Yale University Open Data Access (YODA) Project
Public Comment Response to PCORI'S Data Access and Data Sharing Policy

The Yale University Open Data Access (YODA) Project at the Yale-New Haven Center for Outcomes Research and Evaluation (CORE) fully supports and applauds the development of data sharing strategies by PCORI. Please find our specific comments below.

1. **What should be the retention period for how long PCORI research awardees must retain the full data package for sharing?**

   Ideally, the full data package would be retained indefinitely. However, this notion is highly constrained by available resources. The YODA Project applauds PCORI's commitment to cover reasonable costs associated with maintaining and depositing the full data package in a PCORI suggested repository for a period of at least seven years. For those data packages that are not deposited into repositories, funds should be included within the original awards to cover the cost of study investigators to retain the data package for seven years. After seven years, or if original study investigators are no longer available, whichever comes first, PCORI should set aside funds to continue retaining the data package itself or to cover the costs of depositing the data package into a data repository that is consistent with applicable privacy, security and other legal requirements.

2. **What restrictions on use of PCORI research awardee data are important to include in the data use agreement executed by third parties requesting access to the data (e.g., data will not be re-identified, data will be used only for research and not commercial purposes, data will be used only for research in the same therapeutic area as the original research project through which the data were collected)?**

   Data Use Agreements are an integral tool to sharing data with third parties and should include the following stipulations and requirements:
   - Signature by the Principal Investigator
   - Signature by a representative of the Principal Investigator's primary affiliation
   - Certification that the requestor's credentials and all information provided are true and that there are no debarment issues
   - Certification that the requested data will not be used in pursuit of litigation or for commercial interests
   - Agreement that no distribution of the data to third parties or public posting of the data will be permitted. The requestor must protect the confidentiality of the data, and may not copy, retransmit or reuse the data in any manner other than for the purpose described in the data request.
   - Acknowledgement that violations may be subject to all available legal remedies and will be posted on a public website
   - Agreement that the Principal Investigator (or any other permitted data user) will not attempt to re-identify individuals within the data
   - Agreement that the scope of the analyses will be limited to the specific aims set out in the proposal; additional research objectives will require the submission of a new proposal or an amendment to the previously approved proposal
   - Agreement that any unexpected or serious safety findings must be reported to the original study investigators and PCORI immediately for further evaluation and reporting to appropriate health or regulatory authorities

Yale University Center for Outcomes Research and Evaluation (CORE)
1 Church Street, Suite 200, New Haven, CT 06510

- Agreement that all public dissemination of the findings will be required to include standardized language that reflects that the analyses were based on data made available via PCORI and the original study investigators
- Agreement that the investigator will provide PCORI and original study investigators with a copy of any abstract and manuscript generated from the data request at the time of submission to a scientific research meeting or a biomedical journal. Neither party will have the authority to approve, reject, or revise the investigator's work. Advanced notification may help to identify any relevant safety issues that should be reported to appropriate regulatory officials.
- Acknowledgement of an expiration date of one year after signing the DUA, which can be renewed for additional one-year periods for ongoing projects
- Acknowledgement that the investigator will retain the rights to access data for an additional five years for completed projects for the express purpose of verification of the investigator's publication and to respond to inquiries regarding the publication
- Agreement that any log-in credentials used to access data will be kept confidential

3. What qualifications/credentials should be required for a third party requesting access to PCORI research awardee data (e.g., education level, specific scientific expertise)?

Restrictions should not be placed on the education level or specific scientific expertise of the requestor as long as the following conditions are met:

- Scientific purpose is clearly described
- Data requested will be used to create or materially enhance generalizable scientific and/or medical knowledge to inform science and public health
- Proposed research can be reasonably addressed using the requested data
- Certification that the requested data will not be used in pursuit of litigation or for commercial interests

4. What documentation should a third party requester be required to provide when applying to access PCORI research awardee data (e.g., research question and lay protocol summary, full protocol and statistical analysis plan, institutional review board approval)?

The YODA Project suggests that the following information be required as part of the process to request data access:

- Principal Investigator's name, degree(s), SCOPUS ID (if available), primary affiliation, and contact information, including phone, mailing address, and email address
- Other Key Personnel members' names, degrees, and SCOPUS IDs (if available)
- Funding source and conflict of interest statement for the Principal Investigator and all Key Personnel
- Research Proposal, which includes project title, scientific abstract, brief project background and statement of project significance, specific aims, research methods, narrative summary, project timeline, dissemination plan, and bibliography
- Certification that the credentials and all information provided are true and that the requested data will not be used in pursuit of litigation or for commercial interests

5. Should PCORI establish data repository standards (e.g. security, curation standards)? If so, what should the minimum standards be for a repository to qualify as a PCORI suggested repository for PCORI research awardee data?

The YODA Project feels it would be helpful for PCORI to create data repository standards, as established standards will aid in identifying appropriate repositories for full data packages developed through PCORI-funded research. It may be useful for PCORI to conduct a review of standards that currently exist, including the criteria established by the Data Archiving and Networked Services Data Seal of Approval (https://assessment.datasealofapproval.org/documentation/).

6. **What is the appropriate model for informed consent that should be included in the policy?**

The YODA Project's only comment on informed consent is that it should include language that permits data collected as part of the study to be de-identified, used for future research purposes and shared broadly with researchers not affiliated with the institution conducting the study, which PCORI has proposed in its policy.

7. **Do you have any other comments? Specifically, we seek comment about any of the following sections: (I) Purpose, (II) Applicability, (III) Definitions, and (IV) Policy.**

(III) Definitions, Analyzable Data Set: The YODA Project believe that the proposal could be further strengthened if it were broadened to include collected datasets, not just the analyzable dataset, allowing researchers to use the dataset in its original format. In addition, all deidentified IPD associated with the conduct of the funded study, not just the deidentified IPD underlying any published results, should be shared. As investigators publish multiple articles, multiple data sets for the same study can be shared, creating version control issues and potential confusion.

(III) Definitions, Full Data Package: While the proposal clearly states that the full data package should include the full analyzable data set, full protocol (including initial version, final version, and all amendments), full statistical analysis plan (including all amendments and all documentation for additional work processes), and analytic code, the YODA Project believes this proposal could be strengthened by expanding the metadata even further. The metadata needed to understand and make use of data includes, but is not limited to, blank case report forms, data definitions and specifications, and clinical study reports.

(IV) Policy: Currently lacking from the proposal is specification of penalty(ies) for investigators who do not prepare or follow through with data management and data sharing plans. The YODA Project believes this proposal could be strengthened by explicitly defining penalty(ies), such as exclusion from consideration for PCORI-funded research for a 3-year period. PCORI will need to develop mechanisms for monitoring the data sharing activities of funded researchers. Also, investigators will be responsible for preparing data to be shared, which includes protecting patient privacy, resources will be required for best-practice deidentification (*http://dx.doi.org/10.1136/bmj.h1139; http://dx.doi.org/10.1136/bmj.c181*).