June 24, 2019

ELSI Workshop

Kelly Gebo, MD MPH
CMSO, All of Us Research Program
Acknowledgements

Science Committee

- Keisha Bellamy
- Eric Boerwinkle
- Mine Cicek
- Elizabeth Cohn
- Stephanie Devaney
- Pam Factor-Litvak
- Mona Fouad (co-chair)
- Kelly Gebo
- David Goldstein
- Philip Greenland
- Parinda Khatri
- Christopher O'Donnell
- Lucila Ohno-Machado
- Akinlolu Ojo
- Andrea Ramirez
- Steven Reis
- Dan Roden
- Jordan Smoller (co-chair)
- Steven Steinhubl
- John Wilbanks
- Robert Winn

Methods Co-Chairs

- Assays
  - Maria Argos
  - Steve Thibodeau
- DHT
  - Ed Ramos
  - Chris Lunt
  - James McClain
  - Steve Steinhubl
- EHR
  - Abel Kho
  - Daniella Meeker
  - Andrea Ramirez
- Omics
  - Bruce Korf
- PPI
  - Rob Cronin
  - Brian Ahmedani

Consortium Members

- LCT
- IC Collaborators
- NIH/Leidos Team
  - Keri Althoff
  - Brian Florence
  - David Galey
  - Tram Huyen
  - Sara Lioi
  - Adele Luta
  - Steve Mockrin
  - Heather Sansbury
  - Sheri Schully
  - Carlie Williams
Disclosures

- Research funding: AHRQ, HRSA, NIAID, NIDA, NIA, Robert Wood Johnson Foundation, Johns Hopkins CFAR
- Scientific Consultant: Simon Fraser University
Objectives

● Understand the mission, objectives, and scientific framework of the All of Us Research Program

● Appreciate the data currently being collected within the All of Us Research Program

● Be able to identify research questions where All of Us could serve as a data source
Framingham Heart Study

Enrolled 5209 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 impact of Framingham (and similar cohorts) has been dramatic

https://www.cdc.gov/Mmwr/preview/mmwrhtml/mm4830a1.htm
Our Mission

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

Nurture relationships with one million or more participant partners, from all walks of life, for decades

Deliver the largest, richest biomedical dataset ever, making it as easy, safe, and free to use as possible

Catalyze a robust ecosystem of researchers and funders hungry to use and support it
All of Us Research Program Core Values

- Participation is **open** to all.
- Participants reflect the rich **diversity** of the U.S.
- Participants are **partners**.
- Trust will be earned through **transparency**.
- Participants have **access** to their information.
- Data will be accessed **broadly** for research purposes.
- **Security and privacy** will be of highest importance.
- The program will be a catalyst for **positive change** in research.
Why Diversity?

PERSISTENT BIAS

Over the past seven years, the proportion of participants in genome-wide association studies (GWAS) that are of Asian ancestry has increased. Groups of other ancestries continue to be very poorly represented.

2009
373 studies
1.7 million samples

96% European ancestry

4% Non-European ancestry

2016
2,511 studies
35 million samples

81% European ancestry

19% Non-European ancestry

Popejoy & Fullerton, Nature 2016

Terms for ethnicity are those used in the GWAS catalog. Some have changed between 2009 and 2016 as sampling has increased. Samples of European origin have the most specific descriptions of population ancestry.

4% GWAS represents >33% US population
Innovative Aspects of the *All of Us* Research Program

- Diversity at the scale of 1 million people or more
- Longitudinal, able to recontact
- EHR, surveys, baseline physical evaluation and biospecimens—including genomics
- Focus on participants as partners
- National, open resource for all: open to all researchers with open source software & tools
Participants as Partners
Involved in every step of program development

- What data we collect
- What lab analyses we do
- What research is conducted
- How data is returned
- Partnership with national and local community groups
Scientific Framework
Scientific Framework

Enable research that will:

I. Increase wellness and resilience, and promote healthy living

II. Reduce health disparities and improve health equity in underrepresented in biomedical research (UBR) populations

III. Develop improved risk assessment and prevention strategies to preempt disease

IV. Provide earlier and more accurate diagnosis to decrease illness burden

V. Improve health outcomes and reduce disease impact through improved treatment and development of precision interventions
Example Use Case: Blood Pressure

**Health Equity**
What is the impact of economic stability on rates of screening, likelihood of receiving treatment, and blood pressure levels?

