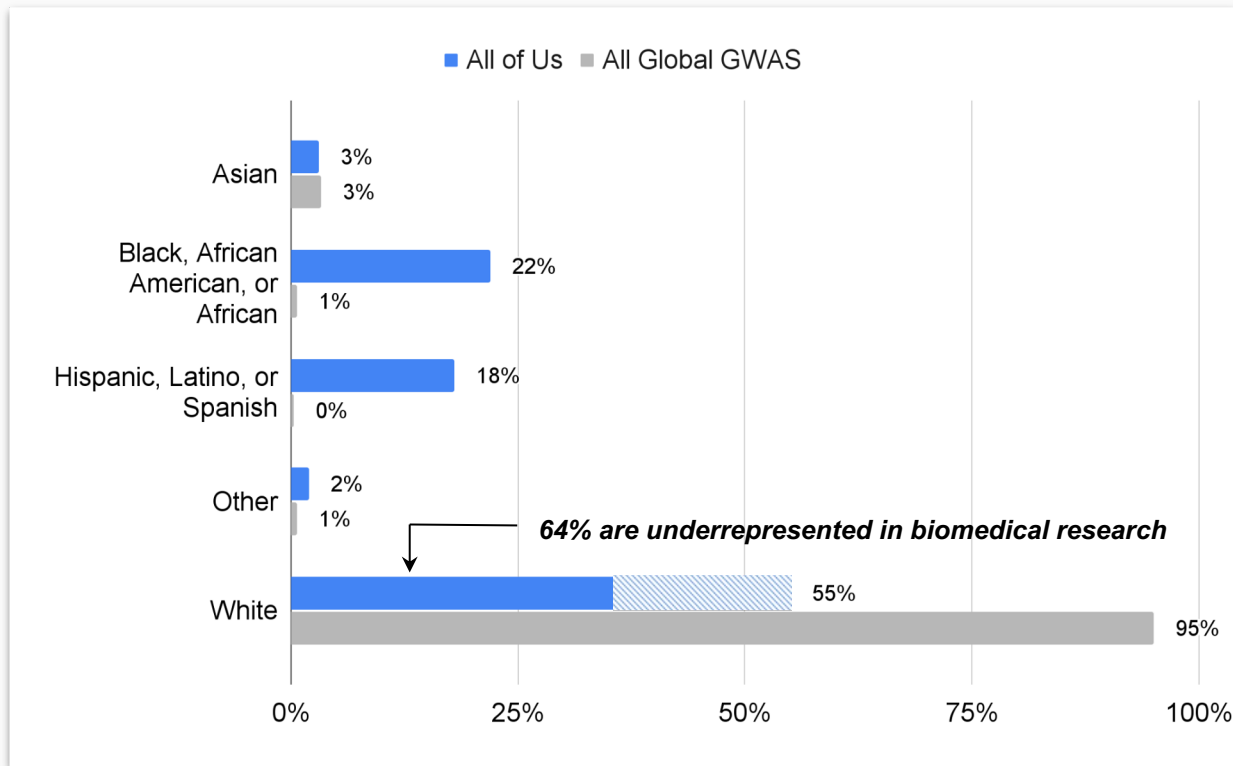
lrWGS = long read whole genome sequence (all African American) **new!**

All of Us is Enhancing Diversity of Genomic Studies

Updated genomic data set

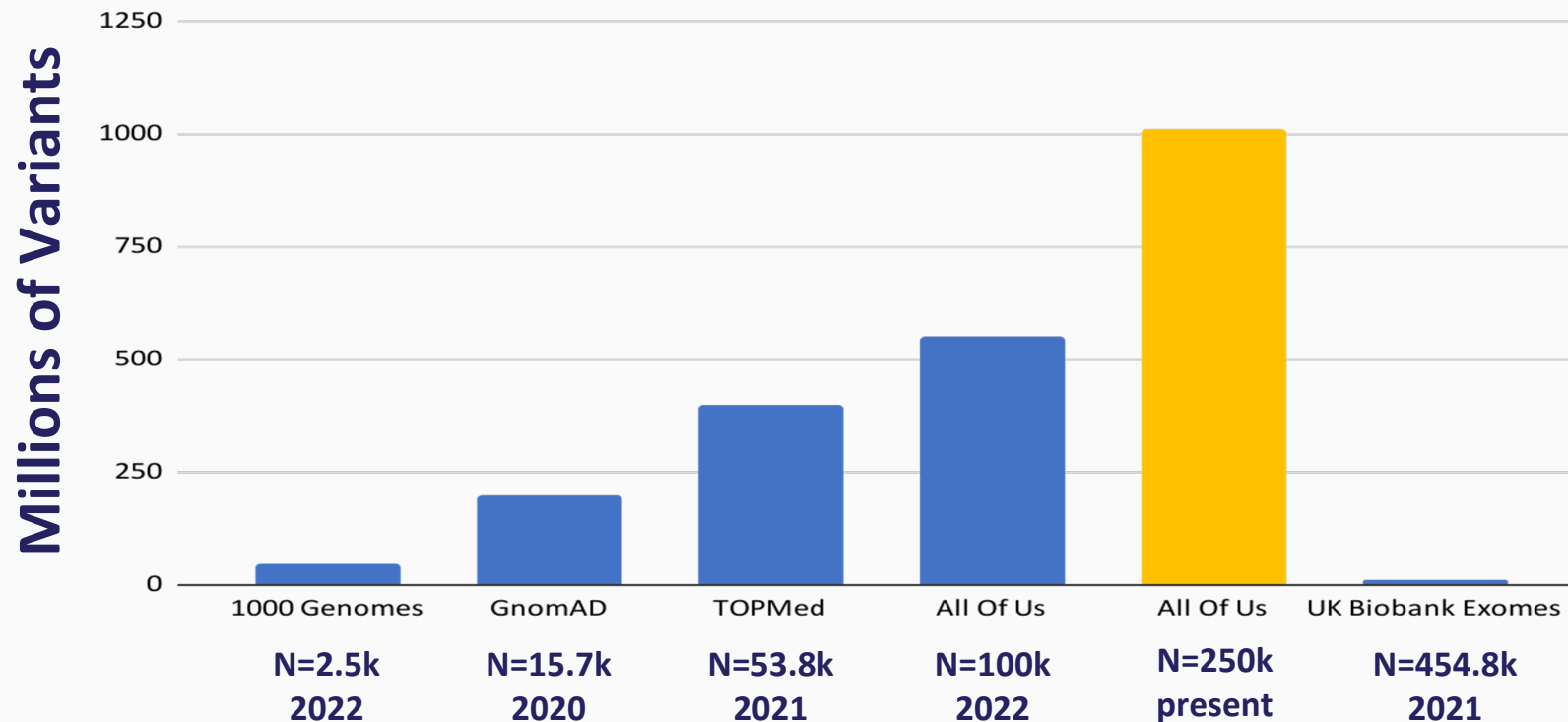
~45% diverse by race/ethnicity

77% underrepresented in biomedical research



All Global GWAS values from www.gwasdiversitymonitor.com. Values current as of November 2023

