Session 2: How Researchers Access and Use All of Us Data

June 13, 2023
Welcome and Introduction

Dr. Karriem Watson
Chief Engagement Officer, All of Us
Logistics

- Links and resources will be posted in the chat
- Please pose questions in the chat
- Please keep yourself on mute - at the end of the program, we’ll address questions in the chat and open up for live questions
- The Questions and Discussion section will be recorded for our notes but not posted online
- Slides and a recording will be posted online afterwards
- Contact the Tribal Engagement team at AOUTribal@nih.gov
Data Governance in the *All of Us* Research Program

13 June 2023
Policy Office Staff
All of Us Research Program
Introductions | Sonya Jooma and Kate Blizinsky

Sonya Jooma, M.A.  
(she/her)  
Deputy Policy Director  
sonya.jooma@nih.gov

Katherine Blizinsky, Ph.D.  
(she/her)  
Policy Director  
katherine.blizinsky@nih.gov
Thank you!

All of Us Participants & Participant Ambassadors

All of Us Program Staff and Partners
- Community Partners
- Consortium Members

All of Us Advisory and Governance Bodies
- ELSI Brain Trust
- Committee on Access, Privacy and Security
- Resource Access Board

All of Us IRB

NIH Office of Science Policy

NIH Office of the General Counsel
Presentation Overview

1. Introduction
   - The History of All of Us
   - All of Us Basics (Review)
   - The Policy Office

2. Considerations
   - Historical Context
   - Important Considerations for Policy-Making
   - Data Governance Principles

3. Data Governance Logistics
   - Database Compilation and Structure
   - Data Access Process
   - Privacy and Security Measures

4. Data Governance Policies
   - Conceptual Model
   - Education
   - Accountability
   - Assistance and Oversight

5. Discussion and Questions
Introduction

- The History of All of Us
- *All of Us* Basics (Review)
- The Policy Office
The All of Us Research Program | From the Beginning

• **2015** | White House announces the Precision Medicine Initiative with the mission to:

> “…enable a new era of medicine through research, technology, and policies that empower patients, researchers, and providers to work together toward development of individualized care.”

- Privacy and Trust Principles
- Security Framework


• **2016a** | Codified and statutorily mandated by the 21st Century Cures Act

• **2016b** | Initial work on the “PMI Cohort Program” begins at NIH, what will become All of Us
The All of Us Research Program
Diversity at Scale

**Diversity of Resources** | Deliver a national resource of deep clinical, environmental, lifestyle, & genetic data from one million or more participants who are consented & engaged to provide data on an ongoing, longitudinal basis

**Diversity of Participants** | Reflect the broad diversity of the U.S.—all ages, races, ethnicities, sexes, genders, SES's, geographies, & health and disability statuses—and over-recruit those historically underrepresented in biomedical research

**Diversity of Researchers** | Build the tools & capabilities that make it accessible to the public and easy for researchers—from citizen and community scientists to scientists from premier university labs—to make discoveries using All of Us
The All of Us Research Program | Comprehensive Research Resources

Participant Surveys
- The Basics
- Lifestyle
- Overall Health
- Personal & Family Medical Hx
- Health Care Access & Utilization
- Social Determinants of Health
- Mental Health and Wellbeing

Electronic Health Records
- Medical Records
- Claims Data
- Pharmacy Data
- Vision and Dental Records

Physical Measurements
- Blood Pressure
- BMI
- Heart Rate
- Height
- Hip Circumference
- Waist Circumference
- Weight

Biospecimens
- Blood
- Saliva
- Urine
- DNA (from blood or saliva)
- RNA

Mobile/Wearable Tech
- COVID Serology
- HbA1c
- Heavy Metals

Assays
- WGS Arrays
- Other Omics

For more on upcoming data types, please visit the Data Road Map at allow-us.org/Roadmap
What does the Policy Office Do?

Policy

noun

pol·i·cy ˈpä-lə-sē
plural policies
often attributive

1  a : prudence or wisdom in the management of affairs
    b : management or procedure based primarily on material interest

2  a : a definite course or method of action selected from among alternatives and in light of given conditions to guide and determine present and future decisions
    b : a high-level overall plan embracing the general goals and acceptable procedures especially of a governing body
Program leadership decides to pursue a program development.

