All of Us RESEARCH PROGRAM | The Future of Health Begins With You

The All of Us Research Program: Cohort Diversity and Data Diversity - American Society of Human Genetics (ASHG) Annual Conference

Tuesday, October 15, 2019
10:45 AM-12:45 PM CDT

Location:
Hilton Americas-Houston Hotel
1600 Lamar Street, Houston, TX 77010
Room Assignment: Lanier Ballroom GH, Level 4

REGISTER

AGENDA

- **Opening Remarks**

  Joshua C. Denny, MD, MS, Vanderbilt University Medical Center

Part I - Overview of the *All of Us* Research Program

- ***All of Us* Research Program Overview and Accomplishments**

  Kelly Gebo, MD, MPH, National Institutes of Health

- **Genomic Data Report Generation and Return of Results for One Million *All of Us* Participants**

  Richard Gibbs, PhD, Baylor College of Medicine

- **Demonstration Project: Understanding the Diversity and Depth of *All of Us* Data**

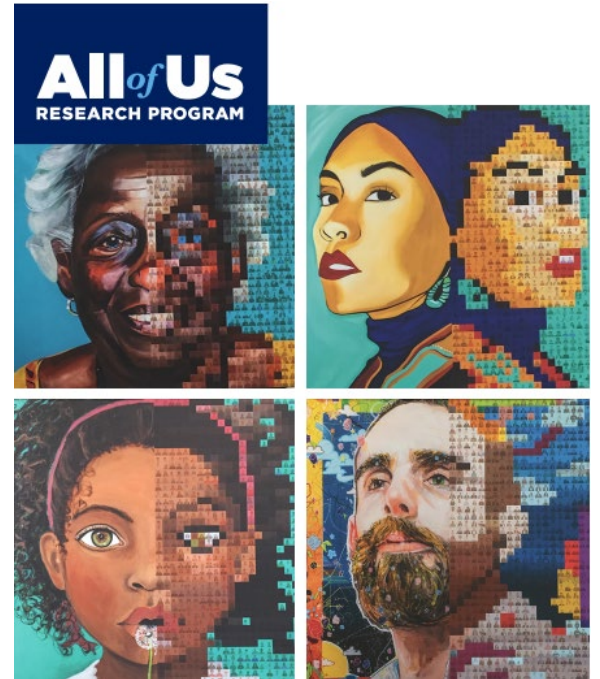 *Jordan W. Smoller, MD, ScD, Massachusetts General Hospital*

- **Q&A with Panel**

Part II - Demonstration and use of the *All of Us* Research Hub

- **Validating the Utility of *All of Us* and Research Hub Demo**

  Andrea Ramirez, MD, Vanderbilt University Medical Center

- **Q&A with Panel**

Artist Credit:
Top Left: Danielle Mastrion, Portrait of Kay (Harlem, NY); Top Right: Vania Soto, In Lak'ech Ala K'in (Kansas City, MO); Bottom Left: Liz Landgren, Addie Cecilia and the Dandelion (Birmingham, AL); Bottom Right: Julius DC Bautista, A More Colorful Tomorrow (Chicago, IL)

Learning Objectives:
This session will:

- Provide an overview of the *All of Us* Research Program and the status of recruiting a cohort of 1 million engaged participants

- Provide information about generation of and access to genomic data and discuss the program's approach to returning genomic results to research participants

- Describe findings from investigation of early demonstration projects using survey instruments and EHRs

- Demonstrate utility and functionality of the *All of Us* public data browser and other features of the Research Hub, a platform that will constitute one of the richest population datasets readily available to investigators.

FEATURED SPEAKERS AND PRESENTATION CONTENT

All of Us Research Program Overview and Accomplishments
Kelly Gebo, MD, MPH

Content: *All of Us* is a participant-engaged, data-driven enterprise supporting research at the intersection of human biology, behavior, genetics, environment, data science with the goal of developing more effective ways to treat disease. The *All of Us* Research Program is building a national research cohort of one million or more participants with rich diversity, particularly including those underrepresented in biomedical research and is currently enrolling all adults 18 and over in the U.S., with a plan to begin child enrollment in the future. The longitudinal data includes collection of electronic health records, participant provided information, physical measurements, biospecimens and digital health data. As of May 2019, data on over 143,000 participants, with over 75% groups historically underrepresented in biomedical research, have been collected. Projects to demonstrate the depth and breadth of currently collected data are underway, and access to this database through a researcher portal is anticipated to be available in the winter of 2019.

