PCORI's Policy for Data Management and Data Sharing

Jason Gerson

Senior Program Officer, PCORI

Research Data Management Conference

New York, NY

December 2019

About PCORI

- PCORI funds comparative clinical effectiveness research (CER)
- PCORI helps people make informed healthcare decisions, and improves healthcare delivery and outcomes, by producing and promoting high-integrity, evidence-based information that comes from research guided by patients, caregivers, and the broader healthcare community.
- Speed the dissemination and implementation and use of CER evidence
- Influence research funded by others to be more patient-centered

PCORI's Open Science Initiatives

Process for Peer Review of Primary Research and Public Release of Research Findings

(Approved by the Board February 2015)

Public Access to Journal Articles Presenting Findings from PCORI-Funded Research Policy

(Approved by the Board April 2016)

Policy for Data Management and Data Sharing

(Approved by the Board September 2018)

Guiding Values & Overarching Purposes

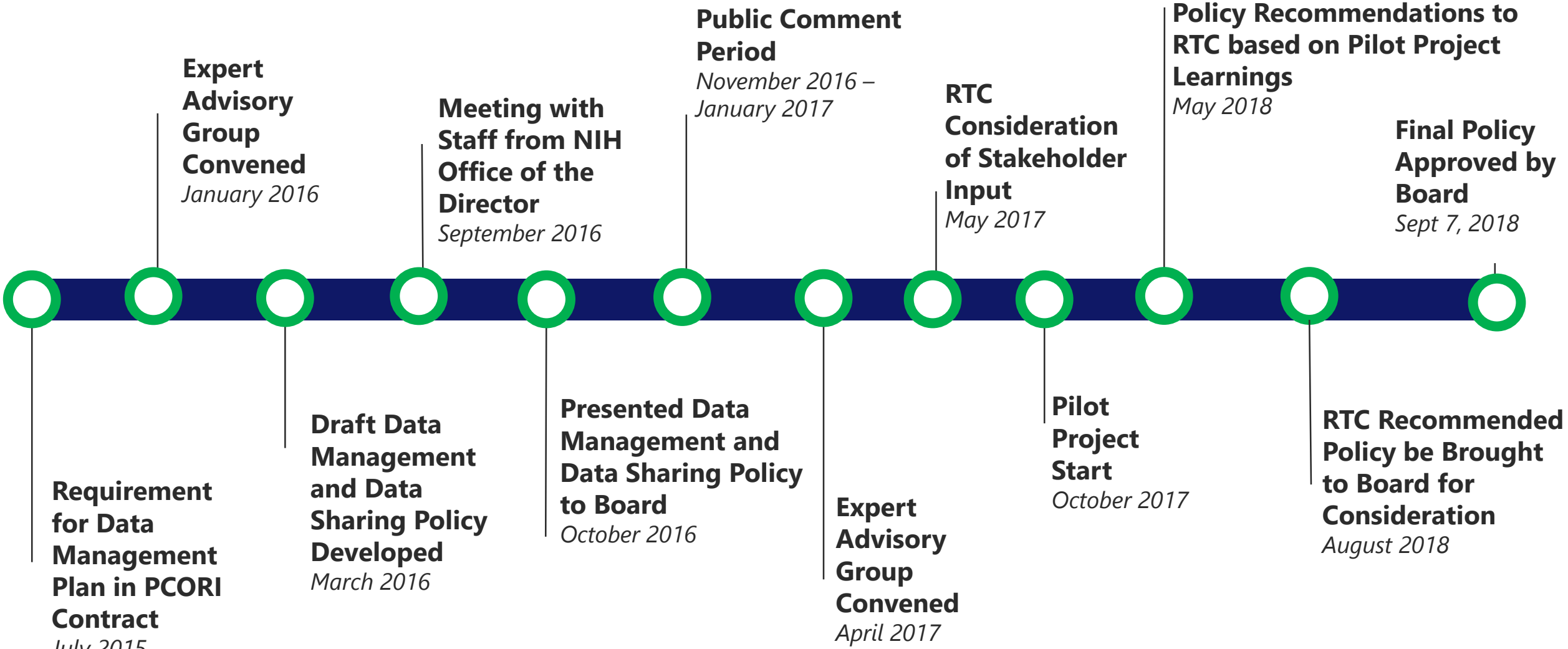
Guiding Values:

- Maximize the utility and usability of data collected in research projects that PCORI funds
- Encourage scientifically rigorous secondary use of clinical research data to foster scientific advances that will ultimately improve clinical care and patient outcomes

Overarching Purposes:

- Facilitate reproduction of original analyses to increase the integrity of PCORI-funded research findings
- Enable conduct of additional research analyses using data from PCORI-funded research projects

Timeline of Policy



Policy Development: Learning By Listening and Doing

- We gathered input from a variety of sources and through a number of activities, including:
  - Expert advisory group
  - Public comment period
  - Pilot project
  - Peer funders and regulatory agencies
  - PCORI's Research Transformation Committee (subcommittee of PCORI Board of Governors)