Policy Office determines the boundaries within which development occurs.

Program and consortium staff plan, build, and implement the development.

Research Compliance ensures program activities stay within the boundaries.
<table>
<thead>
<tr>
<th>FUNCTIONAL areas</th>
<th>CONSIDERATION</th>
<th>INTERPRETATION</th>
<th>PROMULGATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluating ethical and social obligations for the program</td>
<td>Setting the terms for legal, regulatory, and policy compliance</td>
<td>Establishing and overseeing program-wide policies and principles</td>
<td></td>
</tr>
<tr>
<td>TOPICAL AREAS</td>
<td>CONSIDERATION</td>
<td>INTERPRETATION</td>
<td>PROMULGATION</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>PROGRAM OPERATIONS</td>
<td>Evaluating ethical and social obligations for the program</td>
<td>Setting the terms for legal, regulatory, and policy compliance</td>
<td>Establishing and overseeing program-wide policies and principles</td>
</tr>
<tr>
<td>AGREEMENTS</td>
<td>How the program itself approaches its mission and objectives</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PARTICIPANT PROTECTIONS</td>
<td>How the program enables participants (or their authorized representatives) to exercise their right to make autonomous, informed choices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RESOURCE GOVERNANCE</td>
<td>What the appropriate measures are to protect participants from foreseeable, preventable harm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GENOMICS</td>
<td>How and under what conditions individuals may access and use research resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>How and under what conditions the program shares genetic and genomic information with participants and researchers</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Considerations

- Historical Context
- Important Considerations for Policy-Making
- Data Governance Principles
Recognizing Disenfranchisement and Historical Abuses
how can we rebuild
TRUST?
Considerations for Meaningful Inclusion

- Work within the legal, regulatory, and ethical framework
- Balance responsibilities: participants’ interests and maximal use of their donation
- Provide mechanisms for engagement across axes of diversity
- Exercise cultural humility
- Respect participant autonomy
- Ensure risks and benefits are distributed equitably
All of Us Data Access Framework | Principles

Data is a non-scarce resource
- Therefore, data should be as accessible as possible
- Accessibility includes researchers outside of academic medical centers and industrial biosciences

Participants are research partners
- Must protect people to the greatest extent possible

Barriers to access must be periodically reappraised
- Aim for minimum necessary barriers to access
- No group of users should have privileged access to the resources
Data Governance | Logistics

- Database Compilation and Structure
- Data Access Process
- Privacy and Security Measures
The *All of Us* Research Program Database

**Ingress**
Collection of data from a wide variety of sources

**Harmonization**
Standardization of data formats to ensure consistency across participant records

**Refinement**
Processing of the data to preserve participant privacy and usability

**Curation**
Filtration of data to make it viewable in the appropriate part of the database
The *All of Us* Research Hub | Data Access Tiers

- **Public Tier**: Summary Statistics Aggregate Counts
- **Registered Tier**: Limited Participant-Level Data
- **Controlled Tier**: Participant-Level Data Additional Data Types (e.g. genomic data)

**Data Browser**

**Researcher Workbench**
The All of Us Research Hub | The Researcher Workbench

**Registered Tier**
*Limited Participant-Level Data*
- Physical Measurements
- Some Electronic Health Record Data
- Survey Data
- Fitbit Data

**Many Data Transformations**
- Date Shifting
- Generalization (e.g. State-Level Geographic Information)
- Ontological “Roll-Ups”

**Controlled Tier**
*Granular Participant-Level Data*
- Registered Tier Data PLUS Additional EHR Fields
- Genomic Data

**Fewer Data Transformations**
- Unshifted Dates (except Year of Birth)
- More Precise Geographic Information (3-Digit Zip Codes)
- Fewer Generalizations and Ontological “Roll-Ups”
The *All of Us* Research Hub | Access Process

1. Learn more about the data available
2. Check for a valid institutional agreement
3. Register as a researcher
4. Verify identity
5. Complete the *All of Us* Responsible Conduct of Research Training
6. Attest to the Data User Code of Conduct
The *All of Us* Researcher Workbench uses a "data passport" model to give registered researchers broad access to the Researcher Workbench rather than granting data access on a project-by-project or question-by-question basis.