Discovery of Rare Variation



All of Us Data Includes Range of Diseases

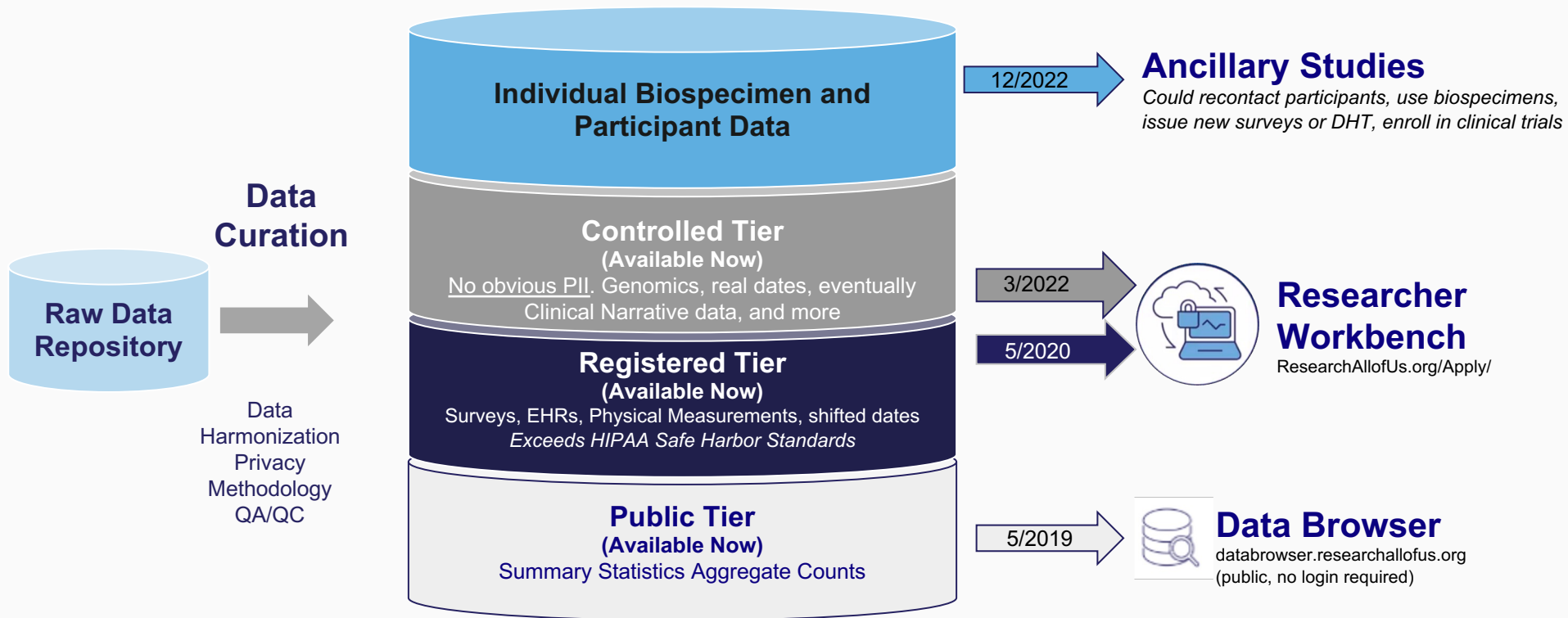
Conditions	Domain	Participants
Heart Disease	Heart	89,180
Obesity	Endocrine	80,980
Type 1 Diabetes	Endocrine	6,740
Type 2 Diabetes	Endocrine	56,340
Any cancer	Cancer	50,720
Asthma	Pulmonary	50,160
Chronic Obstructive Pulmonary Disorder	Pulmonary	24,940
Epilepsy	Neuro	9,320
Stroke	Neuro	940
Rheumatoid Arthritis	Autoimmune	8,660
Osteoarthritis	Autoimmune	99,880

Conditions	Domain	Participants
Depressive Disorder	Mental Health	82,660
Bipolar Disorder	Mental Health	16,200
Dementia	Mental Health	6,020
Human Immunodeficiency Virus	Infectious Disease	5,500
COVID-19*	Infectious Disease	58,000*
Alcoholism	Abuse	20,380
Opioid Usage	Medication	155,000
Age-Related Macular Degeneration	Eye	4,740
Hearing loss	Hearing	30,400
Falls	Aging/Nursing	4,860

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

Researchers

Researcher Data Access

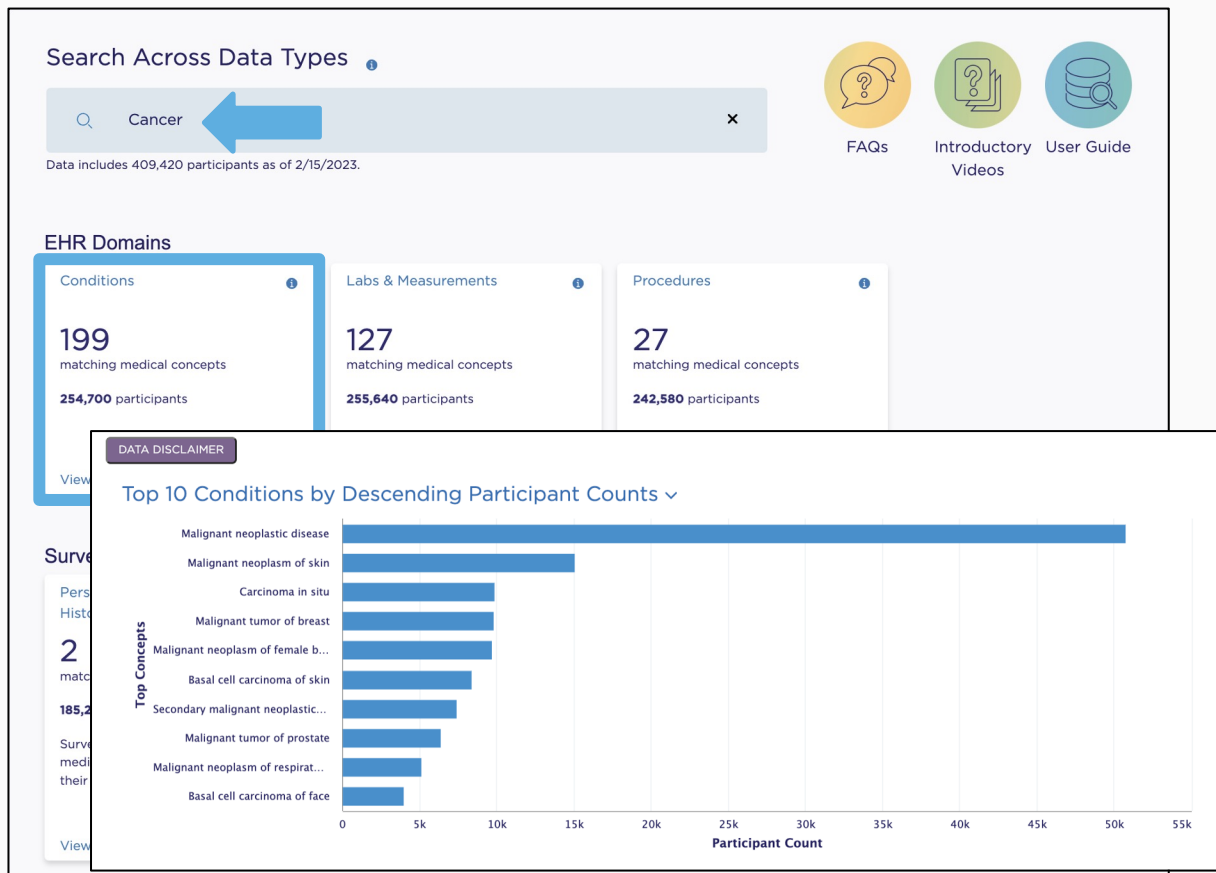


All of Us Research Hub: Public Data Browser

Summary statistics of:

- EHR Data (Conditions, Drug Exposures, Lab & Measurements, Procedures)
- Genomic Variants
- Survey Questions (including COVID-19 surveys)
- Physical Measurements
- **Open Access (no login required)**

Data Browser



All of Us Researcher Workbench: Access to Row-Level Data for Analysis



CONFIRM YOUR INSTITUTION'S AGREEMENT

Before you can create an account, your institution must have a Data Use and Registration Agreement (DURA) in place with *All of Us*. [Confirm DURA](#).



