The Precision Medicine Initiative Cohort Program – Building a Research Foundation for 21st Century Medicine

Precision Medicine Initiative (PMI) Working Group Report to the Advisory Committee to the Director

September 17, 2015

Kathy Hudson, PhD (NIH)
Rick Lifton, MD, PhD (Yale)
Bray Patrick-Lake, MFS (Duke)
Josh Denny, MD, MS (Vanderbilt)
“And that’s why we’re here today. Because something called precision medicine ... gives us one of the greatest opportunities for new medical breakthroughs that we have ever seen.”

President Barack Obama
January 30, 2015
To enable a new era of medicine through research, technology, and polices that empower patients, researchers, and providers to work together toward development of individualized treatments.
## PMI Proposed Support: FY16

<table>
<thead>
<tr>
<th>Agency</th>
<th>$ Million</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIH</td>
<td>$200</td>
</tr>
<tr>
<td>• Cancer</td>
<td>$70</td>
</tr>
<tr>
<td>• Cohort</td>
<td>$130</td>
</tr>
<tr>
<td>FDA</td>
<td>$10</td>
</tr>
<tr>
<td>ONC</td>
<td>$5</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>$215</strong></td>
</tr>
</tbody>
</table>
PMI Working Group of the Advisory Committee to the NIH Director (ACD)

- Working Group Charge: develop a vision for the PMI Cohort Program (PMI-CP) and advise on the design of a longitudinal national research cohort of ≥1 million volunteers

- Leverage existing cohorts, start from scratch, or hybrid?
- How to capture the rich diversity in the U.S. population?
- What data types should be included?
- What policies need to be in place for maximal benefit?
Advisory Committee to the NIH Director
Working Group On Building A Research Cohort
For Precision Medicine

Co-Chairs:
Richard Lifton, MD, PhD, Yale Univ School of Medicine, New Haven, CT
Bray Patrick-Lake, MFS, Duke Univ, Durham, NC
Kathy Hudson, PhD, National Institutes of Health

Members:

• Esteban Gonzalez Burchard, MD, MPH
  University of California, San Francisco

• Tony Coles, MD, MPH
  Yumanity Therapeutics, Cambridge, MA

• Rory Collins, FMedSci
  University of Oxford, UK

• Andrew Conrad, PhD
  Google X, Mountain View, CA

• Josh Denny, MD
  Vanderbilt University, Nashville

• Susan Desmond-Hellmann, MD, MPH
  Gates Foundation, Seattle

• Eric Dishman
  Intel, Santa Clara, CA

• Kathy Giusti, MBA
  Multiple Myeloma Res Foundation, Norwalk, CT

• Sekar Kathiresan, MD
  Harvard Medical School, Boston

• Sachin Kheterpal, MD, MBA
  University of Michigan Medical School, Ann Arbor

• Shiriki Kumanyika, PhD, MPH
  Perelman School of Medicine, Philadelphia

• Spero M. Manson, PhD
  University of Colorado, Denver

• P. Pearl O’Rourke, MD
  Partners Health Care System, Inc., Boston

• Richard Platt, MD, MSc
  Harvard Pilgrim Health Care Institute, Boston

• Jay Shendure, MD, PhD
  University of Washington, Seattle

• Sue Siegel
  GE Ventures & Healthymagination, Menlo Park, CA
Inputs

- **Workshops**
  - April 28-29: Unique Scientific Opportunities for the National Research Cohort (NIH)
  - May 28-29: Digital Health Data in a Million-Person Precision Medicine Initiative (Vanderbilt University, Nashville, TN)
  - July 1-2: Participant Engagement and Health Equity (NIH, Bethesda, MD)
  - July 27-28: Mobile and Personal Technologies in Precision Medicine (Intel Corp., Santa Clara, CA)

- **Requests for Information**
  - Building the cohort
  - Strategies to address community engagement and health disparities

- FNIH Survey of public perceptions of precision medicine cohort
- White House Privacy and Trust Principles
Why?

- Discover new biomarkers predictive of individual risk of future disease for many common diseases
- Understand individual variation in response to therapies
- Study populations reflecting diversity of the US population
- Accelerate research across many areas of health and disease
- Participant engagement and ongoing contact allows follow-up studies to advance understanding of disease mechanisms and targeted clinical trials.
## Why now?