Genomic Data Report Generation and Return of Results for One Million *All of Us* Participants
Richard A. Gibbs, PhD

Content: This presentation will give an overview of the plans for genomic testing and return of results (ROR) in *All of Us*. The generation of accurate genotype and DNA sequence data in a secure environment is straightforward, but the optimal path for ROR to *All of Us* participants is challenging. The program is identifying the classes of information that are most actionable and useful for participants and developing a support framework, including links to genetic counseling and support, while building infrastructure for secure data management. The policy demands include contemplation of data security, individual privacy, exposure to risk information as well as the individual's rights to access their genetic data. The support framework links access to education and counselling with ROR. The infrastructural components provide variant interpretation, ahead of securely returning the reports to participants. Using automated procedures derived from related programs, *All of Us* will provide reports that combine high level summaries, together with more detailed information that can be used in clinical care. The automated procedures allow routine re-reporting when variant interpretations are updated through new knowledge. From previous work, we expect 2-3% of participants will be candidates for return of genetic data related to single gene markers with relevance to their ongoing health care. For other classes of information, including pharmacogenomic data, ancestry and polygenic risk score markers, all individuals can be informed by their genetic data. The pilot phase of the program aims to return results to at least 25,000 participants, and ultimately to scale to the option for ROR for all participants who desire the information.

Understanding the Diversity and Depth of *All of Us* Data

Jordan W. Smoller, MD, ScD

Content: Enrollment in *All of Us* involves two major avenues of geographically-diverse ascertainment: individuals recruited through health provider organizations (HPOs) and direct volunteers (DV) who may enroll from anywhere in the U.S. We will describe the *All of Us* recruitment and enrollment procedures and assessments, including efforts to ensure inclusion of individuals who have been historically underrepresented in biomedical research (UBR). The collection of diverse data types (electronic health records, surveys, biospecimens, physical measurements, and digital health data) will provide a uniquely rich longitudinal resource to enable broad research on risk and resilience factors, disease outcomes, and treatment response to advance precision medicine. The Data and Research Center (DRC) led by Vanderbilt University Medical Center is charged with demonstrating the quality, utility, and diversity of *All of Us* data and tools. With the broader *All of Us* Consortium, the DRC has designed Demonstration Projects aimed to describe strengths and opportunities within the data types including replication of known, previously published findings. This process allows interrogation of the data and development of research best practice examples as well as an initial characterization and growth projections of the *All of Us* data set and does not aim to discover new disease classifications, relationships, or risk associations. Successfully implemented Demonstration Projects provide a valuable teaching and training opportunity to users of *All of Us* data that include the diverse audience of researchers enabled by broad data access policies.

Validating the Utility of *All of Us* and Research Hub Demo

Andrea Ramirez, MD

Content: The *All of Us* Research Program is committed to the ambitious mission of collecting health data from a million or more participants to accelerate scientific discoveries in precision medicine. Delivering on this mission will result in the largest biomedical research dataset ever assembled. At this scale, the cost of downloading a dataset and storing it locally is expensive and impractical. Moreover, few researchers will have the necessary computational or security infrastructure to perform local analyses on such a large dataset. To address challenges in cost, accessibility, and security, the *All of Us* Data and Research Center (DRC) aims to "bring researchers to the data" by creating a cloud-based analytical platform that will help users test hypotheses in silico and readily and efficiently work within an environment that meets robust security and regulatory requirements. This platform, known as the *All of Us* Research Hub, will include layers of functionality to address researcher needs at various stages of research design and execution. This demonstration will provide an overview of the initial Research Hub analysis suite, including tools for study cohort identification and characterization for feasibility testing, data exploration and covariate selection, and analysis.

Learn more at https://www.joinallofus.org/en