COMPLETE THE MANDATORY TRAINING

The training focuses on conducting responsible and ethical research using the Researcher Workbench. Additional training is required to access the Controlled Tier. [Learn more](#).



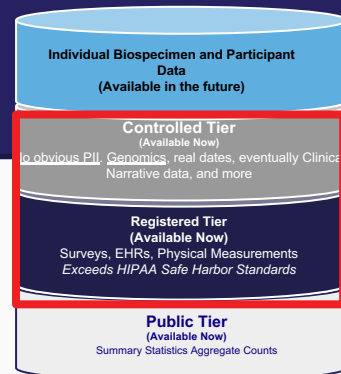
CREATE AN ACCOUNT AND VERIFY IDENTITY

After creating your Researcher Workbench account, you will be asked to verify your identity through login.gov. [Learn more](#).



SIGN THE DATA USER CODE OF CONDUCT (DUCC)

This agreement outlines the program's expectations for researchers who use the Researcher Workbench and describes how program data may be used. [View the DUCC](#).

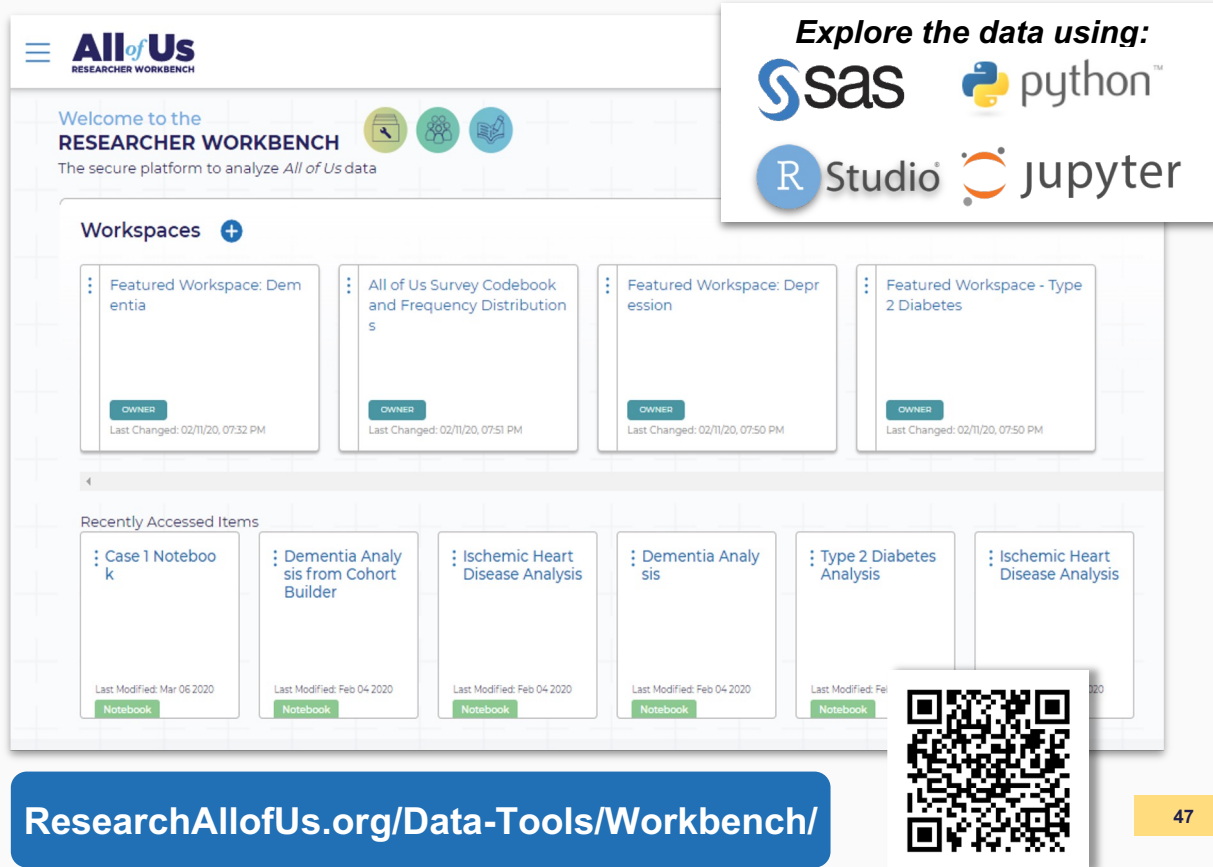


Researcher Workbench

All of Us Researcher Workbench: Access to Row-Level Data for Analysis

Researcher Workbench

- Cloud-based platform
- Graphical tools for creating cohorts and datasets
- Open to researchers at academic, not-for-profit, and health care institutions
- **Passport access model** with institutional agreements for all researchers
- **No separate IRB approval needed**



The screenshot displays the All of Us Researcher Workbench interface. At the top, the logo "All of Us RESEARCHER WORKBENCH" is visible. Below it, a welcome message reads: "Welcome to the RESEARCHER WORKBENCH The secure platform to analyze All of Us data". To the right, a box titled "Explore the data using:" lists supported tools: SAS, Python, R Studio, and Jupyter. The main content area is divided into two sections: "Workspaces" and "Recently Accessed Items". The "Workspaces" section shows four featured workspaces, each with a title, a status (OWNER), and a last changed timestamp. The "Recently Accessed Items" section shows six recently accessed notebooks, each with a title, a status (Notebook), and a last modified timestamp. A QR code is located in the bottom right corner of the interface.

Explore the data using:

SAS Python R Studio Jupyter


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: Feb 04 2020
Notebook

ResearchAllofUs.org/Data-Tools/Workbench/



March 2021, Patrick Wu Defends His PhD Dissertation: Repurposing drugs using gene expression signatures and EHR data



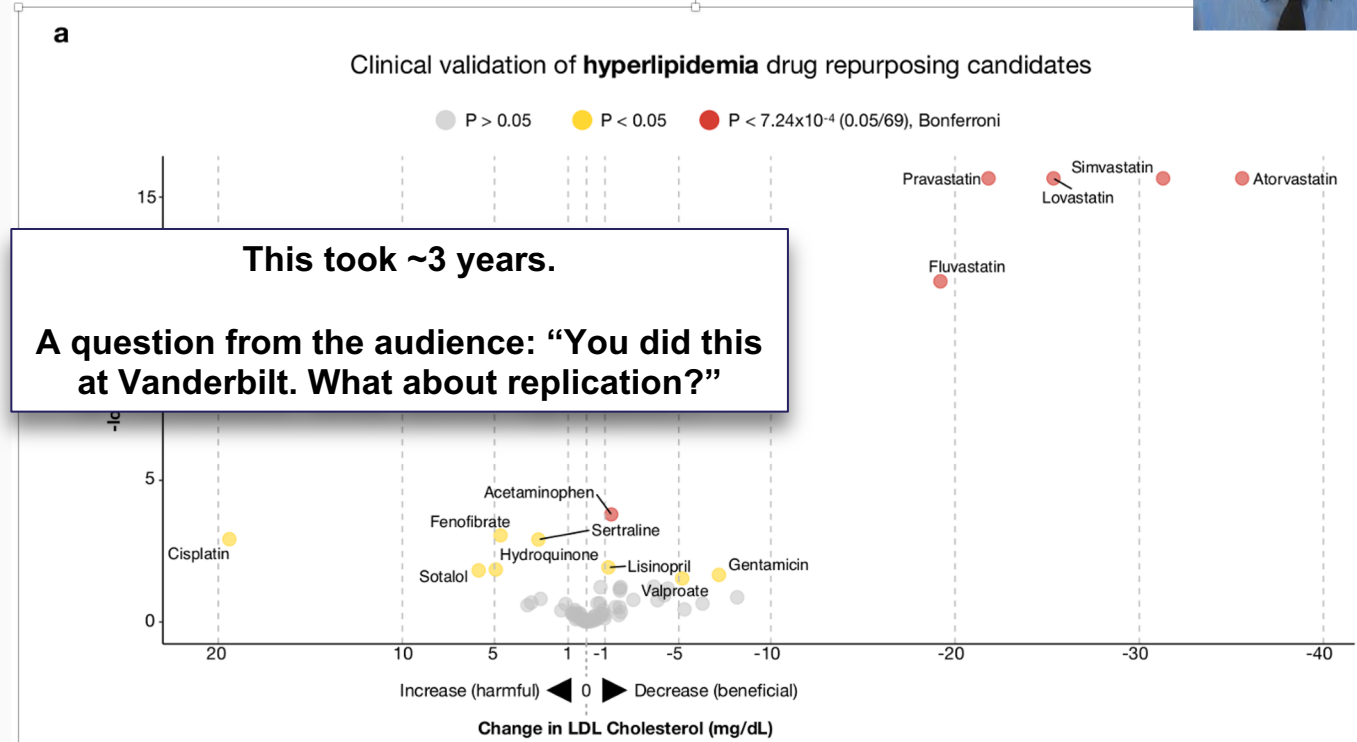
gene expression data identifies
drugs that might lower
cholesterol