- Once they complete the requisite steps, researchers can create workspaces in the tier or tiers to which they have access—no pre-approval required
- When researchers set up their workspaces, they are required to provide public-facing project descriptions on the platform
- Researchers must ensure that their research complies with the program’s data use policies
Partner organizations must meet strict data security standards before they may collect, transfer, or store information from participants.

All participant data is encrypted and obvious identifiers are removed from data used for research.

Researchers must register with the program, complete ethics training, and agree to a code of conduct for responsible data use.

Independent reviewers to check plans and test systems on an ongoing basis to ensure the program has effective security controls in place.
Data Governance | Policies

- Conceptual Model
- Education
- Accountability
- Assistance and Oversight
All of Us Data Access and Use | Conceptual Model

**Education**
Raising Awareness and Setting Expectations in the Responsible Conduct of Research Training

**Accountability**
ID Verification; mandatory compliance with the:
- Data User Code of Conduct
- Data Use and Registration Agreement*
- Data Access and Use Policies

**Preventing Misuse**
**Building Trust**

**Oversight**
Continual monitoring of curated data enclave by Data and Research Center; periodic audits, internally- and publicly-initiated reviews conducted by the Resource Access Board

**Assistance**
User-initiated reviews and assistance from the Resource Access Board available at any stage
Education

Raising Awareness and Setting Expectations in the Responsible Conduct of Research Training
Education | Considerations for Training Development

What's different about All of Us?

- Focus on UBR recruitment
- Diversity of data user backgrounds
- Data users do not interact with research participants

What do we want to achieve?

- Instill in data users a responsibility to All of Us participants
- Contextualize Belmont Principles for research with All of Us data
- Include contemporary conversations on responsible research practices
- Provide tools for data users to put content into practice

How do we accomplish it?

- RCR learning objectives
Education | The RCR Training Modules (Near Future Distribution)
Accountability

ID Verification; mandatory compliance with the...
- Data User Code of Conduct
- Data Use and Registration Agreement*
- Data Access and Use Policies
Data Use and Registration Agreement (DURA)
Institutional contract outlining institutional responsibility for affiliated researcher(s)

Data User Code of Conduct (DUCC)
Individual agreement outlining individual use responsibilities
Accountability | *All of Us* 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 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 Code of 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.
- I will 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.
  - I take full responsibility for any external data, files, or software that I import into the All of Us Researcher Workbench and the consequences thereof.

For more on other data use policies, visit [www.researchallofus.org/data-tools/data-access](http://www.researchallofus.org/data-tools/data-access)
Accountability | Data Use Contracts and Policies (continued)

DURA
Data Use and Registration Agreement
Institutional contract outlining institutional responsibility for affiliated researcher(s)

DUCC
Data User Code of Conduct
Individual agreement outlining individual use responsibilities

Policies
Collection of detailed directives for appropriate use, compliance with which is required by the DUCC

For more on data use policies, visit www.researchallofus.org/data-tools/data-access
Accountability | AI/AN Research Materials Use Policy (Overview)

• Take actions that acknowledge and protect against historical (and contemporary) sources of harm:
  o **MUST** review and agree to abide by the DUCC, Policy on Stigmatizing Research, and Policy on the Ethical Conduct of Research
  o **MUST NOT** attempt to reidentify AI/AN individuals, the region(s) from which they originate, or the tribe(s) to which they belong
  o **MUST** fulfill all technical and logistical requirements for authorized usership, including annual RCR training comprised of the basic modules and any future modules specific to the use of data from AI/AN individuals, as appropriate
  o **MUST** use respectful and appropriate language and preferred terminology, as dictated by the forthcoming guidance

• Comply with all applicable federal, state, and local laws

• Abide by all applicable program policies and procedures for access to data and non-data resources originating from participants who self-identify as AI/AN

• Agree to additional scrutiny of research focusing on AI/AN populations that is conducted using All of Us research resources
All of Us Data Access and Use | Assistance and Oversight

### Continual Monitoring
- Continual monitoring of curated data enclave by Data and Research Center
- Periodic audits, internally- and publicly-initiated reviews conducted by the Resource Access Board

### Assistance
- User-initiated reviews and assistance from the Resource Access Board available at any stage
All of Us Research Hub | Who is using the data now?

