

All of Us Research Program Institutional Onramp Event Summary Report October 3, 2022

Welcome and Keynote

*Karriem Watson, D.H.Sc., M.S., M.P.H., Chief Engagement Officer, All of Us Research Program
Joshua Denny, M.D., M.S., Chief Executive Officer, All of Us Research Program*

Welcome: The All of Us Research Program

Dr. Watson welcomed the attendees and presented the goals of the day's program:

- Share the mission and objectives of the *All of Us* Research Program
- Highlight how institutions can leverage the *All of Us* Research Program to enhance research capacity
- Share resources available to institutions and researchers
- Hear from the attendees

Dr. Watson shared the challenges he faced as an early-stage investigator conducting cancer disparities research in accessing secondary datasets that reflected the diversity of the populations he was studying. He is excited about the intentionality of the *All of Us* Research Program in creating a diverse database.

The *All of Us* Research Program is part of the NIH Precision Medicine Initiative. The *All of Us* database is one of the largest, richest, and most diverse biomedical datasets of its kind. Some of the program's highlights are the following:

- Its goal is to invite one million or more people to enroll in the program over a 10-year period.
- Data are currently available from more than 372,000 participants.
- Of these participants, 80% are from groups typically underrepresented in biomedical research, and 45% are from racial and ethnic minority groups.
- Other groups underrepresented in biomedical research include sexual and/or gender minorities, those with low educational attainment, those in rural areas, and those with disabilities.

The *All of Us* Research Program is also intentional about:

- Creating a demographically diverse researcher cohort that promotes responsible and ethical use of data, returns value to participant communities, and accelerates research impact.
- Encouraging students and early-stage investigators to bring fresh, creative perspectives and innovative research outcomes.
- Ensuring access for researchers from various institutions and organizations to establish a truly equitable resource for all.
- Reaching institutions and researchers who have historically been underrepresented in large genomic studies.

As of October 2022, there are more than 3,000 registered researchers nationwide across 400 institutions and across a range of institutional roles and career stages. They represent more than 260 academic institutions as well as more than 60 nonprofits and 50 health care organizations. There are more than 2,400 active research projects that cover conditions that include cardiovascular disease, hypertension, mental health, cancer, and diabetes.

The Researcher Workbench offers many benefits for both researchers and institutions.

Registered users have access to:

- The research platform, data, and tools
- \$300 in initial compute credits to apply to their research project
- User support materials and staff
- Opportunities to showcase work, including publishing team science, dissertations, poster presentations, class projects, and community/citizen science

Support for institutions includes *All of Us*–related funding, including the following:

- [Notice of Special Interest: Administrative Supplements to Advance Precision Medicine Using the *All of Us* Research Program’s Data](#)
- [Notice of Special Interest \(NOSI\): Administrative Supplements to Support “All of US” and Health Disparities-Related Pilot Research Projects at NIMHD-Funded Research Centers in Minority Institutions \(RCMI\)](#)
- [Educational Hub for Enhancing Diversity in Computational Genomics and Data Science \(U24 Clinical Trials Not Allowed\)](#)

Dr. Watson introduced Dr. Denny, the keynote speaker. Dr. Denny is the Chief Executive Officer of the *All of Us* Research Program. He has been involved in the program since its inception, first as a member of the advisory committee to the NIH director of the Precision Medicine Initiative working group, which developed the initial *All of Us* scientific blueprint. He led the program’s initial prototyping project and served as the principal investigator (PI) for the *All of Us* Data and Research team. Before joining NIH, Dr. Denny was a professor of biomedical informatics and medicine, director of the Center for Precision Medicine, and vice president for personalized medicine at Vanderbilt University Medical Center (VUMC). He was a leader in the development of genome-wide association studies (GWAS) and phenotype risk scores, and he was instrumental in implementing genomics in general practice using the Implementing Genomics in Practice (IGNITE) Network. Most importantly, he is a staunch advocate of ensuring that the *All of Us* database is a huge resource to academic partners, students, and other stakeholders.

Keynote: The *All of Us* Research Program Mission

Dr. Denny thanked Dr. Watson and opened with a statement of the *All of Us* Research Program’s mission: to accelerate health research and medical breakthroughs, enabling individualized prevention, treatment, and care for all of us. The goal is to be inclusive and to recruit participants and researchers from diverse backgrounds.