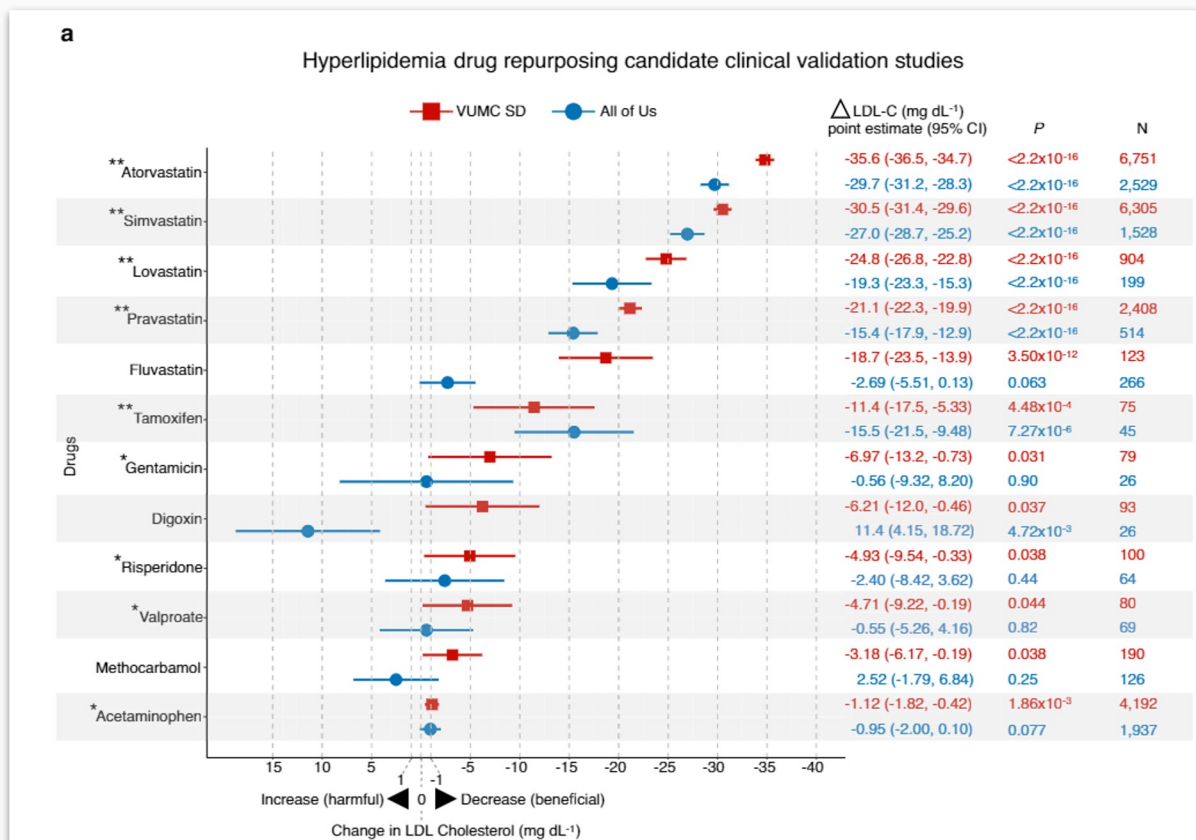
found 69 candidate drugs



study those drugs in the
Electronic Health Record data
at Vanderbilt



One Week Later, he emails his advisor this figure including *All of Us* results...



All of Us Research Program's Commitment to Researcher Diversity (as of May 1, 2024)



10,680+
Registered
Researchers



10,640+
Active
Projects



410+
Publications
using *All of Us*
data



760+
Organizations
43 Historically
Black Colleges &
Universities
69 Hispanic
Serving Institutions



80
International
Institutions



**Top
Conditions
Studied**
Cardiovascular
disease,
Hypertension,
Mental Health,
Cancer, Diabetes

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



Baylor's UBR Faculty Summit

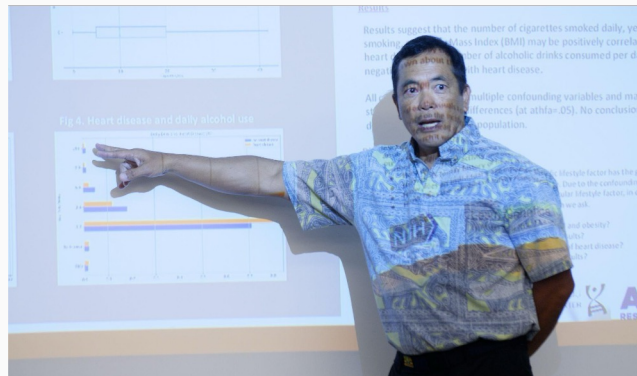
Over 78% of our researchers are underrepresented in the biomedical workforce - including over 30% diverse by race and ethnicity

You can even find *All of Us* at High School Science Fairs now!

High School Students are now using the *All of Us* Researcher Workbench

5 High Schools

15+ projects currently underway



Utah's Summer Teacher Institute 2023

Annual multi-day workshop designed to introduce high school teachers to *All of Us*



Research Team Sci

Can Antidepressants F

Predictive Analysis of Sjogren's Syndrome: A Machine Learning Analysis

Clonal Hematopoiesis of Indeterminate Potential and Solid Tumors in Breast Cancer

He presented his work on SDOH and cardiovascular risk in diabetics. (Paul FYI -- re AMIA and his nursing home project 🤖). Prett

Jack Cummins

High School Student, Manchester Essex Regional H

Jack Cummins is an 11th grader with a passion for research, coding, and bioinformatics. He has been selected to present his research at the Sigma Xi IFORE conference. His research projects applying natural language processing to medical research on predictive modeling and Jack is thrilled to be part of the *All of Us* community! At Manchester Essex Regional High School, he has won several awards and is president of the math team. In addition, he worked with his technology teacher to create a new Python course for the high school. Jack is also a leader in his high school community, as president of the class of 2025 and captain of the tennis team, founder of the French Club and a member of the chess team. He is currently working on a project, where he collaborates with local nursing and rehabilitation facilities, by visiting nursing homes and

Output

Growing Scientific Impact



researchallofus.org/publications/

Clinical Research | www.jasn.org

OPEN

Genetic Inhibition of APOL1 Pore-Forming Function Prevents APOL1-Mediated Kidney Disease

