NIH Data Management & Sharing Policy Requirements Workshop

including UNMC resources for compliance

AUGUST 2022
Agenda

• What, Who, and When?
• Why?
• How?
• Available Resources: NIH, UNMC, other
• Questions and Discussion: Panel
NIH Policy for Data Management and Sharing

• To foster good data stewardship as well as data sharing:
  – A Data Management & Sharing Plan (DMP) must be submitted with all competing NIH applications submitted Jan 25, 2023, or after
  – Plan will be reviewed & approved but not impact grant scoring at this time
  – Investigators then must be compliant with/follow their ICO-approved Plan (if not, may affect future funding)
Who, what, when, and how long?

- **Who?** SPAdmin will require a data management and sharing plan (DMP) for all competing NIH applications submitted Jan 25, 2023, or after (Form H).
- **What?** In the DMP, the investigator(s) must describe how they will share the scientific data generated from that NIH supported research: published or not.
  - **Scientific data:** “recorded factual material… of sufficient quality to validate and replicate research findings”
  - **Not:** lab notebooks, preliminary analyses, case report forms, drafts of papers, future research plans, peer reviews, colleague emails or physical objects.
- **When?** *must share* by time of 1st publication or end of award.
- **How long?** It depends …..funding mechanism, institution: NU/state (7 years), or repository policy.
Any exceptions to Sharing of Data?

- **Maybe, if…**
  - **Existing restrictions:**
    - Informed consent will not permit or limits scope of sharing or use
    - Explicit federal, state, local, or Tribal law, regulation, or policy prohibits disclosure
    - Existing or anticipated contract restrictions with other parties
  - **Confidentiality of participants:** Despite deidentification, privacy or safety of research participants potentially compromised as available protections are insufficient
  - **Datasets cannot practically be digitized** with reasonable efforts
- **Not an adequate justification:** no suitable repository exists, data too small, unpublished, or unlikely to be used
- **Additional considerations:**
  - **Tribal sovereignty:** NIH supports Tribal right to determination of responsible management/sharing tribal data
  - **SBIR/STTR Program** permits withholding data for 20 years, as stipulated in agreements and consistent with program goals
Why does NIH Want Data to be Shared?

• **Advance rigorous and reproducible research**
  – Enable validation of research results
  – Make high-value datasets accessible
  – Accelerate future research directions
  – Increase opportunities for citation and collaboration

• **Promote public trust in research**
  – Foster transparency and accountability
  – Demonstrate stewardship over taxpayer funds
  – Maximize research participants’ contributions
  – Support appropriate protections of research participants’ data
What is required in a DMP?

DMP is often short: Current NIH limit of 2 pages

Recommended elements:

– **Data type**: what data will be preserved and shared
– **Related tools, software, code**: what will be needed to access and manipulate data
– **Standards**: what standards will be applied to data and metadata
– **Data preservation, access, timelines**: where the data will be stored (which repository), how it will be found as with a unique identifier, and rough time line of when & how long data will be available
– **Access, distribution, reuse considerations**: any requirements around data access, distribution, or reuse
– **Oversight of data management**: who will monitor/manage compliance (TBA)

Institute Specific requirements: Links on [https://sharing.nih.gov](https://sharing.nih.gov)

RFA or PAR Specific requirements: within announcements
Useful templates and DMP creation tool

DMPTool: https://dmptool.org

We are a member and have single sign-on (SSO)
GDS Policy preceded the current data sharing policy and has specified elements at https://sharing.nih.gov/

- 2007 - NIH GWAS Policy
- 2014 - NIH GDS policy
- 2022 – RFI on Long-Term Considerations for Genomic Data
Genomic Data Sharing (GDS) Policy

• Who should submit a GDS plan?
  • All Research projects that generate large-scale human or non-human genomic data. Some exceptions noted in PARs/RFAs
  • Included in the Resource Sharing Plan section

• When to submit?
  • Proposal time: General plan regarding how genomic data is shared, include fee structure for data storage and usage (e.g., if ‘cloud’ storage is involved) in the budget
  • Just-in-Time (JIT): Prior to funding, a detailed plan with types of data to be generated, formats, FAIR compliance, metadata, tools used for accessing data, etc.
  • An institutional Certification Form is required at JIT: Investigator completes the form and forwards to IRB for review; investigator and SO both sign https://sharing.nih.gov/genomic-data-sharing-policy/institutional-certifications/completing-an-institutional-certification-form

