From Research to Practice: New Models for Data-sharing and Collaboration to Improve Health and Healthcare

Joe Selby, MD, MPH, Executive Director, PCORI
Francis Collins, MD, PhD, Director, National Institutes of Health
Philip Bourne, PhD, Associate Director for Data Science, NIH

Moderator: Dwayne Spradlin, CEO Health Data Consortium
May 28, 2014
<table>
<thead>
<tr>
<th>Time</th>
<th>Agenda Item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:00 – 1:10 p.m.</td>
<td>Welcome</td>
</tr>
<tr>
<td>1:10 – 1:20 p.m.</td>
<td>Dr. Joe Selby, Executive Director, PCORI</td>
</tr>
<tr>
<td>1:20 – 1:30 p.m.</td>
<td>Dr. Francis Collins, Director, NIH</td>
</tr>
<tr>
<td>1:30 – 1:40 p.m.</td>
<td>Dr. Philip Bourne, Associate Director for Data Science, NIH</td>
</tr>
<tr>
<td>1:40 – 1:55 p.m.</td>
<td>Question and Answer Session</td>
</tr>
<tr>
<td>1:55 – 2:00 p.m.</td>
<td>Wrap Up and Conclusion</td>
</tr>
</tbody>
</table>
Questions may be submitted at any time

1. Click in the Q&A box on the right side of your screen, type your question into the dialog box, click Send button
2. You can also submit questions via twitter at @hdconsortium

Reminder: for audio, Dial 866-640-4044 - Entry Code: 416641#
Need help? Press *0 on phone to reach the operator
Joe Selby, MD, MPH
Executive Director
PCORI
PCORnet: Harnessing Real-World Health Data in Patient-Centered Research

Joe Selby, MD MPH, Executive Director
PCORI
PCORI’s Mission

PCORI helps people make informed health care decisions, and improves health care delivery and outcomes, by producing and promoting high integrity, evidence-based information that comes from research guided by patients, caregivers and the broader health care community.
PCORI’s Strategic Goals…

- Influence Research Funded by Others
- Increase Quantity, Quality and Timeliness of Research Information
- Speed the Implementation and Use of Evidence
...Set the Stage for PCORNet

- Improve the nation’s capacity to conduct clinical research more efficiently, by creating a large, highly representative, national patient-centered clinical research network with a focus on conducting CER – both randomized and observational.

- Support a learning US healthcare system, which would allow for large-scale research to be conducted with enhanced accuracy and efficiency within real-world care delivery systems.
PCORnet – Toward a Learning Healthcare System

PCORnet STEERING COMMITTEE
- Each Clinical Data Research Network
- Each Patient-Powered Research Network
- Patients
- HHS agencies:
  - NIH
  - FDA
  - AHRQ
  - CDC
  - CMS
  - ONC
  - ASPE
- Medical product and device manufacturers
- PCORI and Coordinating Center

COORDINATING CENTER

11 CLINICAL DATA RESEARCH NETWORKS

18 PATIENT-POWERED RESEARCH NETWORKS

PROJECT MANAGEMENT OFFICE

TASK FORCES
- GOVERNANCE
- DATA PRIVACY
- ETHICS & REGULATORY
- DATA STANDARDS & SECURITY
- HEALTH SYSTEMS INTERACTIONS
- PATIENT & CONSUMER ENGAGEMENT
- PATIENT GENERATED OUTCOMES
- CLINICAL TRIALS
- RARE DISEASES
- BIOREPOSITORIES
- OBESITY
Geographic Coverage of PPRNs and CDRNs
PCORnet Goals for Phase I

By 18 Months:

- Each CDRN will have a defined set of standardized clinical data that is fully inter-operable with data from other CDRNs; each PPRN will also have a standard database with varying amounts of clinical and patient-generated data.
- PCORnet will have clear policies on decision-making, uses of data, collaboration and knowledge sharing, data sharing, data privacy and security.
- Within each participating CDRN, patients, clinicians and health systems will be actively engaged in governance and use of the network and its data.
- Both CDRNs and PPRNs will have capacity to participate in both large observational studies and pragmatic (simple) randomized clinical trials.
- Networks will demonstrate a readiness to collaborate with researchers from outside PCORnet.
Francis Collins, MD, PhD
Director
NIH
NIH: Data Sharing Challenges and Solutions
Francis S. Collins, M.D., Ph.D.
Director, National Institutes of Health
From Research to Practice: New Models for Data Sharing and Collaboration to Improve Health and Healthcare
May 28, 2014
Value of Data Sharing

- Increases return on investment
- Facilitates additional research
- Helps to validate findings
- Promotes transparency

Many ongoing efforts to increase and facilitate data sharing

- Big Data to Knowledge (BD2K)
- Plan for increasing public access to data
Explosion of Big Data
*By Daily Users of NCBI*

- Daily Page Views: 28 Million
- Daily Users: ~4 Million
- Daily Downloads: 35 Terabytes
- Peak Hits: 7000 Per Second
Data Sharing Challenges and Solutions