Features of the Policy

- Specifies data and data documentation to be shared
- Articulates expectations for data management and data sharing to Awardees
- Provides funding to support Awardees' time and effort to prepare data
- Specifies when data would be made available for third-party requests
- Describes third-party data request review process
- Articulates criteria and review process regarding exemptions from Policy requirements

Features of the Policy: Privacy and Ethical Protections

- Thorough scientific vetting of requests for data
- De-identified data only (in accordance with the HIPAA Privacy Rule (45 C.F.R. §164.514(b))
- Prohibition against reidentification of individuals
- Prohibition against redistribution of data
- Data from participants whose informed consent permits data collected as part of the research project to be de-identified and to be used for secondary research purposes and shared with researchers not affiliated with the institution

Data Deposition: Overview

- Deidentified data only in accordance with the HIPAA Privacy Rule (45 C.F.R. § 164.514(b))
- Full Data Package: Analyzable Data Set, Full Protocol, metadata, data dictionary, full statistical analysis plan, and analytic code
  - Share data from participants whose informed consent permits data to be used for secondary research purposes and shared with researchers not affiliated with the Awardee's institution
- Awardees depositing full data package (or applicable data elements) will work with PCORI-designated repository(ies) to curate it
  - Data will be hosted by designated repository(ies), not PCORI
- Awardees will enter into a Data Contributor Agreement (DCA) with the repository. DCA governs the data deposition and establishes the Awardee's rights and obligations

Data Deposition: Requirements for Research Awardees

Targeted and Pragmatic Clinical Studies Funding Announcements	PCORnet Funding Announcements	Broad Funding Announcements
• Deposit full data package (or required data elements, as applicable) in a PCORI-designated repository	• Deposit applicable data elements, such as the full protocol, analytic code used to query PCORnet data, and aggregate level datasets in a PCORI-designated repository	• Maintain full data package for 7 years • PCORI may notify Awardee of its intent to provide funds for the deposition of the full data package in a PCORI-designated repository

Timeframe for Data Availability

- The full data package will be made available for third-party requests only when the

PCORI Final Research Report is made available on PCORI's website

- OR -

One of the research project's primary papers is published in a peer-reviewed journal

<u>WHICHEVER COMES FIRST</u>

Data Requests: Review Process

- Review will help ensure data request has scientific merit by evaluating that:
  - Scientific purpose is clearly described
  - Data requested will be used to develop or contribute to generalizable knowledge to inform science, medicine and/or public health
  - Proposed research can be reasonably addressed using the requested data
  - Requestor team has the appropriate expertise to conduct the proposed research
- Approved requestors will enter into a Data Use Agreement (DUA) with a PCORI-designated repository. DUA specifies the terms and conditions of data use, as well as the responsibilities and obligations of data requestors.

Data Requests: Independent Review Committee

- All data requests will be reviewed by an independent committee. Committee will be comprised of 5 individuals:
  - Representative from the PCORI-designated repository
  - Data scientist
  - Clinical researcher with expertise germane to the data request
  - PCORI staff member
  - Patient representative
- A member of the Awardee research team that generated the requested data will be invited to attend the review as a non-voting participant

Rationale for Data Management Plans (DMPs) Standards

- PCORI's Methodology Committee has developed Methodology Standards (MS) -- specific recommendations for researchers that designate the minimal requirements for following PCOR best practices
  - MS focus on selected methodologies and issues that reflect areas where there are either substantial deficiencies or inconsistencies in how available methods are applied in practice.
- The Methodology Standard (CC-3) for Data Integrity and Rigorous Analyses (IR) was initially silent about data management.
- Good data management is fundamental to ensuring the scientific integrity of clinical research
  - Salutary effect for open science: Ensuring that good data management plans are in place at the outset of a study will facilitate data sharing at its conclusion.
- Many organizations (incl. most federal funders) now require DMPs, and others that have articulated "best practices." Including a Standard re: DMPs was, therefore, non-controversial.
- We added IR-7: **In your proposal, specify a data management plan that addresses, at a minimum, the following elements: collecting data, organizing data, handling data, describing data, preserving data, and sharing data.**

Elements of DMP

- How the data will be obtained or collected
- How the individual data items will be described
- How the data will be safely organized, stored and preserved
- Who will have access to the data set
- Who will have permission to make edits or changes to the data
- What mechanisms you will use at the end of your project to share the data.

- The DMP is a living document and should be reviewed periodically (or any time your research plans change) to ensure that it remains suitable for the research being conducted.
- Incorporated into Methodology Standards in 2017

Some Reflections & Key Takeaways

- Be clear-eyed about the fact that data sharing is still in a relatively nascent phase in clinical research – policy will be *de facto* norm changing

- Implementation of data sharing for clinical research requires careful deliberation about the details – cannot simply direct awardees to deposit "something, somewhere" and declare victory
  - It's not just the dataset… it's the whole data package
  - Data curation and ability of repository to work with awardees is critical

- Be clear-eyed about the fact that, as a single funder, we are only one part of the ecosystem. Sharing of clinical research data requires that other funders and sponsors, academic institutions, publishers, health systems, among others, are working together towards aligning incentives to support data sharing.