<table>
<thead>
<tr>
<th></th>
<th>Ten Years Ago</th>
<th>Now – 2014 (most recent data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of sequencing a human genome</td>
<td>$22,000,000</td>
<td>$1,000 - $5,000</td>
</tr>
<tr>
<td>Amount of Time to Sequence a Human Genome</td>
<td>2 years</td>
<td>&lt;1 day</td>
</tr>
<tr>
<td>Number of smart phones in the United States</td>
<td>1 million (&lt;2%)</td>
<td>160 million (58%)</td>
</tr>
<tr>
<td>EHR Adoption (% hospitals)</td>
<td>20-30%</td>
<td>&gt;90%</td>
</tr>
<tr>
<td>Computing Power</td>
<td>n</td>
<td>n x 16 Deep Learning</td>
</tr>
</tbody>
</table>
Scientific Opportunities in the PMI-CP

- Discover new biomarkers predictive of future disease risk
- Discover determinants of individual variation in response to therapeutics
- Determine quantitative risk estimates in the population by integrating environmental exposures, genetic factors, and gene-environment interactions
- Integrate mHealth and sensor technologies
- Determine clinical impact of loss-of-function mutations on clinical outcome
- Discover new classifications and relationships among diseases
- Enable targeted clinical trials of subjects with rich clinical data
- Make ‘big data’ broadly available to investigators
Estimated disease incidences and prevalences in one million people

<table>
<thead>
<tr>
<th>Disease</th>
<th>Expected prevalent cases</th>
<th>Incident cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5 years</td>
</tr>
<tr>
<td>Type 2 Diabetes</td>
<td>135,658</td>
<td>40,411</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>73,723</td>
<td>21,315</td>
</tr>
<tr>
<td>Asthma</td>
<td>62,149</td>
<td>17,292</td>
</tr>
<tr>
<td>COPD</td>
<td>48,728</td>
<td>15,396</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>39,273</td>
<td>14,981</td>
</tr>
<tr>
<td>Epilepsy</td>
<td>33,426</td>
<td>4,161</td>
</tr>
<tr>
<td>Breast cancer (female)</td>
<td>20,470</td>
<td>12,068</td>
</tr>
<tr>
<td>Stroke</td>
<td>16,016</td>
<td>8,969</td>
</tr>
<tr>
<td>Lupus</td>
<td>14,659</td>
<td>3,283</td>
</tr>
<tr>
<td>Dementia</td>
<td>13,373</td>
<td>7,028</td>
</tr>
<tr>
<td>ADHD</td>
<td>13,039</td>
<td>7,213</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>9,407</td>
<td>3,745</td>
</tr>
</tbody>
</table>
Assembling the PMI Cohort

- One million or more volunteers:
  - Be recontactable
  - Collect EHR data, provide biospecimen, survey, and complete a baseline exam

- Longitudinal cohort, with continuing interactions, recontactable for secondary studies

- Two methods of recruitment
  - Direct volunteers
    - Anyone can sign up
  - Healthcare provider organizations (incl. FQHCs)
    - Consider HPO diversity, robustness of EHR, patient follow-up
Assembling the PMI Cohort

Broadly reflect the diversity of the U.S.
- Groups that are underrepresented
- All states of health and disease
- All areas of the U.S.
- All life-stages
- Special policy considerations
  - enrolling children
  - decisionally impaired
  - participants who become incarcerated
“...I’m proud we have so many patients’ rights advocates with us here today. They’re not going to be on the sidelines. It’s not going to be an afterthought. They’ll help us design this initiative from the ground up, making sure that we harness new technologies and opportunities in a responsible way.”

President Barack Obama
January 30, 2015
FNIH Survey of public opinion on a large US cohort study

- 79% agree cohort probably/definitely should be done
- 54% would probably/definitely participate in the cohort

What motivates participation?
- 82% interested in receiving results of study
- 62% wish to help advance health research

- 71% said participants should be partners with researchers
PMI-CP Focus on Engagement

- Highly interactive and proactive participant model
  - Participant representation in governance, design, conduct, dissemination, evaluation
  - Build a strong foundation of trust
- Participant engagement and communication activities should be centrally coordinated
- Consent is with PMI Cohort Program
  - Basic consent to be part of the cohort
  - Broad consent for secondary use
  - Consent is adaptable over time for new components
  - Future option to join supplementary/complementary studies
- Single IRB for PMI-CP
- Return of results and access to data
  - Aggregate results should be available to all participants
  - Individuals may set preferences for return information
Information Flow In