- Genomic Data Sharing
- Clinical Data Sharing
- Human Subjects Protection
Data Sharing Challenges and Solutions

- Genomic Data Sharing
- Clinical Data Sharing
- Human Subjects Protection
Cost of Sequencing a Human Genome
September 2001–January 2014

4,008
NIH Genomic Data Sharing (GDS) Policy

- Expands expectations to share genomic data under the current NIH Genome-Wide Association Studies (GWAS) Policy to large-scale non-human and human genomic data
- Ensures the broad, responsible sharing of genomic research data
  - Responsibilities of investigators submitting data
    - Provide data sharing plan to NIH with grant application
    - Submit data in a timely manner
    - For human data, obtain consent for data to be used for future research purposes and shared broadly and submit Institutional Certification
  - Responsibilities of investigators accessing and using data
    - Terms and conditions for research use of controlled-access data
    - Conditions for use of unrestricted-access data
- Final will be implemented in January 2015
More to come?
Genomic Sequencing in the Clinic

- Authorized Platform: Illumina’s MiSeqDx
- FDA cleared two CF tests that use the Illumina platform
  - Panel of 139 mutations
  - Sequencing assay
- Paves the way for more genomic technologies to gain regulatory clearance
- Will allow for the development and use of new genome-based tests

MiSeq Benchtop Sequencer
(Credit: Illumina)
Data-sharing Challenges and Solutions

- Genomic Data Sharing
- Clinical Data Sharing
- Human Subjects Protection
Publication of Clinical Trial Results

- NIH-Funded trials published within 100 months of completion
- Less than 50% are published within 30 months of completion

Publication of Clinical Trial Results

NHLBI Clinical Trial Data: Time to Publication by End Point

Unadjusted rate ratio, 5.47 (95% CI, 3.74–7.98); P=0.001
Adjusted rate ratio, 2.11 (95% CI, 1.26–3.53); P=0.004

ClinicalTrials.gov: Public Benefits

- Enhance patient access to enrollment in clinical trials
- Prevent unnecessary or unwitting duplication of trials, especially those found to be unsafe
- Honor ethical obligation to participants (results inform science)
- Mitigate bias (non-publication of negative results)
- Inform future research and funding decisions
- Increase access to data about marketed products
- Facilitate use of findings to improve health

All contribute to public trust in clinical research
Data Sharing Challenges and Solutions

- Genomic Data Sharing
- Clinical Data Sharing
- Human Subjects Protection
Revisions to the Common Rule

Rationale for the reforms: human subjects research is changing

- Growth in research volume
- Increase in multi-site studies
- Increase in health services and social science research
- New technologies: e.g., genomics, imaging, informatics
- Increased role of private sector
- Increased sharing of specimens and data

The nature and volume of potential research data is one key rationale for reforms
Common Rule Reforms – July 2011 ANPRM

**Enhancing Protections**
- Require consent for research with biospecimens/data
- Enhance data security and information protection standards
- Extend protections to all research conducted at federally-funded institutions

**Reducing Burden**
- Promote use of broad consent for future research with biospecimens/data
- Broaden exemptions for low risk research
- Eliminate redundant IRB reviews and reduce impact of IRB reviews
Philip Bourne, PhD
Associate Director for Data Science
NIH
Some Observations

- **Good News**
  - Data sharing offers unprecedented opportunities to improve healthcare
  - We have a plan
  - We are beginning to quantify the issues
  - We have some of the best data scientists in the world to work on the problems
Some Observations

Bad News
- Sustainability will not be possible without change
- OSTP have defined the *why* but not the *how*
- We do not know how the data we currently have are used
- It is difficult to estimate supply and demand

Good News
- Data sharing offers unprecedented opportunities to improve healthcare
- We have a plan
- We are beginning to quantify the issues
- We have some of the best data scientists in the world to work on the problems
We have identified 5 programmatic themes and associated deliverables …
The Biomedical Research Digital Enterprise

**Associate Director for Data Science**

**Scientific Data Council**

**External Advisory Board**

**Programmatic Theme**

- Sustainability
- Education
- Innovation
- Process
- Collaboration

**Deliverable**

**Commons**
- Cloud – Data & Compute
- Search
- Security
- Reproducibility Standards
- App Store

**Training Center**
- Coordinate
- Hands-on
- Syllabus
- MOOCs

**BD2K**
- Community
- Centers
- Training Grants
- Catalogs
- Standards
- Analysis

**Modified Review**
- Data
- Resource Support
- Metrics
- Best Practices
- Evaluation
- Portfolio Analysis

**Communication**
- IC’s
- To Researchers
- Federal Agencies
- International Partners
- Computer Scientists

**Example Features**

- Cloud – Data & Compute
- Search
- Security
- Reproducibility Standards
- App Store
- Coordinate
- Hands-on
- Syllabus
- MOOCs
- Community
- Centers
- Training Grants
- Catalogs
- Standards
- Analysis
- Data Resource Support
- Metrics
- Best Practices
- Evaluation
- Portfolio Analysis
- IC’s
- To Researchers
- Federal Agencies
- International Partners
- Computer Scientists
The Biomedical Research Digital Enterprise

