

DATA MANAGEMENT AND SHARING PLAN FOR NIH FUNDED RESEARCH

OMB Number: 0925-XXXX

Expiration Date: XX/XX/XXXX

Public reporting burden for this collection of information is estimated to vary from 30 to 60 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid OMB control number. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to: NIH, Project Clearance Branch, 6705 Rockledge Drive, MSC 7974, Bethesda, MD 20892-7974, ATTN: PRA (0925-XXXX). Do not return the completed form to this address.

Under the Final NIH Policy for Data Management and Sharing

((<https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html>), NIH requires researchers to prospectively plan for how scientific data generated from NIH-funded or conducted research will be managed and shared through submission of a Data Management and Sharing Plan (Plan). The purpose of this template is to facilitate development of a Plan. Upon NIH approval of a Plan, NIH expects researchers and institutions to implement data management and sharing practices as described.

Data management is a vital aspect of high-quality research; it both contributes to impactful research findings and enhances the utility of data for future use. The completed Plan will be reviewed by NIH staff to ensure that researchers are taking steps to maximize the appropriate sharing of scientific data. The Plan is expected to evolve during the life of the project; therefore, significant changes should be submitted to NIH for approval, in the form of an updated Plan. Plans may be shared publicly and should not include proprietary information. References to template instructions and policy information can be found at:

<https://www.nichd.nih.gov/about/org/od/odss>

DATA MANAGEMENT AND SHARING PLAN FOR NIH FUNDED RESEARCH

OMB Number: 0925-XXXX

Expiration Date: XX/XX/XXXX

PART I: General Information (To be completed by all applicants)

New or Revised Plan? Choose an item.

Plan version number: Click or tap here to enter text.

Plan Submission Date: Click or tap to enter a date.

PI name: Click or tap here to enter text.

Point of Contact for DMS plan: Click or tap here to enter text.

Project/Application/Protocol ID: Click or tap here to enter text.

PART II: Data Management Sharing Plan Details

Add one row for each proposed repository/mechanism to share respective data types.

COLUMN 1 Name of repository/ mechanism (list URL if available)	COLUMN 2 Category of repository/ mechanism (if applicable or leave blank)	COLUMN 3 Broad data type(s) categories to be shared (Select all that apply; optional)	COLUMN 4 Brief description of data from COLUMN 3 selections and/or free text here	COLUMN 5 Estimated Data Deposition and Sharing Dates	COLUMN 6 Estimated data amount to be shared (optional)
	Choose an item.	Choose an item.		Submission: Release:	Choose an item.
	Choose an item.	Choose an item.		Submission: Release:	Choose an item.

*Does the Genomic Data Sharing (GDS) Policy apply? ☐ YES ☐ NO

If YES:

DATA MANAGEMENT AND SHARING PLAN FOR NIH FUNDED RESEARCH

OMB Number: 0925-XXXX

Expiration Date: XX/XX/XXXX

Will the datasets be shared according to GDS policy but no later than the time of publication or end of the project, whichever is sooner?	<input type="checkbox"/> YES <input type="checkbox"/> NO
Will an NIH-supported repository be selected for data subject to GDS?	<input type="checkbox"/> YES <input type="checkbox"/> NO
Has an Institutional Certification (IC) been submitted with the application or Just-In-Time that meets GDS criteria?	<input type="checkbox"/> YES <input type="checkbox"/> NO
If responses to (1), (2) or (3) for GDS is NO, explain:	

PART II: Data Management Sharing Plan Description

Element 1: Data Type	
Will <u>all</u> scientific data generated by the research project be shared in a data repository that makes data available to the larger research community? If Yes, go to Element 2. If No, explain the rationale that determines which scientific data will not be shared.	<input type="checkbox"/> YES <input type="checkbox"/> NO
Element 2: Tools, Software, Code	
Describe the tools, software, and/or code that are needed to access or manipulate shared scientific data to support replication or reuse, if any.	
Describe how researchers can access the tools, software, and/or code listed above. Describe if "Other."	Choose an item.

DATA MANAGEMENT AND SHARING PLAN FOR NIH FUNDED RESEARCH

OMB Number: 0925-XXXX

Expiration Date: XX/XX/XXXX

Element 3: Standards	
List data or metadata standards or common data elements that will be used applicable to each data type shared.	Write NA if no existing standards.
Element 4: Data Preservation, Access, and Timelines	
Explain if data sharing timelines will not meet expectations of the DMS or other policies, if applicable.	Write NA if timelines will be met per policy.
What types of persistent identifiers/ indexing methods will be used for data releases, to enable findability and citation of shared datasets?	
Element 5: Access, Distribution or Reuse Considerations	
Describe any limitations or factors affecting subsequent access, distribution, or reuse of this data.	
Are there any privacy or informed consent considerations for human data? If Yes, describe including methods to protect privacy and confidentiality.	<input type="checkbox"/> YES <input type="checkbox"/> NO
What type of access will secondary users utilize to access the shared data? Describe if "Other."	Choose an item.
Element 6: Compliance	
Describe how compliance with the Plan will be monitored and managed, frequency of oversight, and by whom.	
Will data management and/or sharing activities be facilitated by individuals outside of the project team? If YES, list individual(s), their organization(s), and describe their role(s) and responsibilities.	