**Wellness & Resilience**
What genomic, environmental, and lifestyle factors underlie the different patterns in age-related trajectories of blood pressure, thereby increasing or reducing the risk of high blood pressure?

**Risk & Prevention**
How do age-related changes in blood pressure in children and young adults impact the development of hypertension and hypertension-related conditions in adulthood?

**Diagnosis**
Does blood pressure from ambulatory blood pressure monitoring (ABPM) or home blood pressure monitoring (HBPM) provide a more accurate estimate of cardiovascular risk?

**Treatment & Outcomes**
What are effective and scalable community-based interventions to improve blood pressure levels, medication prescription and medication fill rates?
Resilience in Chronic Disease and Health outcomes

Risk & Prevention
What (if any) impact do environmental and behavioral factors (e.g. diet/dietary supplement use, air quality, physical activity, etc.) have on resilience to health outcomes in various subpopulations?

Health Equity
How does resilience to negative health outcomes manifest differently in various subpopulations (race/ethnicity, age, sex, etc.)?

Wellness & Resilience
What are the physiological factors that influence (and physiological markers that indicate) resilience to chronic disease and other health outcomes in low- and high-risk populations?

Monitoring & Diagnosis
Can new technologies be used to monitor physiological changes to various stressors in order to identify biological disturbances (or tipping points)? Which biological changes promote successful management of stressors in individuals that show resilience to chronic disease and other health outcomes?

Treatment & Outcomes
Which interventions might enhance resilience outcomes in various subpopulations?
Example Use Case: Wellness

Risk & Prevention
How can environmental exposures cause unhealthy living?

Health Equity
What are the most important factors relating to sociodemographic health disparities in wellness?

Wellness & Resilience
What are the best predictors of quality of life?

Diagnosis
What community health programs/services (e.g., home care service, self-care monitoring and community healthcare service etc.,) can reduce disease burden?

Treatment & Outcomes
How do non-pharmacological interventions impact health/resilience?
Focusing
The “Big 8”
Big 8: Selected Health Areas

1. Cancer
2. Cardiovascular Disease
3. Chronic Kidney Disease
4. Chronic Lung Disease
5. Diabetes/Obesity
6. Mental Health/Cognition
7. Opioid Use/Chronic Pain
8. Wellness
Current Participant Journey
# Current protocol

<table>
<thead>
<tr>
<th>Enroll, Consent &amp; Authorize EHR</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Recruiting 18+ years old initially; plan to include children in 2019</td>
</tr>
<tr>
<td>- Online, interactive consent</td>
</tr>
<tr>
<td>- Includes authorization to share Electronic Health Record (EHR) data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Answering Surveys</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Initial surveys: The Basics, Overall Health, Personal Habits, Health Care Access &amp; Utilization, Family Medical History</td>
</tr>
<tr>
<td>- Additional surveys released on an ongoing basis.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Physical Measurements*</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Blood pressure</td>
</tr>
<tr>
<td>- BMI</td>
</tr>
<tr>
<td>- Heart rate</td>
</tr>
<tr>
<td>- Height</td>
</tr>
<tr>
<td>- Hip circumference</td>
</tr>
<tr>
<td>- Waist circumference</td>
</tr>
<tr>
<td>- Weight</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Provide Biosamples*</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Blood (or saliva, if blood draw is unsuccessful)</td>
</tr>
<tr>
<td>- Urine specimen</td>
</tr>
<tr>
<td>- Biosamples will be stored at the program’s biobank</td>
</tr>
</tbody>
</table>

*Based on diverse sampling and capacity

<table>
<thead>
<tr>
<th>Wearables and Digital Apps</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Share data from wearable fitness devices, starting with FitBit</td>
</tr>
<tr>
<td>- Share data about mood &amp; cardio-respiratory fitness through integrated apps</td>
</tr>
<tr>
<td>- More integrations to come</td>
</tr>
</tbody>
</table>

Coming soon
### Schedule of Assessments Template: For a Participant Enrolling June 24

