

# Onboarding Data and Resource Sharing Questionnaire

---

## Table of Contents

<b>Data and Resource Sharing Questionnaire .....</b>	<b>1</b>
<b>Data and Resource Sharing Checklist.....</b>	<b>6</b>

## Data and Resource Sharing Questionnaire

This questionnaire is a worksheet to guide Demonstration Projects in developing data sharing plans that meet program requirements (see below checklist). This questionnaire is to be used as part of the onboarding process and can be used for planning purposes by other researchers who need to share data.

Instructions/guidance are provided in italics. Please provide responses in the answer column.

Data Sharing Questionnaire	
1. Study information	
Question	Answer
What is the trial name and acronym?	
Who is completing this questionnaire?	
Date of questionnaire completion?	
Please provide a link to the trial's ClinicalTrials.gov registration.	

## Data and Resource Sharing Questionnaire for Plan Development Worksheet

### 2. Data elements and sharing

*NIH Pragmatic Trials Collaboratory investigators will each **share, at a minimum, a final research dataset** upon which the accepted primary pragmatic trial publication is based (from the NIH Collaboratory Data Sharing Policy; see Data Sharing Information Document for additional information from NIH Pragmatic Trials Collaboratory, NIH, and medical journal data sharing policies).*

2a. Please describe all data collected/used for this study. Select all that apply and fill out each column as applicable.

Data	Y/N	If Y, brief description of data	Identifiable? If so, what IDs?	Can it be shared without restriction?	Can it be shared with restriction?	Describe restrictions (e.g., IDs stripped, aggregated info only, etc.) or reason data cannot be shared
• Individual Level Data						
• Primary data collection through informed consent						
• Primary data collection through waiver of informed consent						
• Secondary data use – data collected by researchers of an earlier study						
• Secondary data use -- administrative data obtained from a covered entity (e.g., claims and assessment data from CMS; electronic health records from health care providers, etc.)						
• Other						
• Provider Level Data						
• Other Data (e.g., state policy, market level, Census)						

## Data and Resource Sharing Questionnaire for Plan Development Worksheet

2b. Please describe the analytic dataset that will be released	
Will individuals be identifiable? ____ Yes ____ No ____ N/A	Comments/explanation:
Level of dataset: ____ Individual ____ Provider ____ Other	Brief description of dataset:
If not identifiable, can individuals be differentiated? (e.g., includes a study-generated ID so that multiple events/observations can be attributed to a unique study participant) ____ Yes ____ No	Comments/explanation:
Will providers be identifiable? ____ Yes ____ No ____ N/A	If not identifiable, can providers be differentiated? ____ Yes ____ No
Can the primary analyses be replicated using the released data? ____ Yes ____ No	If no, why not? (e.g., aggregated data; missing elements; etc.)
What value will the data have for other researchers?	

### 3. What precautions/risks need to be considered?

*The NIH Collaboratory Steering Committee recognizes that sharing data derived from clinical care in studies performed in partnership with healthcare systems may, under some situations, **require precautions in addition to those regarding patient confidentiality**, to protect specific interests of collaborating healthcare systems, facilities, or providers. Precautions such as allowing data sharing in more supervised or restricted settings, such as access to researchers who agree to limited pre-approved research goals, may be appropriate to address these needs (from the NIH Collaboratory Data Sharing Policy).*

Question	Answer
What precautions are needed other than those regarding patient confidentiality?	
Have your research partners expressed concerns about how the data will be shared (enclave, repository, etc.)?	
What are the risks to providers and health systems if a less restrictive mechanism is used? (See Data Sharing Information Document for examples from NIH Collaboratory Demonstration Projects.)	

## Data and Resource Sharing Questionnaire for Plan Development Worksheet

### 4. How will the data be shared?

*Consistent with NIH policy and guidance, NIH Pragmatic Trials Collaboratory investigators will choose the **least restrictive method for sharing of research data** that provides appropriate protection for participant privacy, health system privacy, and scientific integrity (from the NIH Collaboratory Data Sharing Policy).*

Question	Answer
What is the least restrictive mechanism you can use for sharing data? (See Data Sharing Information Document for details about these mechanisms.) <ul style="list-style-type: none"><li>• Public archive (least restrictive)</li><li>• Public enclave</li><li>• Private archive</li><li>• Private enclave (most restrictive)</li></ul>	
What specific platform will be used? (See Data Sharing Information Document for example data sharing platforms.)	

### 5. Preparing for data sharing

Question	Answer
When will you share data? Prior to or after publication?	
Please write a draft data sharing statement. (See Data Sharing Information Document for example statements.)	
Do you foresee any obstacles regarding data and resource sharing?	