• Alternative Data Sharing Plan
  • Justification for not complying with the GDS plan
  • Alternative mechanisms to share as much genomic data as allowable
  • Funding IC will consider such plans on a case-by-case basis
Elements of the Genomic Data Sharing plan


• Data Type
  • Human, non-human, both
  • Genomic, transcriptomic, epigenetic, individual, aggregate
  • Methods, study protocols

• Data Repository
  • Human data must go into an NIH-designated repository
  • Non-Human can be submitted any appropriate repository

• Data Submission and Release Timeline
  • Human-6 months or the first publication, unless restricted
  • Non-human – Initial publication or earlier

• IRB approval of GDS plan
  • Institutional Certification provided at JIT, as before

• Data Use Limitations (DULs)
  • Consented or not, specific-disease use

• Statement on the Data Sensitivity for Population Studies
  • Allele frequencies linked to disease risk or other outcomes
  • Analysis statistics – P-values in genotype to phenotype studies
Plan Submission and Review: A Guide

Extramural Grant Awards*

Plan Submission
- With application
- Brief Plan description in Budget Justification with budgeted costs
- Full Plan as separate attachment (Form H)

Plan Assessment
- Peer reviewers comment on (not score) budget
- NIH program staff assess Plans
- Plans can be revised

Plan Compliance
- Incorporated into Terms and Conditions
- Monitored at regular reporting intervals – mechanisms and tools to support oversight under development
- Lack of compliance may factor into future funding decisions

*Analogous requirements for contracts, Other Transaction Awards, NIH Intramural Research Program
Additional Expectations for Plans

SHARING SHOULD BE ...

- Maximized (with justifiable limitations for ethical, legal, or technical factors)
  - All data should be managed; **not all data must be shared**

- Responsibly implemented and consistent with all existing laws, regulations, and policies
  - Plans should outline protection of privacy, rights, and confidentiality

- Prospectively planned for at all stages of the research process
Plan is one thing: what do you need to know to develop and enact a DMP?

• Which Repositor(ies) to use
• What to budget for the proposed plan: deidentifying, curating, moving or storing data may require funds which should be included in the grant proposal
• Saving and alerting SPAdmin on any updates to the plan and sharing specific identifiers that link to where the data is located: there will be institutional and NIH monitoring
Repository Selection

• Use NIH-supported or other established repositories whenever available

• NIH and the Library will help investigators identify appropriate data repositories
  • E.g., use of persistent unique identifiers, attached metadata, facilitates quality assurance

• Certain NIH Institutes or Centers may designate specific data repository(ies)

See Selecting a Data Repository for details
**Repository Selection: Specialized Data Repositories**

- Prioritize use of data-type and discipline-specific data repositories

- Refer to **NIH-supported data repository list** outlining:
  - Repository description (e.g., data-types accepted, research community served, tools available),
  - Supportive NIH IC(s),
  - Whether and when new data are accepted, and
  - How to submit data

- **Examples include:**
  - dbGaP
  - GenBank
  - NIMH Data Archive
  - BioData Catalyst
  - BioLINCC
  - ImmPort

See [Repositories for Sharing Scientific Data](#) for details
Repository Selection: Other Data Repositories

Other potentially suitable options:

- **Institutional repositories:** UNMC’s DigitalCommons for small datasets (<5 Gigabytes)

- **PubMed Central** (small datasets only)

- **General data repositories, including:**
  - Dataverse
  - Dryad
  - Figshare
  - IEEE Dataport
  - Mendeley Data
  - Open Science Framework
  - Synapse
  - Vivli
  - Zenodo

See [Repositories for Sharing Scientific Data](#) for details
Submitting Domain-Specific Data

NIH-supported

  • Each data type and repository has specific formats and metadata collection requirements.
  • Examples
    • Metabolomics/Lipidomics data
      • https://www.metabolomicsworkbench.org/
    • Cell Image Library
      • http://www.cellimagelibrary.org/

Global Repositories
  • PRIDE – Protein Identification Database
    • https://www.ebi.ac.uk/pride/
  • UniCarb-DR – Glycomic MS data repository
    • https://unicarb-dr.glycosmos.org/
  • EMPIAR – Electron Microscopy Image Archive
    • https://www.ebi.ac.uk/empiar/
Allowable Costs