There are three foundational components of this program:

- Nurturing partnerships for decades with at least one million participants who reflect the diversity of the U.S.
- Delivering one of the largest and richest biomedical datasets that is broadly available and secure
- Catalyzing an ecosystem of communities, researchers, and funders who make *All of Us* an indispensable part of health research

As of September 30, 2022, the *All of Us* Research Program has enrolled 535,000 participants, including 374,000 who have completed the initial steps of the program. *All of Us* has access to 324,000 electronic health records (EHRs) and 394,000 biosamples. More than 80% of the participants are underrepresented in biomedical research, including almost 50% who identify as members of racial or ethnic minority groups. Of those, more than 19%

are Black, African American, or African and more than 16% are Hispanic, Latino, or Spanish.

Participants can enroll in the *All of Us* Research Program through a variety of mechanisms, but all involve a common consent mechanism. The data used by the program come from:

- **EHRs.** Participants provide permission to actively share these records using the Participant Portal linkage, or records can be donated on their behalf.
- **Participant surveys.** New participants complete a [Basics Survey](#). About one new survey is released per year.
- **Physical measurements.** A minimal set of these measurements that includes height and weight for the calculation of body mass index, waist and hip circumference, blood pressure, and pulse is collected.
- **Biosamples.** These include blood samples and others. Biosamples have been used to generate sequencing data from DNA but will be used in other ways in the future.
- **Mobile and wearable technologies.** Data from Fitbit devices, can be linked to participants' other data.

Data come into the Raw Data Repository (RDR) from many different EHR systems, clinics, and other varied sources. Raw data are first harmonized. Privacy-preserving technologies are then applied; these are quality-control and quality-assurance mechanisms to ensure that all obvious identifiers are removed to lessen the risk of re-identifying participants.

After data curation, data are organized by tier, representing the degree of [access to the data](#) by level of security.

- **Public Tier.** These data contain broad summary statistics and aggregate counts and are accessible through the Open Access public [Data Browser](#). No login is required.
- **Registered Tier.** These data include EHRs, physical measurements, and shifted dates (a privacy measure). The privacy settings exceed the HIPAA Safe Harbor Standards.
- **Controlled Tier.** These data are highly secure and contain genomics and real dates. They contain no personally identifiable information (PII). Access to these data is restricted and limited to researchers with additional training.

Both the Controlled and Registered Tiers are currently accessible through the [Researcher Workbench](#). Access to data in the Controlled and Registered Tiers is only available to researchers with an institutional agreement in place. Any U.S.-based academic, nonprofit, or health care institution can enter into a Data Use and Registration Agreement (DURA). A fourth tier, Individual Biospecimen and Participant Data, will be available in the future for ancillary and follow-up studies. The first ancillary study, [Nutrition for Precision Health](#), has been launched within this tier.

The Open Access public [Data Browser](#) contains summary statistics on:

- EHR data, including health conditions, drug exposures, and others
- Genomic variants
- Survey responses, including those from COVID-19 surveys
- Physical measurements

Data on the Researcher Workbench is diverse and longitudinal and may, at times, contain 40 years of data through the use of EHRs. As an example, 20,000 adult *All of Us* participants have EHRs dating back to their birth.

As of October 2022, the Researcher Workbench contained:

- Data on more than 374,000 participants
- Physical measurements from more than 311,00 participants
- More than 324,000 EHRs
- More than 372,000 survey responses
- More than 12,800 Fitbit records

Some of the highlights of access (row-level data, Registered Tier) to the Researcher Workbench include the following:

- It is a cloud-based central resource.
- It is a data passport access model. No separate IRB approval is needed. Once institutions go through the DURA process, researchers can create a workspace and begin work.
 - Project descriptions are posted publicly.
 - Participants can flag stigmatizing research.
 - The Researcher Workbench is subject to audits.

Dr. Denny stressed that genomic studies have been enhanced through inclusion.

- In March 2022, 100,000 sets of whole genomes and 165,000 array data were launched within the Controlled Tier.
- 600 million unique variants were observed, including 400 million variants that were shared by other participants, demonstrating the power of genomic data for understanding the phenotypic importance of these shared variants.
- That March data release included the first large genomic dataset with more than 50% of participants who identify with racial and ethnic minority communities.
 - 22% of the genomic dataset was contributed by Black, African American, or African participants.
 - 18% of participants with genomic data identified as Hispanic, Latino, or Spanish.
- In contrast, 96% of GWAS worldwide have been done in populations of European ancestry.