Adriana M. Hung^{1,2}, Victoria A. Assimon³, Hua-Chang Chen^{1,4}, Zhihong Yu^{1,4}, Caitlyn Vlasschaert⁵, Jefferson L. Triozi⁶, Helen Chan⁷, Lee Wheelless⁸, Otis Wilson^{1,2}, Shailja C. Shah⁷, Taralynn Mack^{1,8}, Trevor Thompson⁹, Michael E. Matheny^{1,4,9}, Saranya Chandrasekar¹⁰, Sahar V. Mozaffari⁷, Cecilia P. Chung¹⁰, Philip Tsao^{11,12}, Katalin Susztak^{1,3}, Edward D. Siew^{1,2}, Karol Estrada³, J. Michael Gaziano^{14,15}, Robert R. Graham⁷, Ran Tao^{1,4}, Maarten Hoek^{1,4}, Cassianne Robinson-Cohen², Eric M. Green² and Alexander G. Bick^{1,14} for the Million Veteran Program*

nature cardiovascular research

Letter

High-proportion spliced-in titin truncating variants in African and European ancestry in the All of Us Research Program

Received: 29 September 2023 | Naman S. Shetty¹, Akhil Pampana², Nirav Patel¹, Peng Li³, Garima Arora⁴ & Parag Arora^{2,5}
Accepted: 19 December 2023
Published online: 13 January 2024

Health disparities in the treatment of bipolar disorder

Vladimir Tchikrizov^a, Mark E. Ladner^a, Felicia V. Caples^b, Mitzi Morris^c, Hailey Spillers^c, Christina D. Jordan^a, Joyce E. Balls-Berry^d, Monica J. Taylor-Desir^e, Mark A. Frye^e, Eric J. Vallender^{a,*}

Journal of Personalized Medicine

Perspective

All of Us and the Promise of Precision Medicine: Achieving Equitable Access for Federally Qualified Health Center Patients

Carolyn P. Neuhaus^{1,*}, Danielle M. Pacia¹, Johanna T. Crane², Karen J. Maschke¹ and Nancy Berlinger¹

Genomic data in the All of Us Research Program

<https://doi.org/10.1038/s41586-023-06957-x> | The All of Us Research Program Genomics Investigators*
Received: 22 July 2022

nature medicine

Article

Selection, optimization and validation of ten chronic disease polygenic risk scores for clinical implementation in diverse US populations

March 20, 2023

all of us JAMA Network

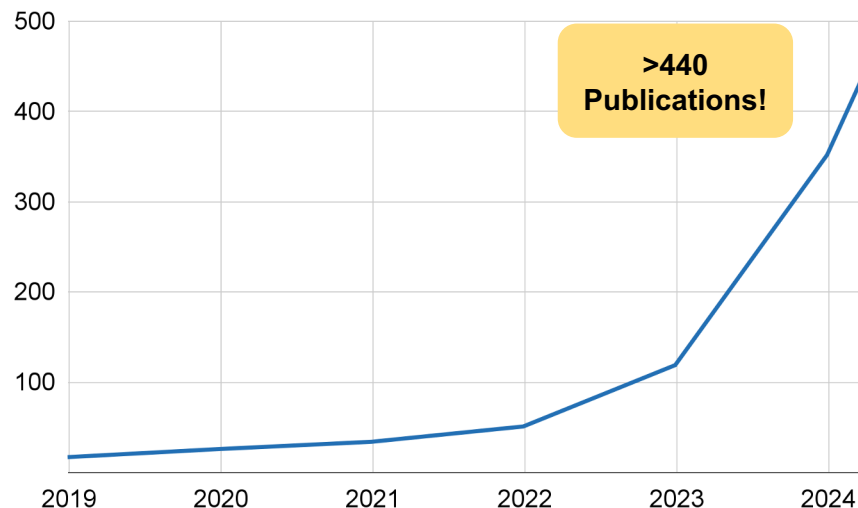
Daily Step Counts Before and After the COVID-19 Pandemic Among All of Us Research Participants

Stacy Desine, MSc¹; Hiral Master, PhD, MPTh, MPH²; Jeffrey Annis, PhD², et al

» Author Affiliations | Article Information

JAMA Netw Open. 2023;6(3):e233526. doi:10.1001/jamanetworkopen.2023.3526

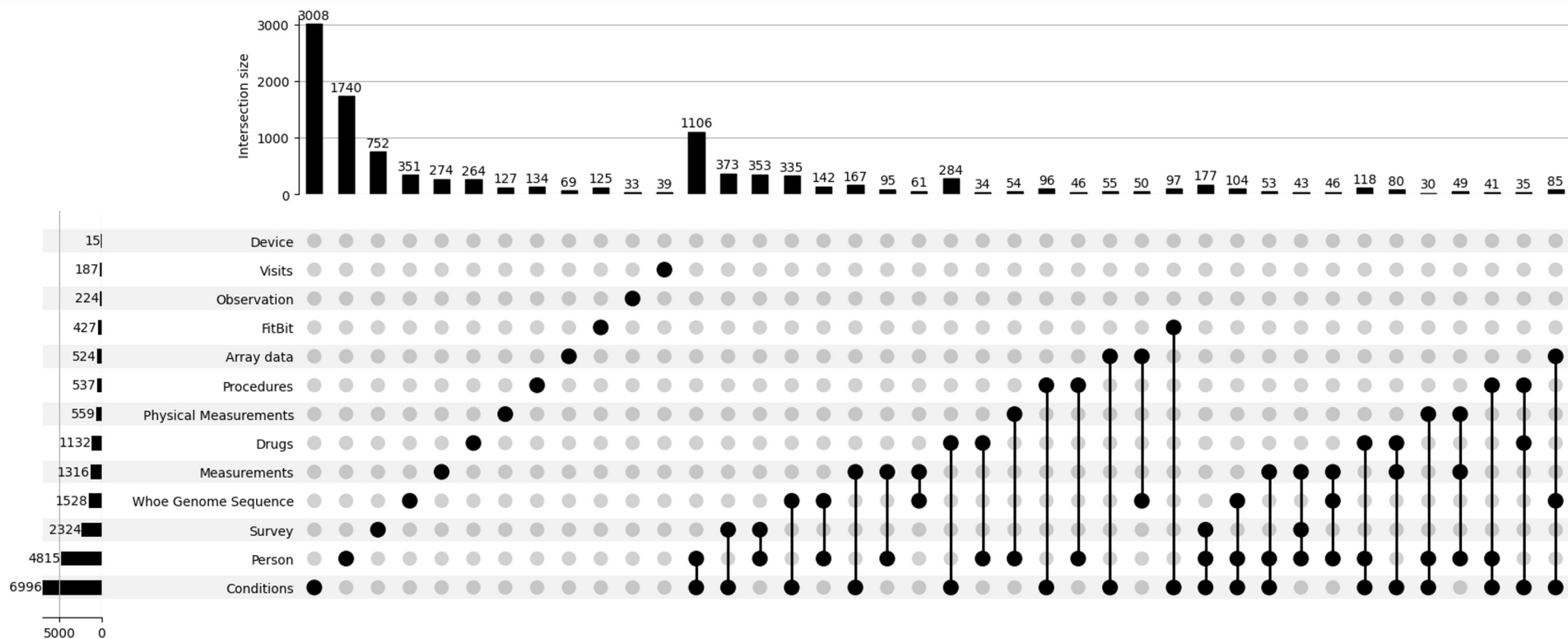
Cumulative Publications Using All of Us Data