• Reasonable costs are allowed in budget requests (incurred during the performance period): please find and budget these
  – Curating data/developing supporting documentation: library or cores
  – Uploading data to dbGAP or other NIH systems biology data: Bioinformatics core costs
  – Assistance with moving other large data sets: discuss with RITO
  – Repository storage: depends but ask for multi-year storage beyond award dates

• NOT considered data sharing costs
  – Infrastructure costs included in indirect costs
  – Costs associated with the routine conduct of research (e.g., costs of gaining access to research data although may be budgeted otherwise)

• NIH hopes to learn more about what constitutes reasonable costs for various data management and sharing activities

See Budgeting for Data Management & Sharing for details
NIH Scientific Data Sharing Website: NIH Resources and Updates

- A central source of guidance for multiple sharing policies
  - e.g., Data Management and Sharing Policy, Genomic Data Sharing Policy

- Resources to learn more about NIH data sharing policies, including:
  - Policy Decision Tool
    - Determine which policies may apply to your research
  - Training Resources
    - Announcements including upcoming webinars and presentations

Sharing.nih.gov
UNMC resources

• To help you manage your data throughout the project
• Library resources
• Assistance with data uploads to specific data repositories
• Website of resources: in progress
Importance of Securely Storing and Managing Your Data Throughout the Project

• Secure cloud data storage options:
  • **Box**: for large data sets or need to collaborate externally, see RITO
  • **Sharepoint**: shareable with other MS institutions; see IT
  • **REDCap**: for some clinical trial or research data sets, not FDA regulated trials
  • **Clinical Trial Management System**: FDA regulated trials (Part 11 compliant)
• **E-lab electronic notebook**: store data locally in a way to be easily found for a specific project, including all your trainees
McGoogan Library resources: Emily Glenn

- **Access to sample data management plans**
  
  McGoogan Library: [https://unmc.libguides.com/rdm](https://unmc.libguides.com/rdm)

  DMPTool’s Public resource: [https://dmptool.org/public_plans](https://dmptool.org/public_plans)

- **Library staff can help you with** metadata planning, data visualization, locating data sets, citing data

- **New Data Services Librarian** to help you use DMPTool, plan development guidance, assistance with finding/selecting a repository, general support with meeting NIH DSMP policy

- **UNMC digital commons (DigitalCommons@UNMC)** potential repository for smaller datasets
Data Repository Assistance: Selection and Uploads (most require fees)

• Genomic and other systems biology data: Bioinformatics and Systems Biology Core can discuss & assist with uploads (fee required)

• Clinical trials data: What can/should be shared depends on informed consent, approval for secondary use, and if it can be deidentified
  • CCORDA can help you deidentify some datasets
  • X-rays or other human images may also require deidentification
  • IRB can help you with access to Clinicaltrials.gov

• Selection of Repositories: discuss with Research Data Librarian

• Uploading files (<5 Gigabytes) to DigitalCommons: Library

• Moving large data sets to other repositories: RITO can assist you
  • Request estimate when developing your DMP and budget accordingly
Join the second part of Webinar Series!

https://sharing.nih.gov/about/learning/DMS-Update
Policy and Supplemental Information:

- **NOT-OD-21-013** – Final NIH Policy for Data Management and Sharing
- **NOT-OD-21-014** – Supplemental Information to the NIH Policy for Data Management and Sharing: Elements of an NIH Data Management and Sharing Plan
- **NOT-OD-21-015** – Supplemental Information to the NIH Policy for Data Management and Sharing: Allowable Costs for Data Management and Sharing
- **NOT-OD-21-016** – Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research
- **NOT-OD-22-195** – Notice of Form H required to describe DMP

NIH Resources:

- **NIH Data Sharing Website** – sharing.nih.gov
- **NIH Office of Science Policy DMS Policy Website** – history and background on the NIH DMS Policy
- **Frequently Asked Questions** – sharing.nih.gov/faq
- **Webinar** – Recording from 2021 Presentation on DMS Policy
- **News & Events** – Latest news and upcoming events

NIH Contacts:

- Questions – sciencepolicy@mail.nih.gov
- Follow us on Twitter – @NIH_OSP, @NIHGrants

Additional Resource:

- **DMPTool** – examples and templates