Dr. Denny emphasized that these genomic data can be paired with rich phenotypic data from EHRs, surveys, Fitbit devices, and more. In addition, the *All of Us* dataset includes a wide range of diseases—such as heart disease, type 2 diabetes, and depressive disorder—from many study domains. This enables researchers to study many different populations and the outcomes of health disparities.

Dr. Denny demonstrated the power of a genome association analysis of type 2 diabetes, using *All of Us* data, that can be completed in 20 minutes.

- A study population was first created with cohort builder, which also defined the cases and controls.
- Dataset variables were then selected, and a genomics analysis tool was selected from the following options: Hail, PLYNK, [R](#), [Python](#), or [Jupyter Notebooks](#).
- In this example, about 5,000 cases with type 2 diabetes were identified, and 23 million variants were analyzed. Associations of variants with particular phenotypes were quickly identified.
- This analysis cost \$37, which is covered by the \$300 of free compute credits offered to new researchers.

Dr. Denny compared this ease of this analysis to one that requires a computing cluster; this GWAS could be undertaken on a laptop at a neighborhood coffeeshop.

Dr. Denny gave examples of topics that can easily be studied and are supported by *All of Us*:

- Health disparities
- Historically underrepresented populations
- Associations between diseases, medications, behaviors, and mental health
- Associations with genomics and pharmacogenomics
- Health outcomes
- Environmental data linked to a three-digit ZIP code

There are six steps involved in gaining access to row-level (Registered and Controlled Tier) data:

- Learning more about the data available through the [Data Browser](#), [data access tiers](#), [Data Snapshots](#), and [Researcher Workbench](#) tools
- Checking to see whether the institution has a [DURA](#) in place
- Registering for a [Researcher Workbench](#) account
- Verifying researcher identification using [Login.gov](#)
- Completing *All of Us* research training
- Signing the [Data User Code of Conduct](#) (DUCC)

Dr. Denny said that on average it takes about a month to set up a new DURA. After that, researchers can register and complete their training within a few hours.

Dr. Denny ended the keynote by emphasizing that the goals of researcher inclusivity are to lower barriers to research and to facilitate access for researchers from diverse institutions and organizations, to establish a truly equitable resource for all.

Panel Discussion: How Institutions Can Benefit from Accessing the *All of Us* Database

Moderator: *Karriem Watson, D.H.Sc., M.S., M.P.H., Chief Engagement Officer, All of Us Research Program*

Panelists: **Baylor College of Medicine:**

Dr. Debra Murray, Office of Community Engagement and Diversity

Dr. Laura Rosales, Administrator, Department of Molecular and Human Genetics

Dr. Kim Worley, Vice Chair for Basic and Translational Research

Panelists: **Xavier University of Louisiana:**

Dr. KiTani Lemieux, Associate Professor, College of Pharmacy

Ms. Kaneisha Akinpelumi, Associate Vice President, Office of Research and Sponsor Programs

Introductions from the Panelists: How Is Their Institution Affiliated With *All of Us*?

Dr. Watson asked each of the panelists to briefly introduce themselves and describe how their institution is affiliated with *All of Us* in advancing the work of their students and faculty, as well as the research community in general.

Dr. Murray explained that her involvement with *All of Us* began with a community engagement grant that was awarded to the Office of Community Engagement to increase the number of underrepresented scientists on the Workbench. Her office works with minority-serving institutions (MSIs) and small universities. The office brings researchers to the Workbench primarily through faculty summits, the first of which was held in May 2022, but they also work within their department to introduce the Workbench.

Dr. Worley said that she is leading the data science training program for *All of Us*, both for the faculty summit participants and for the support organizations. She is encouraging local researchers to engage with the Workbench and to use the data. She noted that there are many researchers working on genetic and genomic applications who could benefit from the cohort information that is available in the *All of Us* platform. Dr. Worley added that, in addition to helping with the DURA needed for access to the *All of Us* platform, her office has also worked out institutional agreements to enable researchers to bring their data to the platform using Google Cloud and Burwood, a cloud provider service organization.