5,700+ Registered Researchers

Comparing registered researchers

<table>
<thead>
<tr>
<th></th>
<th>2/2022</th>
<th>6/2023</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Registered Researchers</td>
<td>1,015</td>
<td>5,705</td>
</tr>
<tr>
<td>Researchers from UBW communities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SGM, race/ethnicity, disability status, &amp; career stage</td>
<td>56%</td>
<td>72%</td>
</tr>
<tr>
<td>Researchers from UBW communities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>By race/ethnicity alone</td>
<td>14.3%</td>
<td>28%</td>
</tr>
</tbody>
</table>

5,300+ Active projects

Top conditions being studied
- Cardiovascular disease
- Hypertension
- Mental Health
- Cancer
- Diabetes

See the research currently underway by visiting researchallofus.org/research-projects-directory
Visit the Support Hub to find answers to common questions about becoming a registered researcher, using the All of Us dataset, and more. **NO LOGIN REQUIRED**

**Find instructional materials about the All of Us Researcher Workbench**
- Includes video tutorials, educational resources, release notes, and more

**Search 280+ articles**
- Find information about data dictionaries, the Controlled Tier directory, how WGS and Array data are organized, and more

**Connect with experts during weekly live office hours**
- Explore our calendar of dedicated office hours where researchers can talk about data types and tools with experts

[researchallofus.org/support](researchallofus.org/support)
The All of Us Resource Access Board (RAB)

User-Initiated Workspace Reviews
Reviews proactively requested by Authorized Data Users to facilitate DUCC compliance

Publicly-Initiated Workspace Reviews
Reviews requested by members of the public through the workspace directory

Periodic Workspace Audits
Cadenced reviews conducted on (mostly) randomly selected workspaces

Internally-Initiated Workspace Reviews
Potential compliance issues flagged by All of Us personnel

Anyone can flag research for RAB review by visiting:
researchallofus.org/research-projects-directory
Assistance and Oversight
ELSI Considerations

Issues of…

● **Subjectivity**
  ○ Who should determine what is and is not compliant? What is and is not stigmatizing?
  ○ What happens when people disagree about what is or is not compliant? What is and is not stigmatizing?

● **Intent vs. Application**
  ○ What should be the limitations on data use where there is good intent but high potential for the downstream misapplication or misuse of findings?
  ○ Can/how do we control downstream use?
  ○ Are there topics that should always be off limits?

● **Protections vs. Rights**
  ○ When should the potential benefit (or harm) to an individual, group, or community outweigh researchers’ rights and freedoms?
Assistance and Oversight | RAB Review Considerations

1. **Occurrence**
   - Has a violation actually occurred?

2. **Type**
   - What is the type of the perceived violation?
     - Legal or Regulatory Violation
     - Re-identification
     - Stigmatizing Research
     - Individual or Group Harm
     - Insufficient Workspace Description
     - Other

3. **Intentionality**
   - Is it reasonable to believe that the perceived violation was intentional?

4. **Scale and Scope**
   - How many individuals or how much information was affected?
   - How significant was the effect?

5. **User History**
   - Has the user always otherwise been in good standing with *All of Us*?

6. **Patterns**
   - Are there patterns in violations in the user’s history or across users (either total or affiliated with a particular institution)?
134 reviews

- Most RAB-identified problems were with clarity and/or detail of Workspace Description or data obfuscation measures.

Among RAB members significant agreement about...
- The potential for research to stigmatize individuals or groups.
- Which workspaces were adequately described.

Researchers have overwhelmingly been receptive and willing to make change(s).
New Challenges Ahead!

