SOP#: RPS-21	**Establishing a Data Management and Sharing Plan**
Version #: 4.0	**Next Review Date: 08/2025**
Approved Date: 08/2023	**Review Interval Period: Biennial**
NCI Clinical Director Signature/ Effective Date:	James Gulley Digitally signed by James Gulley Date: 2023.08.30 12:07:11 -04'00'

POLICY

The NIH Data Management and Sharing (DMS) Policy took effect on January 25, 2023, to promote the sharing of scientific data. Sharing scientific data accelerates biomedical research discovery, in part, by enabling validation of research results, providing accessibility to high-value datasets, and promoting data reuse for future research studies.

Under the DMS policy, NIH expects that investigators and institutions:

- Plan and budget for the managing and sharing of data
- Submit a DMS plan for review when applying for funding
- Comply with the approved DMS plan

For the Intramural Research Program, a Data Management and Sharing (DMS) plan will be required for scientific data from research associated with a:

- ZIA (human and non-human research)
- Clinical protocol that will undergo IC Initial Scientific Review
- Genomic Data Sharing (GDS) project

The single DMS Plan incorporates the requirements of both the DMS and GDS policies.

Investigators are required to develop the following documents **PRIOR** to starting the research:

Human Studies:

1) Data Management and Sharing (DMS) Plan **AND**
2) Institutional Certification (IC) Memo

Non-Human Studies:

1) Data Management and Sharing (DMS) Plan

Applicable data types include:

Genomic data – see Appendix A

Large-scale germline and/or somatic data, such as

- Genome-wide association studies (GWAS)
- Single nucleotide polymorphisms (SNP) arrays
- Genome sequence
- Transcriptomic
- Metagenomic
- Epigenomic

- Gene expression data
- See Appendix A for NCI Center for Cancer Research guidance on required genomic data sharing thresholds.

Imaging Data

- Medical imaging (e.g., ultrasound, MRI, CT)
- Non-medical imaging (e.g., fluorescence microscopy)
- Other (e.g., histopathology)

Biological Data

- Electrophysiology (e.g., sensor data, ECG)
- Biochemical (e.g., X-ray, NMR, AMF, FRET)
- Pre-clinical (e.g., PDX growth curve)

Phenotype Data

- Non-human (e.g., phenotypic features of animal models)
- Human traits (e.g., blood type)
- Demographics
- Clinical data (including specimen information)

Additional Data

- Clinical trial results
- Associated metadata
- Epidemiology/Surveillance
- Administrative
- Algorithm/Simulation
- Social/Behavioral
- Survey/questionnaire

PURPOSE

The purpose of this standard operating procedure is to provide information to comply with appropriate NIH Data management and sharing processes.

RESOURCES

- NIH Data Management and Sharing Policy Website
- NCI Data Management and Sharing submission portal
- CCR Data Management and Sharing Plans website
- NCI Office of Data Sharing website
- NIH Office of Science Policy Scientific Data Sharing website
- NIH Office of Intramural Research 2023 NIH Data Management and Sharing Policy website
- GDS Data Repositories and Trusted Partners

CCR Genomic Program Administrator (GPA)
Kathleen Calzone, PhD, RN, AGN-BC, FAAN
240-760-6168
calzonek@mail.nih.gov

CCR Genomic Program Administrator (GPA)
Abid Al Reza, PhD
240-858-7909
abid.reza@nih.gov

PROCEDURES FOR SUBMITTING A DATA MANAGEMENT AND SHARING (DMS) PLAN AND INSTITUTIONAL CERTIFICATION

NOTE: Please have key protocol or study documents readily available to streamline the process.

Please refer to the *CCR Data Management and Sharing Plan Submission Portal for Clinical Protocols User Manual* below this SOP for using the Portal to submit a DMS Plan and Institutional Certification Memo.

STEP 1: Create a Project

- See the *CCR Data Management and Sharing Plan Submission Portal for Clinical Protocols User Manual* below this SOP for steps to create a project

STEP 2: Creating an Institutional Certification Memo, if applicable

An **Institutional Certification Memo** certifies that:

- data submission and sharing are consistent with the informed consent of the study participants
- documents whether genomic summary results can be open access or require controlled access
- consideration was given to risks to individual participants and their families associated with the shared data
- to the extent possible, consideration was given to risks, groups, or populations related to the shared data
- the principal investigator's plans for de-identifying data sets are consistent with the GDS and DMS policies

See the *CCR Data Management and Sharing Plan Submission Portal for Clinical Protocols User Manual* below this SOP to create the Institutional Certification memo.

- Submitted Institutional Certification Memos are first routed to the GPA for review and will either be returned to the investigator for revision or approved by the GPA.