Dr. Rosales said that she is a co-investigator on an *All of Us* genetics research grant. She was formerly a director of sponsored programs. At today's meeting, she would like to share tips on working with institutions—particularly with sponsored programs offices—to get a DURA. Although the research office is usually the office that initiates the DURA, the sponsored programs office is the primary contact, so working with this office is the best way to get going.

Dr. Lemieux said that the *All of Us* platform better enables Xavier University to carry out its mission as a historically Black college and university (HBCU): to train students to create a more humane society and to serve the medically underserved. To this end, the College of Pharmacy does a lot of community engagement through health fairs and with the Center for Minority Health and Health Disparities Research and Education. The faculty use didactic instruction and an interdisciplinary approach to teach pharmacogenetics and social determinants of health across the curriculum.

In practical terms, Dr. Lemieux said that it is important to equip researchers with meaningful tools and training, including the ability to use R and Python, which are essential for genomics analyses. Gaining access to the Workbench is not enough. She agreed that working with the sponsored research office was critical to getting institutional approval for Xavier University.

Ms. Akinpelumi said that Xavier University is small, but it has many active research projects with a focus on health disparities. She said that gaining access to the *All of Us* dataset was more

challenging than the administrative processes. She offered a few tips to prospective researchers:

- Meet with PIs at the university or institution to discuss the anticipated uses for the data and the processes for getting access to the *All of Us* dataset.
- Determine any unique requirements that the university might have, including training and IT protocol familiarity.
- Get buy-in early on; have an elevator speech prepared that emphasizes the importance of the *All of Us* Research Program and its benefits to the larger researcher community on campus.

R and Python Genomics Analysis Tools Training

Dr. Watson asked Dr. Murray to describe the training involved in using the genomics analysis tools, R and Python.

Dr. Murray said that she became concerned when she became aware of the small number of early-stage scientists who progressed in their research. She addressed this issue by developing a bioinformatics bootcamp. This training is provided to early-career faculty and senior postdocs through the annual Underrepresented in Biomedical Research Faculty Summit (UBR Faculty Summit) and provides specific training in [R and Python](#). Dr. Worley developed a curriculum for the May summit, focusing on the initial steps of engagement with the platform rather than the actual R and Python training that is more widely available. She explained that using the platform allows access to example codes from [example codes from Structured Query Language \(SQL\)](#), and researchers learn how to do the coding to modify those queries. There is a focus on curating the available training resources to quickly point people to the most approachable datasets.

Partnering with HBCUs

Dr. Watson asked Dr. Lemieux to describe the process of partnering with other HBCUs.

Dr. Lemieux said that funding from Pyxis allowed Xavier University to hold Saturday morning training sessions and an intensive bootcamp. Xavier University invited local HBCUs with DURAs as well as faculty on the Researcher Workbench to participate. Undergraduates and pharmacy students participate in these training sessions. Research mentors are part of the *All of Us* cohort, giving students access to the Workbench. Dr. Lemieux said that the importance of this training, in addition to providing familiarity with [Python](#), is that Xavier University serves as a touch point for maintaining diversity on the Workbench.

Administrative Supplements to Support *All of Us* and Health Disparities Research

Dr. Watson asked Dr. Lemieux to describe an example of an administrative supplement and its importance in advancing health disparities research.

Dr. Lemieux explained that a colleague at an [RCMI](#) had been awarded an administrative supplement for Black Bodies Matter to research health disparities related to of amputations among Black people. Generally, wound-healing measures are not offered for serious limb injuries in Black patients, and amputations are promoted. Her colleague's research, using data from *All of Us*, has supported that hypothesis.

Using the *All of Us* Dataset as a Teaching Tool

Dr. Watson asked the panelists whether they had used or are thinking of using the *All of Us* dataset as a teaching tool.

Dr. Worley said that one of Baylor College of Medicine's student groups is developing teaching materials using the *All of Us* platform, and there are a number of sign-ups on the platform using the Researcher Workbench. She also described the [NIH Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability \(STRIDES\) Initiative](#). She thinks that the most approachable data is the public data for public health information.

Dr. Murray expanded on her description of the research team that is interested in developing a curriculum, which will be a part of the master's in public health program. The common curriculum will be developed, then specific plans will be devised for different groups.