## Data and Resource Sharing Questionnaire for Plan Development Worksheet

### 6. What resources will be shared?

As part of the NIH Pragmatic Trials Collaboratory's commitment to sharing, all Demonstration Projects are expected to share data **and resources, such as protocols, phenotypes, videos, training materials, consent documents, and recruitment materials**. We recommend that elements of a final data sharing package include the items listed below. If an element will not be included in the data sharing package, please provide a brief explanation for the omission. Resources can be housed in the [NIH Collaboratory Knowledge Repository](#) (KR), on a repository (i.e., GitHub), or on a study website. We will link to the materials from the Living Textbook. To request posting of materials to the KR, contact [nih-collaboratory@dm.duke.edu](mailto:nih-collaboratory@dm.duke.edu).

Item	Will you publish? Yes, No, N/A If No, justify	Where publish (mark all that apply)		When publish (mark all that apply)		
		NIH KR	Other (specify)	Per manuscript*	Start of study	End of study
Final version of protocol						
Consent documents/process						
Computable phenotypes for outcome measures						
Computable phenotypes for inclusion/exclusion criteria						
Code for generating variables in the analytic dataset from standard sources						
Study questionnaires						
Annotated data collection forms						
Data dictionary (proc contents) for public use dataset						
Data dictionary (proc contents) for all data used in study with annotation regarding limitations on sharing each element						
Code for generating the tables present in a particular manuscript*						
Instructions on how to obtain data that were unable to be released (e.g., CMS data files)†						
Other						

\*For example, PROVEN developed a process of submitting supplemental material for each manuscript published. They store the information in Brown's Digital Repository with a manuscript-specific URL that is published within the manuscript. They include the code that generated the manuscript's tables.

†For example, the PROVEN team refers the reader to [www.resdac.org](http://www.resdac.org) for the use of CMS data files and lets them know the file types and years used for its study since they cannot release those data.

Prepared by: NIH Collaboratory Coordinating Center

Version: April 7, 2022

# Data and Resource Sharing Checklist

---

## Background

All NIH Pragmatic Trials Collaboratory Demonstration Projects will be expected to review this checklist as part of the onboarding process so they understand what will be expected. They will complete the checklist at closeout.

As part of the NIH Pragmatic Trials Collaboratory's commitment to sharing, all of its Demonstration Projects are expected to share data and resources, such as protocols, phenotypes, videos, training materials, consent documents, and recruitment materials. We recommend that elements of a final data sharing package include the items listed in the checklist below. If an element will not be included in the data sharing package, please provide a brief explanation for the omission. Resources can be housed in the [NIH Collaboratory Knowledge Repository](#) (KR), on a repository (i.e., GitHub), or on a study website. We will link to the materials from the Living Textbook on each project's Demonstration Project page and through a separate Data and Resource Sharing section. To request posting of materials to the KR, contact [nih-collaboratory@dm.duke.edu](mailto:nih-collaboratory@dm.duke.edu).

Note: There will **not** be a dedicated space on the NIH Collaboratory website for posting analytic datasets; rather, we will post a hyperlink to the data sharing repository chosen by each project. In the Data Sharing Information Document, the EHR Core provides a partial list of existing data sharing platforms. The accompanying Data Sharing Information Document also contains information on data sharing requirements for the NIH Pragmatic Trials Collaboratory, NIH, and medical journals; information on data sharing mechanisms and platforms; and examples from Demonstration Projects.

## Data and Resource Sharing Checklist for Plan Development – Part 1

Data and Resource Sharing Checklist
<b>1. Study information</b>
Trial name and acronym:
Checklist completed by:
Date:
Link to ClinicalTrials.gov registration:
Link to study website:

## Data and Resource Sharing Checklist for Plan Development – Part 2

Data and Resource Sharing Checklist		
2. Resource location		
Item	Provide hyperlink or indicate if item will be stored in the KR	If item will not be shared, please provide a brief explanation for the omission
<b>Publications</b>		
Link to protocol paper		
Link to main outcome paper		
Link to other study-related publications		
<b>Study tools</b>		
Final version of the protocol, including summary of changes		
Consent documents or consent process		
Computable phenotypes for outcome measures		
Computable phenotypes for the inclusion/exclusion criteria		
Code for generating variables in the analytic dataset from standard sources		
<b>Datasets and documentation</b>		
Annotated data collection forms		
Link to public use dataset		
Data dictionary (proc contents) for public use dataset		
<b>Other resources</b>		