Disciplinary spread



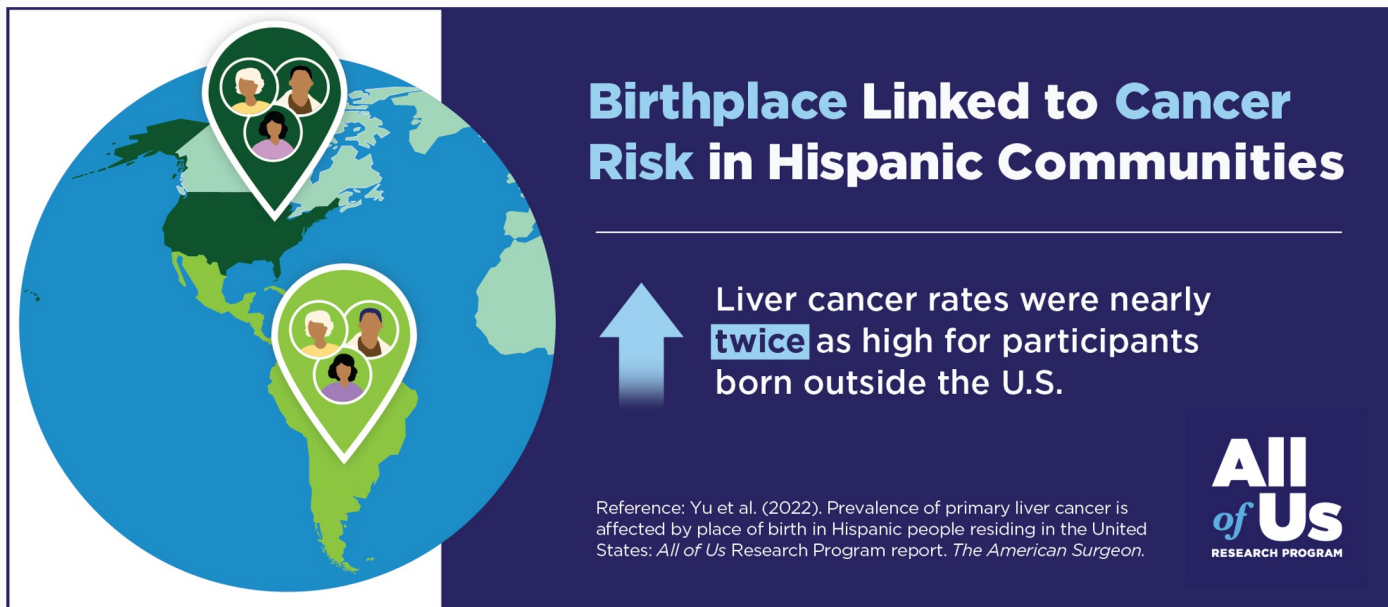
Data types included in cohorts



Plot restricted to show only intersections with counts of at least 30. Person = Demographics. Device is from EHR.

Created by Lina Sulieman

What stories are we telling already about the power of participant data?



This study highlights the importance of **diversity in health research** **within** underrepresented groups. Studies like this one helps health care providers move away from “one-size-fits-all” approaches to care.

Read more about this [here](#).

What stories are we telling already about the power of participant data?



Linked data Fitbit devices with data from electronic health records (EHRs).

The researchers **analyzed more than 1,700 medical conditions**

Taking at least 8,200 daily steps was associated with a lower risk for certain health conditions. This included chronic diseases like high blood pressure, obesity, and diabetes.

The study also uncovered **new links** between daily steps and lower risk for **sleep apnea, depression, and acid reflux**.

Many people shared **four years or more** of their activity data.

Read more about this [here](#).

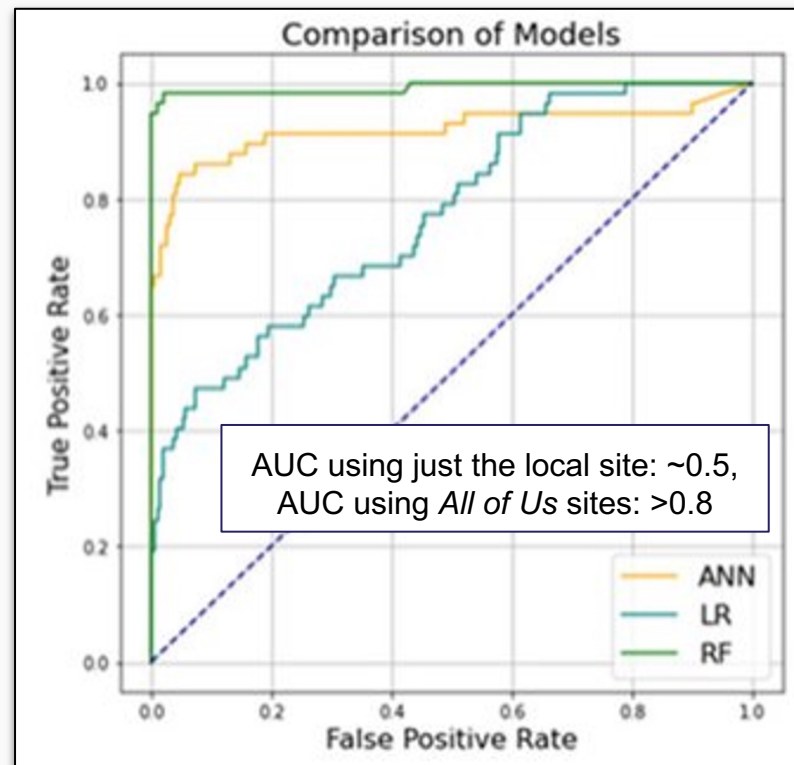
Machine Learning in *All of Us* to Predict Glaucoma Outcomes

Predictive Analytics for Glaucoma Using Data From the *All of Us* Research Program



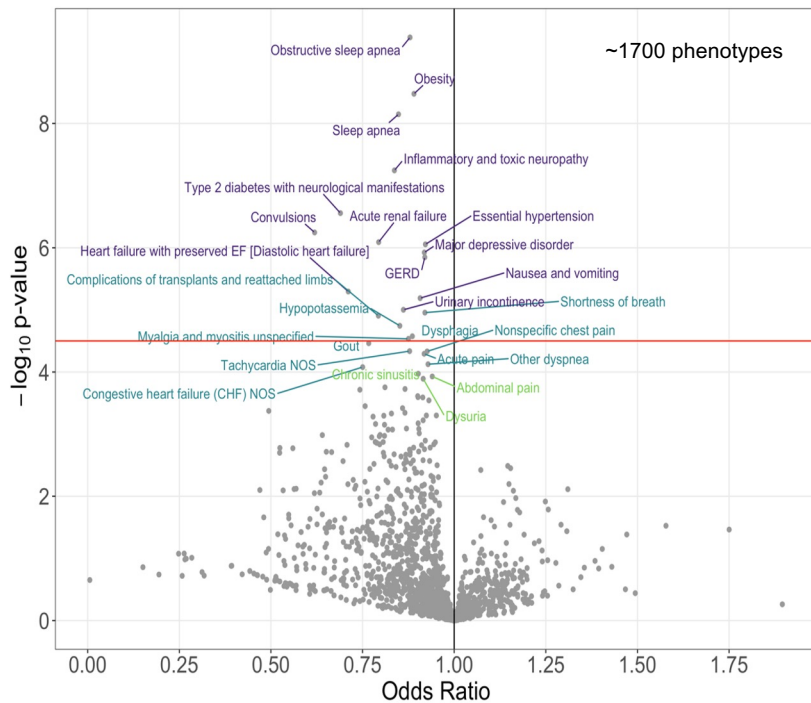
SALLY L. BAXTER, BHARANIDHARAN RADHA SASEENDRAKUMAR, PAULINA PAUL, JIHOON KIM, LUCA BONOMI, TSUNG-TING KUO, ROXANA LOPERENA, FRANCIS RATSIMBAZAFY, ERIC BOERWINKLE, MINE CICEK, CHERYL R. CLARK, ELIZABETH COHN, KELLY GEBO, KELSEY MAYO, STEPHEN MOCKRIN, SHERI D. SCHULLY, ANDREA RAMIREZ, AND LUCILA OHNO-MACHADO, ON BEHALF OF THE *ALL OF US* RESEARCH PROGRAM INVESTIGATORS

Conclusion: Models trained with multisite, multi-vender EHR data *All of Us* data achieved **superior performance** compared to using single-center data.