Dr. Watson said that one use of the *All of Us* Research Program would be as a part of team science. For example, an online program with a weak genomics program might partner with a computer science division; there may be an opportunity to pair a bioinformatics researcher with a health disparities researcher. *All of Us* is glad to partner with the University of Utah to talk about engagement with high school students, resulting in the first high school with a DURA. The high school teacher involved will be using the *All of Us* dataset in the classroom this fall.

During the COVID-19 pandemic, Ms. Akinpelumi initiated an ongoing process that involved working cooperatively across the Xavier University campus. Dr. Lemieux added that the Xavier University president, a scientist, is very supportive of forward-thinking science. That attitude has been helpful in promoting synergistic opportunities across the campus.

Dr. Murray said that Baylor College of Medicine offers an external seed award that would offer some funds to researchers who are using *All of Us* data.

Dr. Rosales noted that faculty on university campuses are often unaware that their institution holds a DURA, so it is important for *All of Us* to promote the resources available.

Questions in Chat

- Several people asked how *All of Us* engages American Indian/Alaska Native (AI/AN) communities. Dr. Watson explained that there is a Tribal engagement workstream, and *All of Us* has been very intentional about engaging with AI/AN communities and other Indigenous populations. For example, *All of Us* has just onboarded a Federally Qualified Health Center in Hawaii that has been engaging Indigenous populations for decades. In terms of access and data, the program had consultations with tribal leaders to discuss ways to safeguard AI/AN data.
- Dr. Denny added that the program recently engaged in its first large-scale Tribal consultation. Information is available in the [Final Tribal Consultation Report](#). *All of Us* arrived at some guidance, including not recruiting on Tribal land. AI/AN populations may consent to participate in the program, but those data have not been released. AI/AN participants self-identify as such. Dr. Denny said that *All of Us* looks forward to further engagements and conversations. In addition, he is excited about partnering with Tribal colleges, universities, and researchers who can identify important questions. He added that there is an open [research opportunity announcement \(ROA\)](#) for engagement partnerships, and *All of Us* is especially interested in using this award to advance work with AI/AN communities.
- Several people had questions about the tools recommended for [R and Python](#). Dr. Watson said that there is a researcher engagement work stream and a two-page instruction guide. The *All of Us* Data and Research Center hosts regular virtual [office hours](#).
- An attendee asked how the *All of Us* Research Program is supporting frontline staff at the various recruitment sites to use the *All of Us* Researcher Workbench. Dr. Watson described one scenario: A research coordinator might approach the PI with a particular research question in mind. He said that the researcher engagement team is beginning to think about working with engagement leads across sites to promote workforce development.

Data Use and Registration Agreements (DURA): An Overview

Cassie Springer, M.P.H., Vanderbilt University Medical Center (VUMC)/Data and Research Center (DRC)

Ms. Springer works at VUMC with the *All of Us* access team in the data and research team. She assists researchers from direct request to the completion of the [DURA](#) within their institutions.

There are six steps involved in registering to become an *All of Us* researcher and gaining access to the *All of Us* Researcher Workbench [Controlled and Registered Tier](#) data:

- Learning more about the data available through the [Data Browser](#), [data access tiers](#), [Data Snapshots](#), and [Researcher Workbench](#) tools
- Checking to see if the institution has a [DURA](#) in place
- Registering for a [Researcher Workbench](#) account
- Verifying researcher identification using [Login.gov](#)
- Completing *All of Us* research training
- Signing the [Data User Code of Conduct](#)

Three portions of the *All of Us* institutional contracting process should be discussed:

- The timeline
- The responsibilities of the signing institution to remain compliant
- The typical stumbling blocks to access

Access to the *All of Us* Researcher Workbench is available to researchers with institutional signoff of the DURA or with an individual agreement for institutions that prefer a different method. Any U.S.-based academic, nonprofit, or health care institution can gain access to the Workbench. Government entities, such as other NIH institutes and centers, are also eligible. As of October 2022, 411 institutions had agreements in place.