  Revision Required

  - The submitter and/or PI will be notified via email that the submission requires revision and includes a hyperlink specific to the document.
  - Open the document and review the items in the Rejection Comments section and address each item.
  - Select Save, followed by Submit. The submitted Institutional Certification will be routed to the GPA for re-review.

Approved by the GPA

- The submitter and/or PI will be notified via email that the GPA approved the submission.
- The GPA will route the GPA-approved Institutional Certification to the Scientific Director for review and approval.

Approved by the Scientific/Clinical Director

- The submitter and/or PI will be notified via email that the GPA approved the submission.
- The GPA will submit the GPA-approved IC memo to the CCR Scientific Director for review and approval.
- The submitter will be notified, via email, of approval by the Scientific/Clinical Director.
- This completes the Institutional Certification memo approval process.

STEP 3: Creating a Data Management and Sharing Plan

The **Data Management and Sharing Plan** provides information on the proposed research that will generate data for which the policy applies. The plan must address six core elements: Data type; Related Tools, Software, and/or Code; Data Standards; Data Preservation, Access, and Associated Timelines; Access, Distribution, or Reuse Considerations; and Oversight of Data Management and Sharing. There is also an option to include other elements as applicable.

See the *CCR Data Management and Sharing Plan Submission Portal for Clinical Protocols User Manual* below this SOP for the completion and submission of your plan.

- Submitted DMS Plans are first routed to the GPA for review and will either be returned to the investigator for revision or approved by the GPA.

Revision Required

- The submitter and/or PI will be notified via email that the submission requires revision and includes a hyperlink specific to the document.
- Open the document and review the items in the Rejection Comments section and address each item.
- Select Save, followed by Submit. The submitted DMS plan will be routed to the GPA for re-review.

Approved by the GPA

- The submitter and/or PI will be notified via email that the GPA approved the submission.
- The GPA will route the GPA-approved DMS Plan to the Scientific Director for review and approval.

Approved by the Scientific/Clinical Director

- The submitter and/or PI will be notified via email that the GPA approved the submission.
- The GPA will submit the GPA-approved DMS Plan to the CCR Scientific Director for review and approval.
- The submitter will be notified, via email, of approval by the Scientific/Clinical Director.
- This completes the DMS Plan approval process.

Appendix A: NCI Guidance on Genomic Data Sharing

Examples of projects for which the NCI anticipates data sharing (*regardless of study design*) include, but are not limited to:

	# of Specimens	
	Human *(including human cell lines)*	**Model Organisms, Non-Human Cell Lines, Infectious Organisms**
SNP array data from >500K single nucleotide polymorphisms (SNPs) *(e.g., GWAS data)*	1,000	500
DNA sequence data from < 100 genes or regions of interest *(e.g., targeted sequencing)*	1,000	500
DNA sequence data from ≥ 100 genes or regions of interest *(e.g., targeted sequencing, whole exome sequencing, whole genome sequencing)*	100	50
Genome-wide RNA sequencing (RNA-seq) data *(e.g., transcriptomic data)*	100	50
Genome-wide DNA methylation data *(e.g., bisulfite sequencing data)*	100	50
Genome-wide chromatin immunoprecipitation sequencing (ChIP-seq) data *(e.g. transcription factor ChIP-seq, histone modification ChIP-seq)*	100	50
Metagenome (or microbiome) sequencing data *(e.g., 16S rRNA sequencing, shotgun metagenomics, whole-genome microbial sequencing)*	100	50
Metatranscriptome sequencing data *(e.g., microbial/microbiome transcriptomics)*	100	50

NOTE: The number of samples includes distinct individuals, species, strains, samples, treatments, time points, and tissues. For example, data from 25 patients at four time points after treatment would reach a 100-sample threshold, as would data from 50 tumor-normal comparisons. Guidance on Genomic Data Sharing for rare diseases and rare cancers

Guidance issued from the Office of the Director, CCR mandates sharing of genomic data for projects examining rare diseases and rare cancers. There are **no minimum thresholds** to meet for such projects. The Trans-NCI Genomic Data Sharing Working Group has adopted the definition of rare disease – a disease that affects **less than 200,000 persons** in the United States, that has been set forth by the U.S. Food and Drug Administration (FDA).

Examples of smaller-scale projects that the NCI would likely mandate data sharing for include, but are not limited to:

- Projects examining rare cancers, rare-cancer-related outcomes, or rare cancer subtypes
- Projects focusing on under-studied populations

Examples of Research Outside the Scope of the GDS Policy:

Examples of NIH-funded research or research-related activities that are outside the Policy's scope include, but are not limited to, projects that do not meet the criteria in the above examples and involve:

- Instrument calibration exercises
- Statistical or technical methods development