<table>
<thead>
<tr>
<th>Assessment</th>
<th>7/19-9/19</th>
<th>10/19-12/19</th>
<th>1/20-3/20</th>
<th>4/20-6/20</th>
<th>Year 2</th>
<th>Year 3 reassessment</th>
<th>Year 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical Measurements:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>Height, weight, blood pressure, heart rate, hip circumference, waist circumference</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Biospecimens:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TBD</td>
</tr>
<tr>
<td>Blood (50 mL), urine (30 mL), saliva</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PPI</strong></td>
<td>Basics, Lifestyle, and Overall Health</td>
<td>Healthcare Access and Utilization, Personal Medical History, and Family History</td>
<td>Mental Health</td>
<td>TBD</td>
<td>Reconfirm Basics, Lifestyle, and Overall Health</td>
<td>TBD</td>
<td></td>
</tr>
<tr>
<td><strong>Omics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Genotyping and WGS</td>
<td></td>
</tr>
<tr>
<td><strong>EHR</strong></td>
<td>Consent to obtain EHR (day 0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>DHT</strong></td>
<td>BYOD Fitbit</td>
<td>Project Fitbit (TBD)</td>
<td>Apple (TBD)</td>
<td>Mood App</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
All of Us Research Hub

PUBLIC

Research Hub Website
- Data Browser
- Survey Explorer
- Data Snapshots

Restrict

Data Access Application

RESTRIC

Researcher Workbench
- Cohort Builder
- Workspaces
- Help Desk

COMING SOON!
Data Access | Data Access Principles and Framework

• Data available to **all types of users**
• Employ a **cloud-based analysis platform**
• Access will be **tiered**
• Users will be granted **data passports**
• Project information will be made **public and auditable**
• Developing policies on **access to samples & cohort**
Journey to Protocol Roadmap
## Protocol Plan and Timeline

<table>
<thead>
<tr>
<th>Task</th>
<th>Who</th>
<th>Due</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Determine Selected Health Areas*</td>
<td>SciCom</td>
<td>April 5</td>
<td>✓</td>
</tr>
<tr>
<td>B1. Identify candidate variables associated with each Selected Health Area</td>
<td>SciCom</td>
<td>May 1</td>
<td>✓</td>
</tr>
<tr>
<td>B2. Cohort Gap Analysis</td>
<td>Kelly/NIH</td>
<td>May 8</td>
<td>✓</td>
</tr>
<tr>
<td>C. Harmonize candidate variables</td>
<td>NIH/Leidos</td>
<td>May 17</td>
<td>✓</td>
</tr>
<tr>
<td>D. Determine “core” variables</td>
<td>NIH to recommend to SciCom</td>
<td>May 23</td>
<td>✓</td>
</tr>
<tr>
<td>E. Create list of “core” datatypes from “core” variables</td>
<td>NIH</td>
<td>May 24</td>
<td>✓</td>
</tr>
<tr>
<td>F. Assign status (in protocol, in development, from RPW) to “core” datatypes</td>
<td>NIH/Leidos</td>
<td>May 24</td>
<td>✓</td>
</tr>
<tr>
<td>G. Filter by “in protocol” and “in development” from “core” datatypes</td>
<td>NIH/Leidos</td>
<td>May 24</td>
<td>✓</td>
</tr>
<tr>
<td>H. Recommend methods to collect datatypes in remaining list</td>
<td>Methods co-chairs</td>
<td>Week of May 27</td>
<td>✓</td>
</tr>
<tr>
<td>I. Prioritize optimal datatype/method pairs to NIH (NIH chaired WebEx)</td>
<td>SciCom with Methods co-chairs</td>
<td>Week of June 3</td>
<td>✓</td>
</tr>
</tbody>
</table>

### Decision Points

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes/No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is this variable associated with one or more of the Selected Health Areas?</td>
<td>No</td>
</tr>
<tr>
<td>Is the AoU platform the right platform to capture this variable (on 1 million+ people or a significant subset of the population)?</td>
<td>Yes</td>
</tr>
<tr>
<td>Is there currently an optimized method for collecting this datatype?</td>
<td>No</td>
</tr>
<tr>
<td>Does this data type/method leverage and balance expertise across AoU program and can be implemented now?</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### Selected Health Areas
1. CVD/Cerebrovascular Disease
2. Cancer
3. Diabetes and Obesity
4. Opioids, Marijuana, Alcohol, and Pain
5. Mental Health and Cognition
6. Chronic Lower Respiratory Diseases
7. Chronic Kidney Disease
8. Wellness

### Methods Criteria
1. Affordable at scale
2. Low participant burden
3. Does not pose a safety risk
4. Currently available, used, valid & reliable
5. Advances innovation in data collection

### Decision Pathways
- Variable should be considered for an ancillary study or by an NIH IC
- Variable should be considered for an ancillary study or by an NIH IC
- Consider methods pilot project funded by AoU or an NIH IC
- Bin Near term
- Bin Medium or Long term
## Protocol Plan and Timeline, cont’d.