- The DURA process begins with the requester submitting an [Access Request Form for their institution that](#) includes the name of an institutional contracting contact who can execute the agreement. A member of the VUMC access team will reach out to the contracting contact to complete the process.
- Each step in the agreement timeline can take one to two business days.
- After the executed agreement is received by the VUMC team, it typically takes just two days for researchers to be able to get started in the Researcher Workbench.
- Through the DURA, the institution accepts responsibility for the actions of data users from the institution, including compliance with the program roles outlined in the DUCC.
- A typical obstacle to the DURA process is a lack of responsiveness from the contracting contact.
- After the researcher has received notice that their agreement has been signed on their behalf, they may begin the process of creating an account and registering. Obstacles at this stage include issues with Login.gov or the ID verification step. If any issues occur, the researcher should reach out to the support team at support@researchallofus.org.

Frequently asked questions:

- Can my institution specify different email domains for the Registered and Controlled Tiers?
 - Yes. However, access is not permitted with Gmail or Yahoo email domains.
- How are researcher counts monitored?
 - There is a monthly user reporting system that enables a team to see who has data access within the Registered and Controlled Tiers under an institutional agreement.

- Can students with an institutional email gain access?
 - Yes.

Institutional Onramp Event Breakout Room 1 Meeting Summary October 3, 2022

Breakout Room 1: The DURA Researcher Request Process and Troubleshooting

Moderators: *Cassie Springer, M.P.H., and Ryan Hale, M.A.*

The breakout session opened at 1:36 p.m. ET.

Mr. Hale said that the agenda for the breakout session was first to review the researcher request process. He would then demonstrate the [Researcher Workbench](#) request form, followed by describing the new “Access and DURA Support” tile in the [User Support Hub](#), and demonstrating of the [Login.gov](#) bypass process.

Researcher DURA Request Process Overview

Mr. Hale reviewed the steps involved in this process:

- Researchers can check the *All of Us* Research Hub to determine whether their institution has already completed a [DURA](#).
- If an agreement has not been completed for their institution, the researcher should fill out an [Access Request Form to start the DURA process](#).
- After the DURA agreement is signed, the researcher can [register for the Researcher Workbench](#) using their approved institutional email.
- To access the dataset, the researcher must complete all required registration steps, including:
 - Verifying their identity via [Login.gov](#)
 - Completing the [Responsible Conduct of Research](#) training
 - Signing the [DUCC](#)

Demonstration of the Researcher Request Form

Mr. Hale demonstrated some of the features on the Researcher Request Form.

- [Institutions with agreements](#) in place are listed, along with their access level. Researchers can use the search bar to find their institution among the 411 that are currently registered.
- If the institution does not have an agreement that includes access to data or requires individual agreements, the researcher can submit a request.
- The institutional access team will reach out to the contracting contact.

Access Tile Summary

- The [User Support Hub](#) is now open for public access. Previously, the User Support Hub was available only to registered researchers.
- The “[Access & DURA Support](#)” tile in the User Support Hub offers information about how to gain institutional access.
- Helpful resources in the “Access & DURA Support” tile include the following:
 - FAQs about accessing the Workbench for [researchers](#) and [institutional contacts](#)
 - [Instructions](#) for navigating the Login.gov verification process
 - Links to other resources located in the *All of Us* Research Hub

Login.gov ID Verification

- As part of the Researcher Workbench registration process, users are required to verify their identity through [Login.gov](#).
- Researchers will be prompted to sign in to their existing Login.gov account or create a new Login.gov account during the registration process.
- Materials needed to verify the user's identity through Login.gov include the following:
 - An institutional email address
 - A state-issued ID
 - A Social Security number
 - A valid phone number associated with the user's name
- Some users have not been able to complete Login.gov ID verification despite having provided all required materials. If this happens, *All of Us* provides a temporary bypass to the Login.gov ID requirement.

Login.gov Bypass Overview

Criteria needed to qualify for the temporary Login.gov bypass include the following:

- Affiliation with an institution with a signed DURA on record
 - A Researcher Workbench account
 - A Social Security number, a state-issued ID and cell phone account with a major cell phone service provider, or a U.S. passport or e-passport
 - An eRA Commons account or confirmed approval from the institutional contact
- Mr. Hale encouraged attendees to direct queries on any issues to aoudurasupport@vumc.org.

Closing

Attendees were asked to complete the [post-event evaluation online](#).

Adjournment

The breakout session was adjourned at 2:01 p.m. ET.

Institutional Onramp Event Breakout Room 2 Meeting Summary October 3, 2022

Breakout Room 2: Leveraging the *All of Us* Dataset to Enhance Research Capacity at Your Institution

Moderators: *Rubin Baskir, Ph.D., and Danielle Wilfong, M.S.*

The breakout room opened at 1:36 p.m. ET.

