

NIH Collaboratory

Data Sharing Policy and

Considerations

Adrian Hernandez, MD, MHS
Duke University School of Medicine
Duke Clinical Research Institute

Wendy Weber, ND, PhD, MPH
National Center for Complementary
and Integrative Health (NCCIH)



Data Sharing Mandate (from 2017 RFA):

“The HCS Research Collaboratory program encourages sharing of resources with broad availability **of policies, practices, materials, and tools** to facilitate **collaboration, reuse, and replication**. In addition, the HCS Research Collaboratory program **requires sharing of study data** from Demonstration Projects in a timely manner with appropriate privacy and confidentiality protections, in accordance with the Data Sharing Policy developed by the HCS Research Collaboratory.”

Data Sharing Policy and Considerations are included in the meeting materials.

Why?

- Public health
- Maximizing investment
- Accelerate learning
- Foster collaboration (pay it forward)

Ultimately... transparency, reproducibility, secondary use of medical research, are good for society

Data Sharing and Embedded Research

Gregory E. Simon, MD, MPH; Gloria Coronado, PhD; Lynn L. DeBar, PhD, MPH; Laura M. Dember, MD; Beverly B. Green, MD, MPH; Susan S. Huang, MD, MPH; Jeffrey G. Jarvik, MD, MPH; Vincent Mor, PhD; Joakim Ramsberg, PhD; Edward J. Septimus, MD; Karen L. Staman, MS; Miguel A. Vazquez, MD; William M. Vollmer, PhD; Douglas Zatzick, MD; Adrian F. Hernandez, MD, MHS; and Richard Platt MD, MS

- The ethical responsibility to share data generated by publicly funded research must be balanced against the need to protect *patient privacy* and scientific *integrity*
- Data sharing policies must not dissuade health system participation



NIH Data Sharing Policy and Implementation Guidance

- The privacy of participants should be safeguarded.
- Data should be made as widely and freely available as possible.
- Data should be shared no later than the acceptance for publication of the main study findings.
- Initial investigators may benefit from first and continuing use of data, but not from prolonged exclusive use.

Principles

- Patient privacy: Essential obligation
- Health system privacy: Sometimes necessary

The Balance of Benefits & Risks





Three Documents Reflect Our Policy

1. Onboarding Data and Resource Sharing Questionnaire
2. Closeout Data and Resource Sharing Checklist
3. Informational Document for the Development of Data and Resource Sharing Plans

These three documents are in your booklet



1. Onboarding Questionnaire

Purpose: To help new projects think about their data sharing plan.

- Data elements and data sharing
- What precautions/risks need to be considered
- How will the data be shared
- Preparing for data sharing
- What resources will be shared

1. Excerpt from Onboarding Questionnaire

2. Data Elements and Sharing

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 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						



2. Closeout Data Sharing and Resource Checklist

- This what we expect to be shared
- If data or a resource cannot be shared, we want to know why not
- All links and resources will be shared on the Living Textbook

Living Textbook Mockup

Data and Resource Sharing

As part of the NIH Collaboratory's commitment to knowledge sharing, this page includes links and resources for all completed Demonstration Projects.

Protocols

[ABATE Protocol](#)

[LIRE Protocol](#)

[PPACT Protocol](#)

[STOP CRC Protocol](#)

[TiME Protocol](#)



Datasets and Documentation

[ABATE Dataset](#)

[LIRE Dataset](#)

[PPACT Dataset](#)

[STOP CRC Dataset](#)

[TiME Dataset](#)



Informed Consent Documents

[ABATE Informed Consent](#)

[LIRE Informed Consent](#)

[PPACT Informed Consent](#)

[STOP CRC Informed Consent](#)

[TiME Informed Consent](#)



Other Resources

[ABATE Resource](#)

[LIRE Resource](#)

[PPACT Resource](#)

[STOP CRC Resource](#)

[TiME Resource](#)



BY PROJECT

For a comprehensive list of resources by project, please visit each project's page:

[ABATE](#)

[LIRE](#)

[PPACT](#)

[STOP CRC](#)

[TiME](#)

2. Closeout Data Sharing and Resource Checklist (1 of 2)

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
Publications		
Link to protocol paper		
Link to main outcome paper		
Link to other study-related publications		
Study tools		
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		

2. Closeout Data Sharing and Resource Checklist (2 of 2)

Datasets and documentation		
Annotated data collection forms		
Link to public use dataset		
Data dictionary (proc contents) for public use dataset		
Other resources		

It's not all about the data

- Share policies, practices, materials, and tools
 - Key study documents
 - Publications
 - Study tools
 - Datasets and documentation

Remember: the checklist is in your binder

3. Informational Document

- Requirements from NIH Collaboratory, NIH, and medical journals
- Examples of how Demonstration Projects have shared data
- Data Sharing Mechanisms (public/private; archive/enclave)
- Data sharing platforms (Github, HCUP, Dryad, etc)
- Examples of data sharing statements

3. Includes explanation of data sharing mechanisms

- **Public archive** – send data to any user
- **Private archive** – send data to approved users
 - For any purpose
 - For approved purposes only (e.g. governed by DUA)
- **Public enclave** – any user may send questions to data
- **Private enclave** – some users may send questions to data
 - Automatically executed
 - Moderated (either queries or results)

3. Includes NIH Data Sharing Policy and Implementation Guidance

- The privacy of participants should be safeguarded.
- Data should be made as widely and freely available as possible.
- Data should be shared no later than the acceptance for publication of the main study findings.
- Initial investigators may benefit from first and continuing use of data, but not from prolonged exclusive use.

3. Includes Our Policy

- Share the final research data set
- Protect specific interests of health care systems partners
- Choose the least restrictive method for sharing
- Balance protection of participant privacy, health system privacy, and scientific integrity.
- Collaborate with NIH to implement this data sharing policy

Summary

- Data sharing is a societal good
- Like healthcare, health systems are.... complex
- Collaboratory can lead the way...
- What can we do to accelerate data sharing?



Thank You