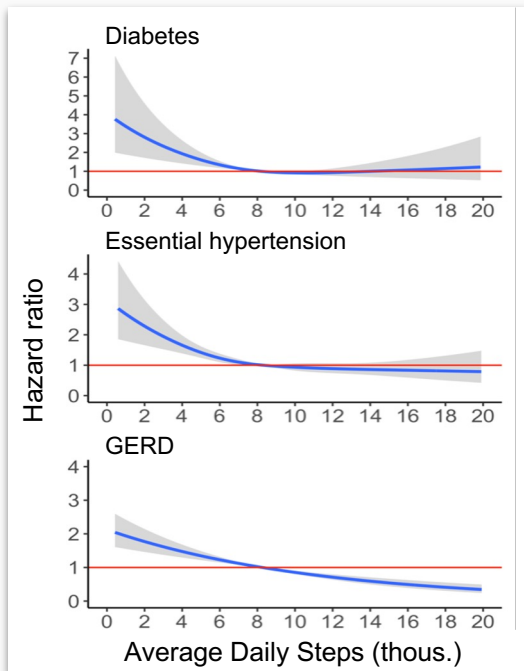


PheWAS of Physical Activity (Step Count) in *All of Us*

Incident Diagnoses after 6 months



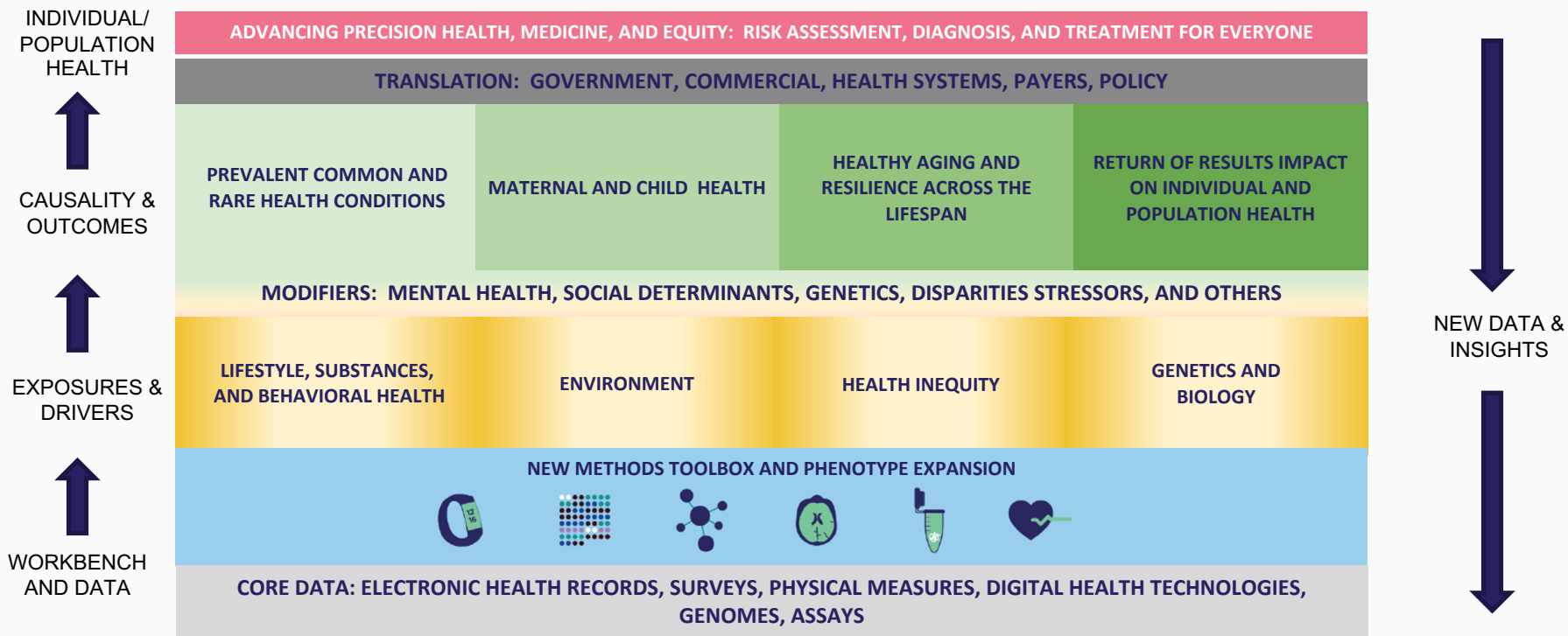
- Analysis uses average of daily steps *prior to diagnosis*.
- Effect size per 1000 steps, adjusted for age, sex, race



Adjusted for age, sex, race, CAD, cancer, BMI, systolic blood pressure, education level, smoking, alcohol use, and *monthly daily steps* (time varying)

The Future

The Scientific Priorities Roadmap for *All of Us*



New Data (v8) Slated for Release Late 2024 to Researchers

Genomic Increases	Data Quantity
Arrays (from 312,900+)	400K+ Arrays <u>(+27%)</u>
srWGS (from 245,350)	400K+ srWGS <u>(+63%)</u>
Long Read Sequences (from 1K)	2.5K+ Long read sequences <u>(+150%)</u>

Data from Mental Health and Well-Being surveys and Exploring the Mind tasks are scheduled for off-cycle release in **January 2025, and data from participant-mediated EHR is scheduled for release in **March 2025**.*

New Data	Data Quantity
Life Functioning Survey	60K+
Surveys from CareEvolution Platform	60K+
Fitbit Data from CareEvolution	10K+
HealthPro Deceased Participant Reporting	4K+
Remote Height & Weight	50K+
Racial and Ethnic Subcategories	620K+
Mental Health and Well-Being Surveys*	110K+
Exploring the Mind Tasks*	35K+
Participant-Mediated EHR*	20K+

All of Us Ancillary Studies: Exploring the Mind with NIMH



**>28.4k unique participants
completed all four tasks**

Exploring the Mind

Help us learn more about behavior and how the brain works.