Dr. Baskir clarified the distinction between the three breakout rooms:

- Breakout Room 1: The DURA Researcher Request Process and Troubleshooting is geared toward individuals who are interested in gaining access to the *All of Us* dataset or who need assistance navigating the DURA process.
- Breakout Room 2: Leveraging the *All of Us* Dataset to Enhance Research Capacity at Your Institution is geared toward individuals who are interested in resources available to researchers through *All of Us*, including support, training, and funding opportunities.
- Breakout Room 3: Utilizing the Researcher Workbench is geared toward individuals who are interested in a more in-depth look at the *All of Us* dataset and how to set up a workspace and navigate the dataset.

Dr. Baskir said that participants could switch between breakout rooms if they wished. Although breakout sessions would not be recorded, written summaries of the sessions would be made available to attendees, along with a summary of the meeting as a whole.

Ms. Wilfong said that *All of Us* conducts informational presentations regularly. If a group of individuals at an institution is interested in scheduling an information session, they should contact the program at support@researchallofus.org. Attendees should also direct any questions to this email address after the call.

Benefits of Registration to Researchers

Ms. Wilfong reviewed the resources provided to registered users:

- A platform and tools, including a cohort-building point-and-click tool to create inclusion and exclusion criteria for a research cohort
- \$300 in initial compute credits to cover costs associated with Google storage and processing researcher data
- User-supported materials and staff to support researchers
- Opportunities to showcase work, such as [Spotlights](#)

Ms. Wilfong said that the *All of Us* dataset can support many types of inquiries and research projects, including team science, dissertations, theses, poster presentations, class projects, and practicing data science methods.

Featured workspaces can help new researchers become familiar with the *All of Us* platform. By duplicating these workspaces, copying and pasting codes, and editing codes, researchers can learn how to use the Researcher Workbench.

Many different kinds of support are available for researchers. Researchers can contact the Support Desk at support@researchallofus.org for assistance. They will receive a response within one to two business days. The DRC also offers monthly [virtual new user orientations](#). In addition, the DRC holds twice weekly office hours that provide live virtual help with questions .

For more information on the support offered, Ms. Wilfong directed attendees to visit the [Support Hub](#). Anyone can access the site, which provides a calendar of office hours, training, and informational events.

Dr. Baskir said that he is a member of the *All of Us* Division of Engagement and Outreach. In addition to supporting researchers in general, the program is dedicated to supporting researchers who are underrepresented in the biomedical workforce with specific funding opportunities. These efforts include the following:

- *All of Us* invested in two administrative supplement funding awards to support research centers in minority institutions.
- [The All of Us Researcher Academy](#) (administered by a partner, RTI International) issued grants to six HBCUs.
- *All of Us* is partnering with the Delta Research and Educational Foundation (DREF) to award funding for capacity building at 10 HBCUs.
- *All of Us* has partnered with multiple National Community Engagement Partners around researcher engagement.

All of Us is committed to both engaging a diverse research community by using the Researcher Workbench and ensuring data access to researchers at all levels. Efforts include the following:

- *All of Us* has partnered with the University of Utah to develop high school curricula by using the *All of Us* dataset
- The program is engaging, mentoring, and supporting nearly 30 students through the Research Scholars Program.
- *All of Us* is working with Baylor College of Medicine to build and support multidisciplinary research teams. Baylor's 2022 UBR Faculty Summit was attended by 34 scholars from 26 institutions, leading to the creation of 34 accounts on the Researcher Workbench.
- The program is facilitating research opportunities for community and citizen scientists with the National Library of Medicine.
- The National Institute on Minority Health and Health Disparities and *All of Us* have made multiple NOSI awards that are funding research at the Ponce Health Sciences University, Xavier University of Louisiana, and Howard University.

Dr. Baskir described the following open and upcoming funding opportunities:

- [UBR Faculty Summit](#) seed awards to support multidisciplinary research teams
- Baylor's external pilot awards to establish diverse college and university research teams
- A [research opportunity](#) to support engagement, communications, and enrollment partnerships with *All of Us*
- A [funding opportunity](#) for new investigators to promote workforce diversity in genomics, bioinformatics, or bioengineering and biomedical imaging research

Dr. Baskir said that more information about these funding opportunities can be accessed through the QR codes on the slides and in the funding announcements.

