Research Hub Overview | Introduction to the *All of Us* Data Browser (20 min)

Josh Denny, MD, MS, FACMI
Outline

- Data and Research Center (DRC) Background
- Research Hub Overview
  - Goals & User base
  - Public and restricted access
  - Data Browser Tour
  - Release Timeline
Background

The Data and Research Center (DRC) aims to:

- Acquire, organize, & provide secure access to scientific data
- Provide research support & analysis tools to facilitate use of scientific data

Who We Are:

- PIs: Josh Denny, MD, MS, FACMI (VUMC), David Glazer (Verily), Anthony Philippakis, MD, PhD (Broad Institute)
- VUMC, Broad, Verily, Columbia University and investigators from Northwestern University and University of Texas Houston
Responsibilities

- Operationalizes all data (receipt, curation, access)
- Executes data security agreements
- Develop secure researcher ecosystem & tools

Deliverables

- Data repositories (raw, curated)
- HealthPro (Enroll/id participants, sample collection & shipment support, participant status & look-up)
- EHR upload tools, quality control, & support
- Researcher Workbench access & analysis tools/support
- Operational dashboards (tools for monitoring enrollment, data tracking, audits)
Research Hub Interaction

https://www.researchallofus.org/

Static Data Snapshots

High-level summary statistics for cohort including:
- Enrollment volume over time
- Cohort demographics (age, race, etc.)

Research Hub Public Site

- Learn about the program
- Learn about access and billing
- Learn about our tools
- Learn about the data

Public Data Browser

Browse aggregate data

Research Workbench

Create a Workspace
- Cohort Builder
- Notebooks for Analysis

Register as an AoU Researcher & Apply for Data Access

More Tools
(To be determined)

This visual is an overview of how the user can interact with the Research Hub, and is not a site map or system architecture.
Research Hub Components

Launched May 2019

RESTRICTED

Researcher Workbench
(COMING SOON)

- Cohort Builder
- Notebooks
- Workspaces
- Help Desk

DATA PASSPORT

PUBLIC

- Data Browser
- Survey Explorer

Launched May 2019

Coming Winter 2019
Data Tiers

- Registered tier Requirements
  - Institutional sign-off

- Identity Verification
- Research Ethics Training
- Attestation to the Data Use Agreement

- Publicly Accessible (No login requirements)

Controlled
- Amended Participant-Level
- Additional data types

Registered
- Amended Participant-Level

Public
- Summary Statistics
The Data Browser

The Data Browser is:
- Interactive view of the public tier participant data
- Counts & visualizations of electronic health record (EHR) data, survey responses, & physical measurements taken at enrollment

Goals:
- Allow researchers to generate hypothesis & assess study feasibility
- Allow researchers, participants & public to explore the data

Status:
- Launched May 6th, 2019 – displays data for 116,460 participants
- Data refreshed with each new CDR release
The Workbench (approved users)

The Workbench is:
- Restricted access cloud-based environment where approved users conduct analysis using registered or controlled tier data
- Users must be approved through data access application process

Suite includes:
- Project-specific collaborative workspaces
- Graphical tools for exploring and selecting subsets of scientific data
- Jupyter Notebooks for data analysis
- Tools and services for research support

Status:
- Registered tier to publicly launch this Winter – to display data for ~200,000 participants
Research Support

- **Help Desk**
  - General & technical support
  - Community forums
  - Knowledge base

- **Training Materials**
  - Tool user guides
  - Instructional videos
  - Data model tutorials
  - Workspaces w/ example analyses
  - Reusable “code snippets”

- **Data Dictionary**
  - PDF as part of detailed data documentation
  - Integrated directly within Workbench tools
Research Hub Releases

Q2 2018

- Research Hub v1
- Splash page with overview of Research Hub components
- **Key feature:** Sign up for updates

Q2 2019

- Research Hub v2
- Expanded website featuring more granular pages for Research Hub components
- **Key features:** Data Browser + Survey Explorer

Winter 2019

- Research Hub v3
- Full suite of pages released, showcasing full life cycle of the Research Hub experience
- **Key features:** Workbench access and tools