Preparing for NIH-Mandated Data Sharing by Partnering with Arcus

Children's Hospital of Philadelphia®
Department of Biomedical & Health Informatics

Presentation overview

- What's in the new NIH Data Management and Sharing policy?

- Responsibly managed internal data sharing via Arcus

- Research data management and data collection recommended practices

NIH Data Management and Sharing Policy

- Beginning on 2023-01-25, the new NIH policy for data management and sharing will go into effect
- This policy applies to all NIH funded research *that generates scientific data*
- It will affect any grant applications or contract proposals submitted on or after the effective date
- The policy has two requirements:
  - 1. All applications must include a Data Management and Sharing Plan
  - 2. All funded efforts must comply with the terms of their approved Data Management and Sharing Plan

Children's Hospital of Philadelphia®
Department of Biomedical & Health Informatics

Data Management and Sharing Policy: Basics

- Data should be made available no later than the time of publication or the end of the performance period, whichever comes first
- No specified time frame for how long data should be available (consult agreements & industry norms)
- "Sharing" does not mean "public"
  - Researchers should share to the extent permitted
  - Restricted data / Data Use Agreements (DUA) are OK
  - NIH may request/require changes to submitted sharing plans
- Enforcement: compliance with plans can affect funding for current and future awards
- Costs for data management and sharing can be included in budget requests

Internal Sharing through the Arcus Data Contribution Process

- The Arcus Archives are the canonical repository for research data at CHOP's Research Institute.

- Contributing your data to Arcus also means that Arcus will steward the data as part of its long-term preservation. The responsibilities involved will be documented in a Data Contribution Agreement (DCA) between Arcus and the PI. The DCA will also define access rights to the data, ensuring the contributor continued access and defining the parameters by which other CHOP researchers may access the data.

- Depositing research data in the Arcus Archives via the data contribution process allows researchers to gain practical experience fulfilling two of the principal requirements of the NIH Policy for Data Management & Sharing, by positioning Arcus as an internal repository where data can be securely preserved, curated, discovered, and disseminated in a manner that mirrors the best practices and desired aims of the broader research community.

Responsibly Managed Data Sharing in Arcus

- Contributors have the flexibility to make data internally available to approved Arcus users via these broad set of conditions, modeled after NIH data repositories like dbGAP:

  - without access restrictions.

  - only after a purpose-specific embargo of [X months/years].

  - only after my approval according to scientific review criteria established during the contribution process.

  - only as a de-identified or limited data set.

- Data is only released within the Arcus Environment and cannot be directly downloaded.

- Depending on the sharing and attribution terms, the contributor may be involved in the request review process.

Continuing the Data Sharing Discussion

- Arcus Library Sciences worked with Arcus Privacy analysts and the IRB on exemplary language to allow for indefinite data retention in Arcus for both active and completed studies without requiring the burdensome step of reconsenting all study participants.

- Arcus is highly interested in fortifying the partnership between the Arcus Archives and the Biorepository Core, especially in regards to increasing secondary use of sequencing data and making rich metadata about these kinds of assets more widely available. Collaborating with researchers in your subject area can help us achieve this goal. Please reach out!

Children's Hospital of Philadelphia®
Department of Biomedical & Health Informatics

Research Data Management

- RDM is about managing your data so you can get the most out of it
  - Easier compliance with repository standards
  - Better knowledge transfer during staff turnover
  - Remove/reduce bottlenecks
  - Easier combination with other data / collaboration with other researchers
- RDM is an area of practice and knowledge many have not been taught
  - Not commonly included in most curricula
  - Can be learned and improved upon!
- Arcus Research Data Management Services is available to advise on implementing RDM practices in your research

Research Data Management Recommended Practices

- Recommended Practices Vs. Best Practices
  - There is no one universal best
  - Focus on finding what works for your research and your team
- Low Hanging Fruit: get the most bang for your buck
  - Files
    - Establish and document directory structure
    - Standardize file names
    - Establish and document file naming conventions
  - Processes
    - Document processes such as data collection, analysis pipelines, transformations
  - Data Quality
    - Establish naming conventions for variables
    - Maintain data dictionaries for data sets
    - Consider using Common Data Elements (CDEs)

Children's Hospital of Philadelphia®
Department of Biomedical & Health Informatics

Research Data Management Recommended Practices

- And much more!
  - Survey methodology
  - REDCap data collection instrument standardization
  - README-based project documentation
  - Versioning
- This takes time
  - Plan for added time to implement whatever practices you decide on
  - Starting small & adding to practice is much better than starting big and not keeping up

Contact Us

Arcus.RDM@chop.edu

arcus-support@chop.edu