He encouraged researchers to look through some of the tools at [the Research Hub](#), including the Data Browser, which provides access to deidentified, aggregate information.

In the Researcher Workbench's Registered and Controlled Tiers, registered researchers can look at individual cohorts of participants and geographic and demographic information. The Controlled Tier includes three-digit ZIP code information.

Discussion

- Dr. Baskir said that *All of Us* is creating a national data resource. The program's goal is to provide this data resource to researchers. Dr. Baskir said that the program's platform supports educators and emerging researchers and fosters a learn-by-doing approach.
- Individuals do not need separate registrations for different uses of the dataset.
- Dr. Baskir explained that *All of Us* does not currently combine with other datasets. Although the program is interested in doing so, that is not yet an option. Ms. Wilfong said that the program is working toward this goal, but there is currently no timeline for reaching it. Dr. Baskir suggested that individuals involved in other cohort studies with datasets that might be linked to *All of Us* in the future email him directly at rubin.baskir@nih.gov or contact the Division of Engagement and Outreach at allofusengagement@nih.gov with additional questions.
- Ms. Wilfong recommended that attendees [subscribe to the All of Us Research Roundup newsletter](#) for program announcements.
- Dr. Baskir encouraged attendees to complete the [post-event evaluation form](#).
- Dr. Baskir explained that the qualifications of a minority-serving institution come from the U.S. Department of Education.
- Ms. Wilfong encouraged attendees to keep in touch with the program. *All of Us* is especially interested in the work completed by using the program's dataset.
- Dr. Baskir directed attendees to the resources in the [Support Hub](#). He said the goal is to make the process for using the dataset as easy as possible.
- Individual researchers have to wait until the DURA process is complete before they can sign up on the Researcher Workbench. Once the data use agreement is in place, fulfilling the other requirements of the registration process takes only a couple of hours.
- Written summaries of the breakout sessions will be made available to the attendees. A recording of the rest of the onramp event, excluding the breakout sessions, will be made available.

Adjournment

The breakout session was adjourned at 2:19 p.m. ET.

Institutional Onramp Event Breakout Room 3 Meeting Summary October 3, 2022

Breakout Room 3: Utilizing the Researcher Workbench

Moderators: *Ashley Green, VUMC/DRC, and Gage Rion, M.Eng., VUMC/DRC*

Discussion

This 25-minute breakout session provided an opportunity for attendees to learn how to use the Researcher Workbench.

Mr. Rion leads the help desk team, and Ms. Green is the lead outreach coordinator for the DRC.

Mr. Rion took attendees through the [Researcher Workbench](#), highlighting resources and going over the structure of the Workbench.

Attendees learned how Workspaces spend allocated funds. Mr. Rion also shared that newly registered researchers receive \$300 in initial compute credits to conduct research.

Lastly, attendees learned about support resources. Mr. Rion explained how users can use the tiles on the [User Support Hub to find more information](#). For example, the “[Access and DURA Support](#)” tile takes users to more information on gaining institutional access to the Workbench. Attendees first learned

how to log in using the [Researcher Login](#) button at the top of the [Research Hub](#) homepage.

Then, attendees learned about the following features that are available after logging in:

- The Quick Tour and Videos at the bottom of the landing page that provide 2.5-minute tutorials
- How to create a new workspace
- How to search what is publicly displayed in the [Research Projects Directory](#)
- Using inclusion and exclusion criteria to [build a cohort](#)
- Building datasets with [cohorts and concept sets](#)
- Using the [Jupyter Notebook](#) space to analyze the data and customize the analysis
- The [costs](#) associated with cloud computing and the use of compute credits
- The two ways to get to the [User Support Hub](#) from the Workbench: from the left-hand drop-down menu and from the public [Research Hub](#)

Questions from attendees included the following:

- How are Observational Medical Outcomes Partnership common data model and Observational Health Data Sciences and Informatics tools integrated into the system?
 - *All of Us* utilizes the [Observational Medical Outcomes Partnership \(OMOP\) common data model](#) (CDM) to organize and standardize raw data for collaborative research and large-scale analytics.
- Can additional tools or libraries be uploaded or integrated?
 - External files can be uploaded for use in notebooks. [Read more](#).