7th Annual Big Data in Biomedicine Symposium

Next Generation Data Commons

Transforming Human Health

Oct 26, 2018
1) How Data Commons are changing the way that Large Biomedical Datasets are analyzed and shared and how can it be used to accelerate research transformation?

2) The All of Us, Pediatric data commons, Imaging data commons and others hold a promise to acquire some of the largest ever collected longitudinal data. These are ambitious, multifaceted programs integrating data from multiple sources including genomics:
   - Once collected, what are the planned infrastructure capabilities to support cleaning, analyzing storing curating data to making it easy for the research community to use these assets as well as for participants to access their information—what are the challenges?
   - Are there enough genetic counselors to support participants with “return on information”?
   - What are some unique challenges for genetic vs clinical vs imaging data commons?

3) What emerging standards supporting data commons are under development?
Create technologies that inspire people and their communities to make data-driven decisions to achieve the best health outcomes possible.
Vibrent Background

- A graduate of NIH SBIR program. Over a 7 year period built a Digital Health and Precision Medicine Platform

- Winner of $90+ Million NIH grant to create the largest deployment of digitally enabled longitudinal cohort research for All of Us Precision Medicine Research program

- Digital Health Platforms for prevention, public health, biomedical research, digital clinical trials, medical devices, digital therapeutics, life sciences and health systems

- Engaging global research and collaboration partnerships by building an ecosystem to improve health of all
Digital, Multi-omics, Analytic Health Technologies, Data Collection and Data Curation

• Many diseases, many health conditions, multiple co-morbidities
• Multiple data points, disparate data sources & many factors define comprehensive outcomes.
• Human ingenuity & Vibrent’s technology platforms bridge the gap for today and for future generations.
# Technology Infrastructure for Accelerating Health Research and Care

<table>
<thead>
<tr>
<th>Verticals</th>
<th>Horizontals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>Cardio-vascular conditions</td>
</tr>
<tr>
<td>Cardiovascular conditions</td>
<td>Chronic pain</td>
</tr>
<tr>
<td>Chronic pain</td>
<td>Immuno &amp; inflamma conditions</td>
</tr>
<tr>
<td>Immuo &amp; inflamma conditions</td>
<td>Infectious disease</td>
</tr>
<tr>
<td>Infectious disease</td>
<td>Mental health &amp; addiction</td>
</tr>
<tr>
<td>Mental health &amp; addiction</td>
<td>Metabolic conditions &amp; obesity</td>
</tr>
<tr>
<td>Metabolic conditions &amp; obesity</td>
<td>Movement/musculo-skeletal</td>
</tr>
<tr>
<td>Movement/musculo-skeletal</td>
<td>Neuro/cognitive conditions</td>
</tr>
</tbody>
</table>

- **Prevention & Wellness**
- **Reducing Disparities/Improving Access**
- **Genomics**
- **mHealth**
- **Environmental Exposures**

**Verticals** = High-level condition/disease areas; **Horizontals** = Cross-cutting areas of interest/themes
Definition of Precision Medicine is expanding – Health Research Informatics to Consumer Health Informatics

Countless diseases and conditions. Billions of data points. Myriad correlation variations.

genes
environment
lifestyle
behaviors

Technology, Informatics and Data Science packaged into relevant consumer-centric experiences can turn information to knowledge, knowledge to insights and insights to action.
Within the area of digital health emerging methods and approaches to address clinical needs and advance Precision Medicine

Enabling Personalized Health requires collection, processing and machine learning of data from broad sources.
Vibrent Approach to Health Innovation – Collaboration Between Medical Science, Health Research and Technology

Digital Health Collaboration with broad health R&D eco-system – health policy makers, academia, research institutes, patient advocates, innovators, pharma, health networks, payers, established organizations and public health programs pioneering big ideas for health and hope.
An Introduction to the

All of Us℠ Research Program

NIH
National Institutes of Health
“My hope is that this becomes the foundation, the architecture, whereby in 10 years from now we can look back and say that we have revolutionized medicine.”

—President Barack Obama
NIH All of Us Research Program – Kick off with Dr. Collins

Phase 1
Human Genome Project

Phase 2
Precision Medicine Initiative

Phase 3
Scientific Breakthroughs Using Data
NIH *All of Us* Research Program – Scope and Scale

- 1-10 Million participants across the United States
- 10-40 years engagement
- Reflecting broad diversity of the U.S.
- Data will inform 1,000s of research studies
- $2 Billion, national program
NIH *All of Us* Research Program

- **Mission:** To accelerate health research and medical breakthroughs, enabling *individualized prevention, treatment,* and *care* for all of us.

- **How:** Deliver a *national resource* of deep clinical, environmental, lifestyle, and genetic data from *one million participants* who are consented and engaged to provide data on an ongoing, *longitudinal* basis (60+ years!).

- **Priority:** Reflect the *broad diversity* of the U.S. – all ages, races/ethnicities, gender, SES, geo, and health status – by over recruiting those *underrepresented* in biomedical research.