PART III: Additional Information (optional)

If additional policies apply (e.g., Clinical Trials Access Policy, FOA-specific requirements), describe additional information required to meet the policy: [Click or tap here to enter text.](#)

DATA MANAGEMENT AND SHARING PLAN FOR NIH FUNDED RESEARCH

OMB Number: 0925-XXXX

Expiration Date: XX/XX/XXXX

Provide any additional information or context for readers and reviewers of your Data Management and Sharing Plan:

Click or tap here to enter text.

Appendix. Data Types Glossary

Disclaimer:

- Data types defined here as structured information are for this template only. They are categorized into very broad "data buckets." Some data types and their sub-categories (e.g., Patient Registry under Epidemiology and Population Sciences Data) may overlap with other data types and their sub-categories (e.g., Clinical Data). Please check all that may apply.
- The broad data type "buckets" and their sub-categories encompass wide scopes, e.g., clinical data here serves as a broad term for data including but are not limited to intervention responses, clinical trials, EHR, etc.
- The data type "buckets" and sub-categories may not be comprehensive. Selection of pre-defined data types in the drop down menu in COLUMN 3 is optional for researchers. If they choose to describe their data types in free text, they can leave COLUMN 3 blank and describe their data in COLUMN 4. Describe "Other" data types not listed elsewhere in COLUMN 4 as well.
- Examples of data type content that are query-able are:
<https://datacatalog.ccdi.cancer.gov/resource/Kids%20First%20Data%20Resource>; <https://cedcd.nci.nih.gov/select>;
<https://healthcaredelivery.cancer.gov/data/>; <https://portal.gdc.cancer.gov/projects>

Glossary:

1. **Omics Data** includes DNA/RNA sequence data, epigenetic data, transcriptomic data, proteomic data, metabolomics data, phenomic data, etc.
 - **Genomics/Genetics:** The study of the structure, function, expression, evolution, mapping and editing of genes/genomes. NCI-GLOSS Definition: The study of the complete genetic material, including genes and their functions, of an organism. In the context of this template, it includes DNA and RNA sequence data, as well as epigenetic regulation such as DNA

DATA MANAGEMENT AND SHARING PLAN FOR NIH FUNDED RESEARCH

OMB Number: 0925-XXXX

Expiration Date: XX/XX/XXXX

methylation, DNA-protein interactions, chromatin accessibility, histone modifications.

https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C84343

- **Proteomics:** The global analysis of cellular proteins. Proteomics uses a combination of sophisticated techniques including two-dimensional (2D) gel electrophoresis, image analysis, mass spectrometry, arrays, amino acid sequencing, and bioinformatics to resolve comprehensively, to quantify, and to characterize proteins. The application of proteomics provides major opportunities to elucidate disease mechanisms and to identify new diagnostic markers and therapeutic targets.

NCI-GLOSS Definition: The study of the structure and function of proteins, including the way they work and interact with each other inside cells.

https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C20085

- **Metabolomics:** The study of the biological metabolic profile of a cellular specimen in a specific environment at an isolated timepoint. This discipline depicts the physiological states of cells and organisms by focusing on carbohydrates, lipids, and other metabolites. Several analytical techniques are utilized to quantify the metabolic content of specimens such as mass spectrometry and electrophoretic applications.

https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C49019

- **Other Omics:** Other omics data include lipidomics, metagenomics, microbiomics, etc. not listed elsewhere in the data type categories.

2. **Imaging:** A process that makes pictures of areas inside the body. Imaging uses methods such as x-rays (high-energy radiation), ultrasound (high-energy sound waves), and radio waves.

- **Medical Imaging:** Any technology which helps in visualization of any biological process, cell, tissue or organ for the screening, diagnosis, surgical procedure or therapy.

https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C16831

- **Histopathology:** The microscopic study of characteristic tissue abnormalities by employing various cytochemical and immunocytochemical stains.

NCI-GLOSS Definition: The study of diseased cells and tissues using a microscope.

https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C18190

- **Cellular Imaging:** Cellular imaging encompasses the techniques that allow the detection and analysis of cellular organelles and macromolecules. Cellular imaging observations are obtained using light-based or electron-based microscopes and often request further analysis with computer based programming. It provides a powerful tool to study

DATA MANAGEMENT AND SHARING PLAN FOR NIH FUNDED RESEARCH

OMB Number: 0925-XXXX

Expiration Date: XX/XX/XXXX

biological processes including regulation, protein-protein interaction, trafficking, development, cellular structure and morphology, to name a few.