<table>
<thead>
<tr>
<th>Task</th>
<th>Who</th>
<th>Due</th>
<th>Process/Criteria</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>J1. Prepare F2F presentation</td>
<td>NIH/SciCom</td>
<td>June 6-June 11</td>
<td>we are here</td>
<td></td>
</tr>
<tr>
<td>J2. Obtain F2F SC input</td>
<td>SciCom to SC</td>
<td>June 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J4. Continue to prioritize optimal datatype/method pairs</td>
<td>SciCom w/ Methods Co-Chairs</td>
<td>TBD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J3. Obtain selected LCT member review</td>
<td>NIH</td>
<td>TBD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>K. Ensure no populations are alienated</td>
<td>Ambassadors &amp; Community Engagement Partners</td>
<td>TBD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. Prioritize PPEs within each time-frame = Draft Road Map</td>
<td>NIH to DCM</td>
<td>TBD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M. Create Schedule of Assessments &amp; Reassessments with near term PPEs = Protocol V2</td>
<td>NIH to DCM</td>
<td>TBD</td>
<td></td>
<td>*Additional health areas will be considered for future protocol iterations</td>
</tr>
</tbody>
</table>

*Additional health areas will be considered for future protocol iterations*
Variables Under Consideration
Guiding Principles (GP) for Selecting Methods*

- **GP 1**: The method(s) chosen should be parsimonious (affordable at scale; low participant burden; available, reliable, and valid; advances innovation in data collection)

- **GP 2**: The method(s) chosen will collect the datatype from as many participants as possible (e.g. one million participants or a significant subset of them)

- **GP 3**: If we can’t get a datatype from all participants using one method, we use a 2\textsuperscript{nd} method to get the datatype from the remaining participants (e.g., If we don’t get a datatype from all participants using EHR, we use PPI to get the datatype from the remaining participants)

- **GP 4**: The method(s) should address the heterogeneity and quality of the data from different UBR populations

*Note: numbers are used to facilitate discussion, not to indicate priority*
## In Protocol / In Development

### Kidney Disease

#### Diagnosis of Kidney Disease (PPI)

<table>
<thead>
<tr>
<th>Variables for consideration</th>
<th>NEAR-TERM</th>
<th>MID TO LONG-TERM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment of Kidney Disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemodialysis</td>
<td>✓</td>
<td>🍀 (Linkage to USRDS)</td>
</tr>
<tr>
<td>Peritoneal Dialysis</td>
<td>✓</td>
<td>🍀</td>
</tr>
<tr>
<td>AV Fistula placement</td>
<td>✓</td>
<td>🍀</td>
</tr>
<tr>
<td>Complication of Kidney disease (hyper PTH)</td>
<td></td>
<td>🍀</td>
</tr>
</tbody>
</table>
**All of Us Research Program Protocol Development**

**Cancer**

**In Protocol / In Development**
Diagnosis of Cancer, Treatment for Cancer (PPI)

<table>
<thead>
<tr>
<th>Variables for consideration</th>
<th>NEAR-TERM</th>
<th>MID TO LONG-TERM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EHR</td>
<td>Exam</td>
</tr>
<tr>
<td>Diagnosis of Cancer</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Treatment for cancer</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Cancer recurrence</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Tumor characteristics</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Dx of precancerous condition (FAP)</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
Remember

- Proposed variable lists include:
  - *Italicized Blue* are in protocol or in development
  - Black are for consideration
  - Lists are organized by outcome, sociodemographic factor, SDOH, risk factor, exposure, lab tests, genetics and omics

- Remember: This is NOT FINAL

- We are currently seeking feedback
We ask for your feedback to the following questions on the new core datatypes* under consideration:

- What is blatantly missing (variables and methods)?
- What would draw you to this dataset that you don’t see represented (variables and methods)?
- Should something be removed from this list?

