

## Considerations for External Data Sharing Requirements

### 1. Research Data Sharing Requirements

This purpose of this guidance is to inform researchers on how to respond to individual (de-identified) participant-level research data sharing conditional mandates from publications, funding agencies, and consortiums.

The Mass General Brigham IRB fully supports data sharing but must balance the data sharing requirements with privacy and human subject considerations for each protocol on a case by case basis.

### 2.

#### Mass General Brigham IRB Role

- a. No review required for:
  - i. Sharing of methods and protocol-related documents
  - ii. Sharing of summary, aggregate or statistical analyses
- b. IRB review required for sharing individual participant-level research data.

*Please note, any external participant-level data storage or sharing will require a specific plan or protocol amendment if the external sharing has not already been covered by the initial protocol.*

### 3. Information required for IRB Review

- Why the data sharing is being requested. Attach data sharing requests, agreements, contracts, electronic communication (or any other supportive documentation.)
- Describe the mechanism for sharing and if it is a controlled or open access site.
- A LIST OF SPECIFIC data variables to be shared and actual number of individuals.
- If there are any data use restrictions based on the informed consent form or any other considerations, such as sensitive research.

- If there is an increased risk of re-identification based on small scale population or rare disease area.

#### 4. Mechanisms for sharing

##### Infrastructure

- i. Central repository – a site which provides access to data submitted from multiple sites or researchers. There are two types of repositories: controlled and open access. Controlled Access – is a restricted access process in which a scientist requests access to query the data and signs a data use agreement.

1. Examples of this include dbGAP, Vivli (only clinical trials are eligible for this site) or Harvard's Dataverse.

<https://dataverse.org/>

<https://vivli.org/>

*McKinney, B., et al., Extension of research data repository system to support direct compute access to biomedical datasets: enhancing Dataverse to support large datasets. Ann N Y Acad Sci, 2017. 1387(1): p. 95-104.*

*Bierer, B.E., et al., A Global, Neutral Platform for Sharing Trial Data. N Engl J Med, 2016. 374(25): p. 2411-3.*

- b. Open Access is an unrestricted access process in which anyone can request or download the data without any vetting of the use or scientist, or an agreement.

1. Examples of this include a foundation website, or NIH GEO (Gene Expression Omnibus).

##### Recommendations

- a. Many publications or funding agencies have specific data sharing requirements, sometimes open access is mandatory. Privacy, consent and other regulatory issues must be

considered before posting participant level data to central repositories, especially if the data is posted for unrestricted use.

- b. The preferred method for sharing is controlled access as previously described. Completely unrestricted public posting of data on an open access website can be approved under specific conditions, including, but not limited to: participants prospectively provided consent, and the individual participant-level data is 'limited,' so there is no risk of re-identification.

## 5. Alternative methods to central repository sharing

- a. Data Available Upon Request
  - i. Provide PI or other staff member contact information and state the data is "available upon request" – vetted through the PI. If a journal requests a contact who is NOT an investigator, we recommend a senior divisional representative at your site. Note that the IRB cannot be this contact. A data use agreement process may be executed through Research Management or the appropriate office for this scenario.
- b. PI-controlled/operated repository
  - i. The PI may set up a website or REDCap database behind the Mass General Brigham firewall. The PI may work with Mass General Brigham Research Information Security to create a site behind the Mass General Brigham firewall where the materials may be posted and an (electronic) data use agreement made available. This may be automated, if appropriate.
- c. The detailed plan to execute these methods must be reviewed and approved the IRB.

## 6. How to document a data sharing plan for protocol submission or external entity

- a. If you need specific language for a data sharing plan, consider customizing the following (based on the publication requirements) that has been used for NIH grants in the past:

<https://www.partners.org/Assets/Documents/Medical-Research/Clinical-Research/SVP-CAO-memo-NIH-GDS.pdf>

- We are committed to making resources and data from the proposed research available to other investigators in the research community. All data collected for this research will be obtained with IRB review/approval and a waiver of consent/informed consent of study participants to sharing of de-identified data. We will submit insert type of data to be shared as well as relevant associated data (e.g., phenotype and exposure data) to XXX data repository in a timely manner, as indicated by the XXX policy; typically, up to six months after data submission is initiated or at the time of acceptance of initial publication, whichever occurs first. We will also submit any information necessary to interpret the data, such as study protocols, data instruments and survey tools. The identities of research participants will not be disclosed to the repository. We will take appropriate steps to de-identify datasets according to the HIPAA privacy law. [https://privacyruleandresearch.nih.gov/pr\\_08.asp](https://privacyruleandresearch.nih.gov/pr_08.asp)
- Aggregate Data will/will not be available for submission/general research use.
- Insert a paragraph describing any data sharing restrictions: e.g., limited to cardiovascular disease or breast cancer; any restrictions described in the informed consent.