• Recent rollout of the Controlled Tier, with more granular and more sensitive data types
• Growing and diversifying user base
• Addition of new data and other research resources, like biospecimens and direct access to participants
Data Access, Use, and Tribal Engagement
Diverse Perspectives in *All of Us* Program Governance

**AI/AN Governance Representation:**

- Research Program Advisory Panel
- Institutional Review Board
- Steering Committee
- Participant Ambassadors
- Science Committee
- Resource Access Board
- Ethical, Legal, and Social Implications (ELSI) Brain Trust
- Biospecimen Access Policy Task Force (past)

**Suggested Opportunities for Input:**

- Augment the expertise available to the Resource Access Board with Indigenous knowledge for data or research projects focusing on AI/AN peoples
- The program strongly encourages AI/AN researcher participation and will aim to collaborate with Tribal leaders to hold presentations on how to use the *All of Us* Research Hub at tribal events, AI/AN-focused scientific conferences and workshops, and webinars.
Researcher Spotlights on ResearchAllofUs.org

Like this one from Younga “Heather” Lee, PhD, Postdoctoral Research Fellow at Massachusetts General and Harvard/MIT

Meet the Researchers
Studying the link between discrimination and depressive symptoms
August 23, 2022

Meet Dr. Heather Lee, a researcher at Massachusetts General Hospital exploring the mental health impact of everyday discrimination during the COVID-19 pandemic.

https://allof-us.org/HeatherLee
Stay in touch to learn more

Sign up for our bimonthly newsletter

Subscribe to Research Roundup

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

allof-us.org/RRSignup
Create an *All of Us* account

Register to be an *All of Us* Researcher

1. **STEP 1**
   Learn more about the data available

2. **STEP 2**
   Check for your institution's agreement

3. **STEP 3**
   Register as a researcher

4. **STEP 4**
   Verify your identity

5. **STEP 5**
   Complete all of Us responsible conduct of research training

6. **STEP 6**
   Sign data user code of conduct

[alof-us.org/Register]
**Purpose of the ROA**

### Area of Interest 1: Participant Engagement and Enrollment

**Task 1:** Participant outreach and engagement

**Task 2:** Enrollment and retention

**Task 3:** Pediatric outreach and engagement

**Task 4:** Computer access

**Task 5:** Healthcare provider outreach and engagement

### Area of Interest 2: Researcher Engagement

**Task 1:** Researcher Hub awareness and researcher engagement

**Task 2:** Building research capacity

**Task 3:** Promote team science

### Area of Interest 3: Tribal Nations, American Indian and Alaska Native populations, and Indigenous Communities

**Task 1:** Outreach and engagement

**Task 2:** Training and career development

**Task 3:** Education and culturally appropriate resource development

**Task 4:** Research infrastructure and capacity building

### Area of Interest 4: Participant as Partners

**Task 1:** Oversight of consortium participant partner related activities

**Task 2:** Full integration of participants and advocates

**Task 3:** Facilitation of ongoing input from participant representatives

**Task 4:** Assess the impact of participant engagement
Area of Interest 3: Engagement with Tribal Nations, American Indian and Alaska Native Populations, and Indigenous Communities

The All of Us Research Program initiated **tribal consultation in 2019** to engage Tribal Nations about the inclusion of AI/AN populations in this research program. The information received during consultation has been used to guide our work with AI/AN communities and Tribes.

**Technical Objectives**

- Outreach and engagement to AI/AN and Indigenous communities
- Training and career development in research as it relates to precision medicine
- Education and culturally appropriate resource development for AI/AN populations and partners
- Research infrastructure and capacity building within AI/AN and Indigenous communities ensuring sustainability of precision medicine research

*Review SAM.gov for full submission requirements*
Informational Presentation and Discussion Schedule

Session 1: *All of Us* Research Program Overview and Tribal Engagement  
Tuesday, June 6, 2023, 1:00 - 3:00 pm ET

Session 2: How Researchers Access and Use *All of Us* Data  
Tuesday, June 13, 2023, 1:00 - 3:00 pm ET

Session 3: Describing Self-Identified AI/AN Participants in *All of Us* Data  
Tuesday, June 20, 2023, 1:00 - 3:00 pm ET

Session 4: Data and Partnerships in *All of Us*  
Tuesday, June 27, 2023, 1:00 - 3:00 pm ET

*All of Us* Tribal Consultation  
Thursday, September 28, 2023, 1:00 - 3:00 pm ET
Logistics

- Links and resources will be posted in the chat
- Please pose questions in the chat
- Please keep yourself on mute - at the end of the program, we’ll address questions in the chat and open up for live questions
- The Questions and Discussion section will be recorded for our notes but not posted online
- Slides and a recording will be posted online afterwards
- Contact the Tribal Engagement team at AOUTribal@nih.gov