Direct Volunteers

HPO Volunteers

Self-report Measures

mHealth Data

Consent

EHR Data

Baseline Exam

Biological Samples
Information Flow Out

Volunteers

Researchers

Individual Data

Individual Health Info

Ongoing Study Updates

Aggregated Results

Data
## Possible data sources for the PMI Cohort

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Example Data Provided</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self report measures</td>
<td>Diet, substance use, self-report of disease and symptoms (e.g., cognitive or mood assessment)</td>
</tr>
<tr>
<td>Structured clinical data (EHR)</td>
<td>ICD and CPD codes, medication history, laboratory results, vitals, encounter records</td>
</tr>
<tr>
<td>Unstructured clinical data (EHR)</td>
<td>Narrative documents, images, EKG and EEG waveform data</td>
</tr>
<tr>
<td>Biospecimens</td>
<td>Blood sample, microbiome, nail and hair for environmental exposures over time</td>
</tr>
<tr>
<td>mHealth and sensor data</td>
<td>Passively-collected data (e.g., location, movement, social connections), wearable sensor data (activity, calories expended, hours and quality of sleep, time sedentary).</td>
</tr>
<tr>
<td>Healthcare claims data</td>
<td>Billing codes as received by public and private payors, outpatient pharmacy dispensing</td>
</tr>
<tr>
<td>Geospatial and environmental data</td>
<td>Weather, air quality, environmental pollutant levels, food deserts, walkability, population density, climate change</td>
</tr>
<tr>
<td>Other data</td>
<td>Social networking e.g., Twitter feeds, over-the-counter medication purchases</td>
</tr>
</tbody>
</table>
Initial Core Data Set

- Centrally collected and stored in a Coordinating Center
- Align with other data sets when possible
- Leverage existing data standards and common data models when possible

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Data Provided</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self report measures</td>
<td>Diet, substance use, self-report of disease and symptoms (e.g., cognitive or mood assessment)</td>
</tr>
<tr>
<td>Baseline health exam</td>
<td>Vitals (e.g., pulse, blood pressure, height, weight), medical history, physical exam</td>
</tr>
<tr>
<td>Structured clinical data (EHR)</td>
<td>ICD and CPT codes, medication history, select laboratory results, vitals, encounter records</td>
</tr>
<tr>
<td>Biospecimens</td>
<td>Blood sample</td>
</tr>
<tr>
<td>mHealth data</td>
<td>Passively-collected data (e.g., location, movement, social connections) from smartphones, wearable sensor data (activity, hours and quality of sleep, time sedentary).</td>
</tr>
</tbody>
</table>
Data Flow Between Coordinating Center (CC) and Participant Sites

- Core Data
- Direct Volunteers
- HPOs
- Simpler queries matching core data handled by CC (fast, scalable)
- Select subset validation/data cleaning
- Over time, core data grows
Biospecimen Collections

- PMI-CP would collect new biospecimens
  - Anticipate what future uses may be
  - Collect initially from everyone and at subsequent intervals as determined by use cases
  - Start with blood, but should accommodate samples for exposure studies, metabolites, microbiome, etc.
- Quickly establish a central PMI-CP biobank
- Maintain CLIA-compliance in specimen collection and testing where possible
Policy for the PMI-CP

- Policy needs for PMI-CP:
  - Single Institutional Review Board (IRB)
  - Privacy and security
    - Standards for data security
    - Safeguards against unintended data release
    - Penalties for unauthorized re-identifcation
  - Share results and provide access to data
    - Support broad consent in Common Rule NPRM
    - Clarify CLIA and HIPAA
- Special considerations for certain populations
PMI-CP Governance

- Governance structure
  - PMI-CP director
  - Independent Advisory Board
  - Executive Committee
  - Steering Committee with five subcommittees
    - Return of results and information
    - Data
    - Biobanking
    - Resource Access
    - Security
- Maintain interagency coordination
Next Steps

Principles:
- Utilize innovative ways to implement the cohort
- Stay flexible, nimble, cutting edge

Priorities:
- Act quickly to bring in a PMI Director & to “staff up”
- Quickly build infrastructure to support enrollment
  - Communications & engagement
  - Single IRB and consent
  - Data storage & acquisition infrastructure
  - Biobank
- Begin enrollment ASAP
Thank you!

- Working Group Members
- Workshop participants and RFI respondents
- NIH workshop planning teams
- Workshop hosts Vanderbilt and Intel
- Gwynne Jenkins and NIH Staff