**RESPONSE**

- Please email your response to Kelly Gebo at kelly.gebo@nih.gov by COB June 28, 2019

*1M+ participants or a large subset of the population*
Discussion Questions
Innovative Aspects of *All of Us*

- Diversity at the scale of 1 million people: demographically, geographically, medically, and especially those underrepresented in biomedical research

- Diversity of data types collected longitudinally: clinical, environmental, genetic, behavioral, socioeconomic

- Focus on participants as partners: included in governance, invited to co-invent systems and give input into the science, choice to receive all data and information back

- National, open resource for all: open to the public and all researchers, open source software & tools

*All of Us* learns from and partners with other large research programs; sharing knowledge and data is key!
For Discussion

- How do we retain the innovative aspects of All of Us while collecting a complete dataset as possible on one million participants?
  - Designing things for a million, but anticipate others will build focused cohorts for deeper phenotypic assessments
  - Propose to collect new variables by EHR initially and after one year assess the completeness of the data
  - What variables are so important that if EHR is not complete, we would employ another method to complete the dataset (e.g., DHT, PPI, Assays)
    - Spirometry
    - Gait analysis

- Should we hold off on doing assays (as technology will become more advanced, cheaper, and efficient) as suggested by Assays Task Force and external stakeholders?
For Discussion

- Realizing we are trying to minimize the burden to participants by maximizing data collection from EHR and passive data collection methods with DHT

- How do we include important ELSI variables/questions into the All of Us scientific protocol roadmap and next version of the protocol?
  - What are the most important research questions?
  - What variables are needed to answer these?
  - What methods are used to collect these variables?
We’ve done a lot in 2.5 years!

- **3/1/16**: PMI IRB est.
- **10/1/15**: Advisory panel est.
- **1/20/15**: PMI announced in State of the Union Address
- **2/26/16**: PMI pilot S4S awarded
- **9/17/15**: NIH ACD report
- **7/6/16**: AoU research program kickoff
- **5/31/17**: First participant enrolled
- **12/13/16**: 21st Century Cures Act
- **1/31/17**: AoU research program engagement partners awarded
- **5/9/17**: First AoU IRB protocol approved
- **7/1/17**: Community partners awarded
- **7/5/18**: Spanish enrollment launch
- **10/5/17**: 3 new HPOs awarded
- **5/6/18**: National launch, 27,829 core participants

**Timeline Events**

- **2015**
  - PMI Announced in State of the Union Address
  - PMI Research Program Kickoff
  - NIH ACD Report

- **2016**
  - PMI Pilot S4S Awarded
  - AoU Research Program Kickoff
  - First Participant Enrolled
  - 21st Century Cures Act

- **2017**
  - AoU Research Program Engagement Partners Awarded
  - First AoU IRB Protocol Approved
  - Community Partners Awarded

- **2018**
  - Spanish Enrollment Launch
  - 3 New HPOs Awarded
  - National Launch, 27,829 Core Participants

- **2019**
  - 100K

**Key Milestones**

- Discovery and impact for decades
- All of Us network of awardees
- Small groups at NIH
- Researcher portal launch
- Genomics pilot
- Assay pilot

**Additional Notes**

- **As of 2019**: 200K (anticipated)
- **As of 2019**: Small groups at NIH
All of Us consortium members

**DV Network**
(Direct Volunteers)

**HPO Network**
(Health Care Provider Organizations)

**RMCs**
California Precision Medicine Consortium

**Illinois**
Precision Medicine Consortium

**New England**
Precision Medicine Consortium

**Trans-American Consortium**
for the Health Care Systems Research Network

**New York City**
Precision Medicine Consortium

**Southern**
All of Us Network

**SouthEast Enrollment Center**

**All of Us, Wisconsin**
Marshfield Clinic

**University of Arizona**

**University of Pittsburgh**

**FQHCs** (Federally Qualified Health Centers)

**VA Medical Centers**

**Platform Development**

**Genomics Infrastructure**

40
All of Us: Current Community Partners Network
It takes *All of Us*....
For more information...

ResearchAllofUs.org

@AllofUsResearch
#JoinAllofUs

AllofUs.nih.gov

databrowser.researchallofus.org
Questions?