3. **Model Organisms:** A model organism (often shortened to model) is a non-human species that is extensively studied to understand particular biological phenomena, with the expectation that discoveries made in the model organism will provide insight into the workings of other organisms. Model organisms are widely used to research human disease when human experimentation would be unfeasible or unethical. This strategy is made possible by the common descent of all living organisms, and the conservation of metabolic and developmental pathways and genetic material over the course of evolution. Studying model organisms can be informative, but care must be taken when generalizing from one organism to another. There are many model organisms ranging from *Drosophila melanogaster* and yeast to rodents and canines. In cancer, it represents theoretical pre-clinical models before being tested clinically and can encompass cancer cell lines, computational cancer models, genetically engineered mouse models, organoids, patient-derived xenografts (PDXs).
https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C18953
4. **Epidemiology and Population Sciences:** The study of the causes, incidence and distribution of disease in the population and its application for prevention or control.
NCI-GLOSS Definition: The study of the patterns, causes, and control of disease in groups of people and may include Epidemiology, Nutrition, Behavioral Medicine, Genetics, Public Health, Anthropology, Tobacco Control Research, Demography, Bioethics, Sociology.
https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C16556;
https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&version=22.11d&ns=ncit&code=C19491&key=n1932309059&b=1&n=null

Surveillance is the systematic collection, analysis, and interpretation of health data on an ongoing basis. In medicine, the ongoing collection of information about a disease, such as cancer, in a certain group of people. The information collected may include where the disease occurs in a population and whether it affects people of a certain gender, age, or ethnic group.

https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C15719

A patient registry is an organized system that uses observational study methods to collect uniform data (clinical and other) to evaluate specified outcomes for a population defined by a particular disease, condition, or exposure and that serves predetermined scientific, clinical, or policy purpose(s). Different types of registries track specific aspects of care. A registry may focus on a disease or condition, a procedure, or a medical device. The registry defines a patient population, then recruits physicians and other healthcare professionals to submit data on a representative sample of those patients.

https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&version=22.11d&ns=ncit&code=C129000&key=n2025685980&b=1&n=null

DATA MANAGEMENT AND SHARING PLAN FOR NIH FUNDED RESEARCH

OMB Number: 0925-XXXX

Expiration Date: XX/XX/XXXX

5. **Clinical Data (Routine Care and/or Clinical Research Data):** Data obtained through patient examination or treatment. Clinical data in this context pertains to the medical well-being or status of a patient. It encompasses a wide range of data including Electronic Health Record (EHR), clinical reports, clinical trial data, etc. For example, EHR is digital version of a patient's medical history and has the following distinguishing features: able to be obtained from multiple sources; shareable; interoperable; accessible to authorized parties.

https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C142529.

Clinical data also include clinical reports and individual patient data (IPD)

[https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C15783], as well as clinical trials, i.e., information collected during the course of a scientifically-controlled study of the safety and efficacy of a therapeutic agent, such as a drug or vaccine, using consenting human subjects. See also clinical trial information:

https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C142441

6. **Administrative Data (Healthcare):** Refers to information that is collected, processed, and stored in automated information systems. Administrative data include enrollment or eligibility information, claims information, and managed care encounters. The claims and encounters may be for hospital and other facility services, professional services, prescription drug services, laboratory services, and so on.
7. **Bioassays or Measurements:** A qualitative or quantitative analysis performed to determine the amount of a particular constituent in a sample or the biological or pharmacological properties of a drug. NCI-GLOSS Definition: A laboratory test to find and measure the amount of a specific substance. In the context of this template, assays or measurements are broadly defined to include but are not limited to biochemical (e.g., ions and cofactor assays, cell-based assays), biophysical (e.g., circular dichroism, fluorescence, atomic force microscopy), physiological (e.g., EEG, patch clamp), immunological (e.g., ELISA, virus neutralization assays), and pharmacological (e.g., receptor-ligand assays, pharmacokinetic/pharmacodynamic assays) methods. <https://www.cancer.gov/publications/dictionaries/cancer-terms/def/assay>; https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C25414
8. **Social, Psychological and Behavioral Data:** <https://www.healthit.gov/isa/social-psychological-and-behavioral-data>; <https://www.nih.gov/sites/default/files/research-training/about-obssr.pdf>
9. **Wearable/Sensor Data:** A device that responds to a stimulus, such as heat, light, or pressure, and generates a signal that can be measured or interpreted. https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C28220
10. **Software, Coding, Modeling, or Methods Development:**

DATA MANAGEMENT AND SHARING PLAN FOR NIH FUNDED RESEARCH

OMB Number: 0925-XXXX

Expiration Date: XX/XX/XXXX

Definition: A technique which attempts to provide an abstract model of a particular system. It utilizes a mathematical model, which attempts to predict the behavior of the system from a set of parameters and initial conditions.

NCI-GLOSS Definition: In cancer treatment, a process used to plan radiation therapy so that the target area is precisely located and marked.

https://ncithesaurus.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI_Thesaurus&ns=ncit&code=C16461

11. **Associated Metadata:** Data about data; information that describes another set of data. CDISC-GLOSS Definition: Data that describe other data, particularly XML tags characterizing attributes of values in clinical data fields. Metadata can help users find relevant information and discover resources, and also help organize electronic resources, provide digital identification, and archive and preserve resources. In the context of research, associated metadata can be experimental protocols, instrument parameters, model organism phenotypic description, documentations about the described data types.
12. **Other:** Additional data types not pre-defined in the drop down menu for this template.