- **Priority:** Build the *tools & capabilities* that make it easy for researchers from citizen scientists to premier university labs to make discoveries using the *data & bio-samples* and through *ancillary studies* with the cohort.
A Transformational Approach to Diversity

Reflecting the country’s rich diversity to produce meaningful health outcomes for historically underrepresented communities.
Different kinds of consortium partners w/ different expertise & responsibilities

1. Branding & Content
2. Education & Awareness
3. Outreach & Recruitment
4. Enrollment: Informed Consent, PPI, EHR
5. Baseline Evaluation & Biospecimen Collection
6. Sustained Engagement

DV Track (Direct Volunteers)
- WONDROS
- Scripps Translational Science Institute
- Sage
- Vanderbilt University
- National Blood Collaborative (BBB Collaborative)
- QTC
- Walgreens
- BlueCross BlueShield

HPO Track (Health Care Provider Organizations)
- HCM
- vibrent
- CSC
- EMSI Health
- WebMD
- Quest Diagnostics

RMCs
- UC San Diego Health
- UC HEALTH
- Cedars-Sinai
- UC Irvine Health
- UC Davis Health
- Northwestern University
- UIC
- NorthShore
- RUSH
- Massachusetts General Hospital
- Partners HealthCare
- University of Pennsylvania
- University of Virginia
- University of North Carolina
- VA
- The University of Arizona
- Essentia Health
- BayCare St. Joseph
- SPECTRUM HEALTH
- NewYork-Presbyterian

FQHCs
- Community Health Care
- Kaiser Permanente
- VA Medical Centers
- U.S. Department of Veterans Affairs
AllofUs Research Program Awardees and their Role

**PARTICIPANT TECHNOLOGY SYSTEMS CENTER (PTSC)**
Web, mobile and wearables based platforms for participants

Vibrent Health

**DATA AND RESEARCH CENTER (DRC)**
Big data capture, cleaning, curation, and sharing in secure environment

Vanderbilt, Verily, Broad Institute

**HEALTH CARE PROVIDER ORGANIZATIONS (HPOs)**
Clinical and scientific expertise network, enrollment and retention of participants

20+ regional med centers, FQHCs, VA, future awards to grow network

**BIOBANK**
Repository for processing, storing, and sharing bio-samples (35+M vials)

Mayo Clinic

**COMMUNICATIONS & ENGAGEMENT**
Comms, marketing, design expertise; engagement coordination and community partners network

Wondros, HCM, future awards to grow network of community partners

**PARTICIPANT CENTER**
Direct volunteer participant enrollment, digital engagement innovation, and consumer health technologies

Scripps Research Institute (with multiple partners)
Vibrent’s Role is Central to the Success of the AoURP Nationwide Next Generation Research Infrastructure for this Large Consortium

Vibrent is the Participant Technology Systems Center (PTSC).

**Design, build and Operate** digital health platform to enroll and engage 1+ million consumers in a long term longitudinal study of genomics, bio-samples, clinical, environmental, behaviors, lifestyle and sensor data.

Provides tools, dashboards and analytics systems for researchers and citizen scientists.
Participant Life-Cycle & Engagement – Participant Portal, Mission Control, Analytics

Study “work flow” and how the multiple stakeholders participate and are updated

Extensive Multi-Layer Security (Participant & Institution)
Participant Portal: shows visually the integration and complexity of the workflow
Global Learning Health System

Future of Healthcare and Clinical Research

Need for Large Data Commons
Stage 1: Era of Individual Study Silos

n = 100s or 1000s?
Stage 2: Era of Competitive Consortia (larger silos?)

n = 10,000s?
Stage 3: Era of Large Cohorts? (some as silos, some as national resource?)

$n = 1$Ms to $10$Ms?
Stage 4: Era of “America as a Cohort?”

n = 325,791,167
Stage 5: Era of a “Universal Cohort” (no silos? no cohorts?)

n = 7,395,927,821
Designed to Scale Using Machine Learning and AI to Personalize Healthcare for each individual

Governance Level Data

- profile
- preferences
- In-app data
- wearables data
- public sources of data
- survey responses
- EMA responses
- self-reports
- metadata

Support and personalize engagement at scale

Machine learning

personalization

iterative cycle of learning for making adjustments to rules, interface design, language, devices, etc.
1) How Data Commons are changing the way that Large Biomedical Datasets are analyzed and shared and how can it be used to accelerate research transformation?

2) The All of Us, Pediatric data commons, Imaging data commons and others hold a promise to acquire some of the largest ever collected longitudinal data. These are ambitious, multifaceted programs integrating data from multiple sources including genomics:
   - Once collected, what are the planned infrastructure capabilities to support cleaning, analyzing storing curating data to making it easy for the research community to use these assets as well as for participants to access their information—what are the challenges?
   - Are there enough genetic counselors to support participants with “return on information”?
   - What are some unique challenges for genetic vs clinical vs imaging data commons?

3) What emerging standards supporting data commons are under development?

4) Cancer biology knowledge is at the core of Precision medicine and can greatly benefit from quantitative image analysis to discover Imaging biomarkers (radiomics)
   - What is the future of radiomics in the bedside care and how is data being integrated and analyzed with genomics, proteomics and other -omics;

5) Open innovation is now a widely used concept in academia, business, and policy making: what kind of ethical, legal and security challenges should we be considering?
THANK YOU