Programmatic Theme

Sustainability

Education

Innovation

Process

Collaboration

Deliverable

Commons
- Cloud – Data & Compute
- Search
- Security
- Reproducibility Standards
- App Store

Training Center
- Coordinate
- Hands-on
- Syllabus
- MOOCs

BD2K
- Community
- Centers
- Training Grants
- Catalogs
- Standards
- Analysis

Modified Review
- Data Resource Support
- Metrics
- Best Practices
- Evaluation
- Portfolio Analysis

Communication
- IC’s
- To Researchers
- Federal Agencies
- International Partners
- Computer Scientists

Example Features
The Biomedical Research Digital Enterprise

Programmatic Theme

- Sustainability
- Education
- Innovation
- Process
- Collaboration

Deliverable

- Commons
  - Cloud – Data & Compute
  - Search
  - Security
  - Reproducibility Standards
  - App Store
- Training Center
  - Coordinate Hands-on
  - Syllabus
  - MOOCs
- BD2K
  - Community Centers
  - Training Grants
  - Catalogs
  - Standards Analysis
- Modified Review
  - Data Resource Support
  - Metrics
  - Best Practices
  - Evaluation
  - Portfolio Analysis
- Communication
  - IC’s
  - To Researchers
  - Federal Agencies
  - International Partners
  - Computer Scientists
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

Data
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

The Why:

Data  Data Sharing Plans
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

The Why: Data Sharing Plans

The How:
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

Data

The Why:  
Data Sharing Plans

The How:

The End Game:
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

*The Why:*
Data Sharing Plans

*The How:*

*The End Game:*
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

The Why: Data Sharing Plans

The How:
- Sustainable Storage
- Quality
- Scientific Discovery
- Usability
- Security/Privacy

The End Game:
- Knowledge
- Metrics/Standards
- Sustainable Storage

Core Facilities/HS Centers

Clinical/Patient

Data
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

The Why: Data Sharing Plans

- The Long Tail
- Core Facilities/HS Centers
- Clinical /Patient

NIH Awardees

The How:

- Scientific Discovery
- Usability
- Quality
- Security/Privacy
- Metrics/Standards
- Sustainable Storage

The End Game:
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

The Why: Data Sharing Plans

- NIH Awardees
- Rest of Academia

The How:

- Government
- Private Sector

The End Game:

- Scientific Discovery
- Knowledge
- Usability
- Quality
- Security/Privacy
- Metrics/Standards
- Sustainable Storage

Data

The Long Tail

Core Facilities/HS Centers

Clinical /Patient

NIH

Rest of Academia

Private Sector

Government

Knowledge

Usability

Quality

Security/Privacy

Metrics/Standards

Sustainable Storage
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

**The Why:**
- Data Sharing Plans

**NIH Awardees**
- Software Index

**Rest of Academia**
- Data Discovery Index

**Government**
- Standards

**Private Sector**
- BD2K Centers

**The End Game:**
- Scientific Discovery
- Usability
- Quality
- Security/Privacy
- Metrics/Standards
- Sustainable Storage

Core Facilities/HS Centers
Clinical /Patient

The Long Tail
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

The Why: Data Sharing Plans

The How: Data Discovery, Index, Standards

The Commons

NIH Awardees
Rest of Academia

Government
Private Sector

Data Sustainable Storage
Quality
Scientific Discovery
Knowledge
Usability
Security/Privacy
Metrics/Standards
Sustainable Storage

The End Game:

NIH Awardees
Rest of Academia

Clinical /Patient

Core Facilities/HS Centers

Private Sector

BD2K Centers

Software Index

Standards

The Why:

The How:

The End Game:
The Power of the Commons

Commons == Extramural NCBI == Research Object Sandbox == Collaboratory

The Why:
Data Sharing Plans

The How:
- NIH Awardees
- Rest of Academia
- Government
- Private Sector

Cloud, Research Objects, Business Models

Software Index
Data Discovery Index
BD2K Centers

The End Game:
- Scientific Discovery
- Knowledge
- Usability
- Quality
- Security/Privacy
- Metrics/Standards
- Sustainable Storage

The Long Tail

Core Facilities/HS Centers

Clinical /Patient
What Will the Commons Accomplish?

- **Community Building** - support sharing, accessibility, and discoverability of biomedical data and analytical tools
- **Enable Innovation** - data resources co-located with advanced computing resources
- **Provide cost effectiveness** – through economies of scale, new business models, including public private partnerships
- **Provide opportunities for interagency and international cooperation**
BD2K will Empower the Commons

- Data discovery index
- Data/metadata standards
- Software index and software development
- Training centers and grants
- Centers engaged in advanced biomedical data science for the community at large
NIH…
Turning Discovery Into Health
Questions may be submitted at any time

To submit a question:

1. Click in the Q&A box on the right side of your screen, type your question into the dialog box, click Send button
2. You can also submit questions via twitter at @hdconsortium