Available tasks



City or Mountain

React to photographs. (6-10 minutes)

Start



Guess the Emotion

Look at a series of faces. (2-5 minutes)

Start



Now or Later

Choose between example rewards. (5-7 minutes)

Start



Left or Right

Focus on the middle arrow. (5-7 minutes)

Start

Participant Counts

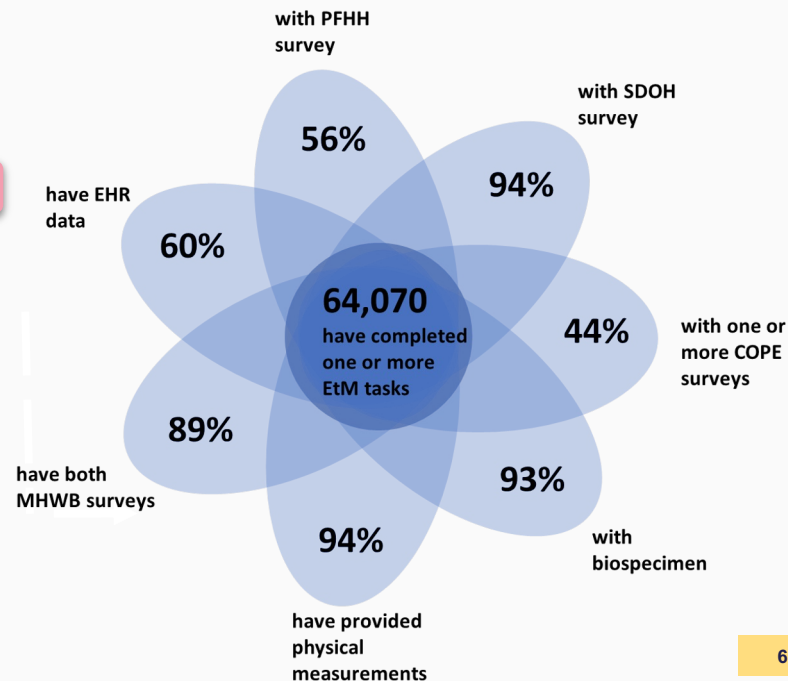
50,220+

59,780+

47,180+

50,020+

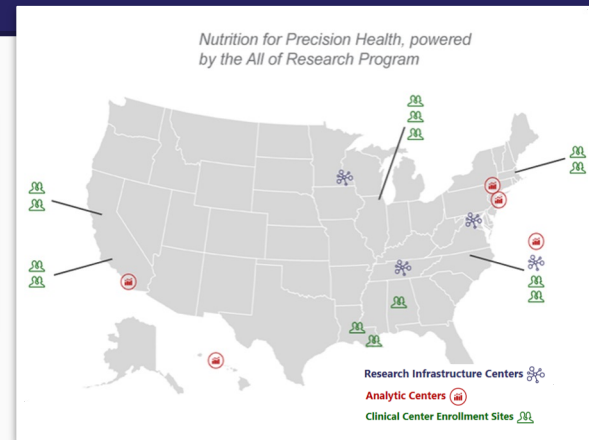
**Overlap with other biomedical data types
available in *All of Us***



All of Us Ancillary Studies: Nutrition for Precision Health

Over 17,000 participants expressed interest in NPH

Researching how nutrition can be tailored to each person's **genes, culture, and environment** to improve health



1



Launched March 2023

Examine baseline diet and physiological responses to meal challenges

2184 consented, 1679 enrolled, 1546 completed

2



Launched October 2023

Examine responses to 3 short-term intervention diets in community dwelling controlled feeding studies

140 consented, 97 enrolled, 22 completed

3



Launched: Fall 2023

Examine responses to 3 short-term intervention diets in domiciled controlled feeding studies

39 consented, 25 enrolled, 6 completed

Why





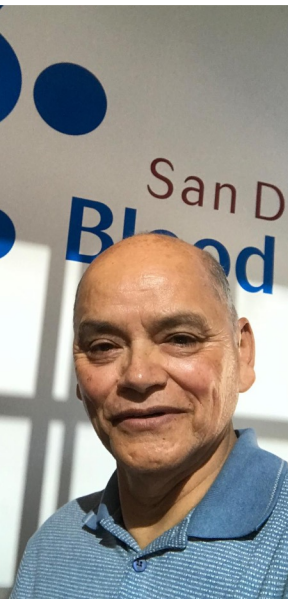




AllofUs.NIH.gov

JoinAllofUs.org

ResearchAllofUs.org



Thank you to our 790K+ participants!

Backup slides

SAS Studio

SAS Studio is now available on the *All of Us* Researcher Workbench. SAS Studio has a point-and-click interface and allows users to perform powerful statistical analyses and data visualizations with *All of Us* data.

Researcher Workbench users do not need to pay for an individual SAS license at this time. The software is provided at no additional cost through the Researcher Workbench. Users will still incur computational costs in the cloud, as with all other analyses.

Visit allof-us.org/SAS or scan the QR code to learn more about using SAS Studio.



RStudio

RStudio is now available on the *All of Us* Researcher Workbench.

RStudio's point-and-click interface and suite of tools, including a variable explorer, an R Markdown editor, a debugger, and an R terminal, make it easier for researchers to use R in the Researcher Workbench.

Visit allof-us.org/RStudio or **scan the QR code** to learn more about using RStudio.



Separate harmonization and collection

Collection
Normalization | is expensive
Storage



4. COLOR OR RACE



Fill one circle.

If "Indian (American)," also give tribe.

If "Other," also give race.

- | | | |
|---|--------------------------------|--|
| <input type="radio"/> White | <input type="radio"/> Japanese | <input type="radio"/> Hawaiian |
| <input type="radio"/> Negro
or Black | <input type="radio"/> Chinese | <input type="radio"/> Korean |
| <input type="radio"/> Indian (Amer.) | <input type="radio"/> Filipino | <input type="radio"/> Other— <i>Print race</i> |
- Print tribe* →

1970 US Census

- _ (ツ) _ / -

LOL



"An extraordinary memoir . . .
It belongs with Atul Gawande's
writings and *When Breath
Becomes Air*." —Adam Grant,
New York Times bestselling
author of *Originals*

CHASING MY CURE

A Doctor's Race to Turn
Hope into Action

A MEMOIR

David Fajgenbaum

Castleman's disease

a rare disorder that involves an overgrowth of cells in
lymph nodes

Abductive reasoning



Cherry angioma

VEGF

Vascular endothelial growth factor



Collection
Normalization | is expensive
~~Storage~~

Preserve provenance

Qualitative collection



Analysis-specific
feature extraction


Reduce collection costs with Chatbots



CL

You

You are an expert health researcher performing an interview to collect the personal health history of a study participant. You should be sympathetic and respectful, without offering any health advice or diagnoses. If you see connections between answers, ask follow up questions that can connect the answers. The interview should start by understanding the topics that are of particular interest to the participant. Ask questions one at a time, and acknowledge the answers that are given. Start from the beginning.



ChatGPT


Understood. Let's begin the interview. I'll start by understanding the topics that are of particular interest to you:

1. Are there any specific health issues or concerns that you feel are particularly important to share or focus on?

CL


You

I have a family history of essential tremors, affecting your father, son, and myself. Notably, my father's tremors worsened with age, impacting his hobbies. My tremors were more prominent when younger but have since reduced. My father had intracranial surgery to implant a deep brain stimulation device, but it was not successful.



ChatGPT

Thank you for sharing that significant aspect of your family health history. It's important to note the familial pattern of essential tremors and the varied response to treatments like deep brain stimulation.



1. Besides the essential tremors, have you or any family members experienced other

Reduce harmonization costs with LLMs

(And direct the harmonization toward the specific analysis)



ChatGPT

Based on the provided doctor's notes for Mrs. Dey, several conditions from the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) could be considered. OMOP CDM provides standardized codes for various medical conditions. Here are some potential OMOP conditions that may be present based on the symptoms described:

1. **Major Depressive Disorder:** The notes indicate signs of severe depression, such as a depressed mood, slowness in physical movement, and slowed speech and thinking.
2. **Dysfluent Speech:** The description of dysfluent, garbled, and loud speech suggests a possible speech disorder.
3. **Anomic Aphasia:** Difficulty in naming objects may be indicative of anomic aphasia, a type of language disorder.
4. **Psychotic Disorder:** The mention of a psychotic or borderline psychotic process, disorganized behavior, and inappropriate anger might align with a psychotic disorder.
5. **Cognitive Impairment:** The presence of disorganized behavior and difficulty in repeating phrases suggests possible cognitive impairment.

It's important to note that assigning OMOP conditions based on a set of symptoms should be done by a healthcare professional who can conduct a comprehensive evaluation. This list is a suggestive starting point based on the symptoms described and would require further clinical assessment and diagnostic processes to confirm.

