SOP#: RPS-21	**Establishing a Genomic Data Sharing Project and Required Documents**
Version #: 2.0	**Next Review Date: 01/2021**
Approved Date: 01/2019	**Review Interval Period: Biennial**
NCI Clinical Director Signature:	William L. Dahut -S Digitally signed by William L. Dahut -S Date: 2019.01.22 08:45:23 -05'00'

POLICY

The Genomic Data Sharing (GDS) Policy applies to all NIH intramural research that generates large-scale human or non-human genomic data as well as the use of these data for subsequent research. Large-scale data include genome-wide association studies (GWAS), single nucleotide polymorphisms (SNP) arrays, and genome sequence, transcriptomic, metagenomic, epigenomic, and gene expression data. See Appendix A for NCI Center for Cancer Research guidance on thresholds for required data sharing.

For studies that will generate large-scale genomic data, investigators are required to develop the following documents **PRIOR** to start the research:

Human Studies:

1) Genomic Data Sharing Plan **AND**
2) Institutional Certification Memo.

Non-Human Studies:

1) Genomic Data Sharing Plan

A **Genomic Data Sharing Plan** provides information on the proposed research that will generate large-scale human and non-human genomic data for which the GDS policy applies.

An **Institutional Certification Memo** certifies that:

- data submission and sharing are consistent with the informed consent of the study participants;
- 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 to groups or populations associated with the shared data and
- the principal investigator's plans of de-identifying data sets are consistent with the GDS policy.

More than one Institutional Certification Memo is required if the research began prior to August 31, 2015 (with or without consent) and continued past the policy implementation date when consent is mandatory. One memo is required for each instance:

1) prior to August 31, 2015 with consent
2) prior to August 31, 2015 without consent; and/or
3) after August 31, 2015 with consent.

Effective November 1, 2018, the NIH Genomic Data Sharing Policy was expanded to enable **unrestricted access** to Genomic Summary Results (GSR). GSR are defined as any systematically computed statistics such as, but not limited to, genotype counts and frequencies and allele counts and frequencies. The GSR are available through unrestricted access except when a population is determined to meet sensitive criteria. Sensitivity is defined as study populations from isolated geographic regions, or with rare or potentially stigmatizing traits which could result in increased privacy or confidentiality risks. The IRB is responsible for making the sensitivity determination with input from the investigators. If a sensitivity determination has not yet been made by the IRB, contact the CCR GPA Kathleen Calzone, 240-760-6168, calzonek@mail.nih.gov

Multi-Institutional studies in which NIH is the coordinating site and will be responsible for data sharing requires an Institutional Certification Memo from each site that is accruing samples/subjects unless NIH is the single IRB of record.

PURPOSE

The purpose of this standard operating procedure is to provide instructions for creating a GDS Project and corresponding Genomic Data Sharing Plan and, for human studies, the Institutional Certification Memo(s).

RESOURCES

- CCR Genomic Data Sharing Policy Website
- CCR Genomic Data Sharing submission portal
- NIH Genomic Data Sharing Policy
- NIH Genomic Data Sharing Website
- NIH Intramural Investigator GDS Responsibilities
- GDS Data Repositories and Trusted Partners
- NCI Genetic Data Sharing Policy Website

CCR Genomic Program Administrator (GPA)

Kathleen Calzone, PhD, RN, AGN-BC, FAAN
240-760-6168
calzonek@mail.nih.gov

GPA Assistant

Position vacant

PROCEDURES

Step 1: Project Information

- All projects (human and non-human) that require compliance with the NIH GDS policy require basic project information.
  - Please refer to the CCR GDS website or contact the CCR GPA if you are uncertain whether your project/protocol falls within the purview of the GDS policy.
- Project Information is entered into the Genomic Data Sharing Portal
  - Sign into portal using your PIV card or NIH Username/password
  - Genomic Data Sharing Plan and Institutional Certification Memo(s) are generated via the portal
- Required Project Information includes:
  - Project Title
  - Principal Investigator name
    - For clinical trials, Principal investigator is considered the protocol Principal Investigator or Lead Associate Investigator. For laboratory-based studies, Principal Investigator is considered the scientist leading the research.
  - Email, phone number, title, and branch will be auto-populated from NIH Enterprise Directory (NED)
    - If there is an error, please contact your Administrative Officer to get NED content updated.
  - Specify the organism type (human, non-human, or both)
  - Insert the Z1A number (can use the look-up function to retrieve your number)
  - Indicate the name/study number of the Parent Protocol/Project
    - Parent Protocols/Projects may be Tissue Procurement studies, OHRSP exemptions, or other umbrella projects in which multiple protocols/projects with different investigators are conducted under the umbrella protocol/project.
- Save the Project Information

Step 2: Establishing User Access

You have the option of giving other NIH staff permission to access the Project and associated forms.

- Go to the User Access ribbon and click NEW
  - Enter the NIH User Name (which will link to full name and email)
  - Access type (Read Only or Read/Write)

Step 3: Select Forms to Complete

- **Human Studies** require a Genomic Data Sharing Plan AND Institutional Certification Memo(s)
- **Non-Human Studies** require a Genomic Data Sharing Plan only

Step 4: Genomic Data Sharing Plan

- A Genomic Data Sharing (GDS) Plan can be established by clicking the Genomic Data Sharing Plan ribbon.
- Click to expand the Ribbon and select NEW to start generating a GDS Plan.
- Elements of a GDS Plan include the following:
    - Principal Investigator information, auto-populated from the Project Information
    - Project/protocol Information
    - Data type
    - Data repository identified for submission
    - Data submission and release timeline
    - Data use limitations – these reflect any data use limitations specified in the consent. There are four main categories:
        - General Research Use – data can be used for any research purpose
        - Health/Medical/Biomedical – data use is limited to a focus on health/medical/biomedical research objectives, excluding the study of population origins or ancestry
        - Disease-Specific – data can be used only for research on a specific disease or related condition
        - Other – specify other criteria
- Genomic Data Sharing Plan review and approval
    - The GPA will review and approve the GDS Plan
    - For plans that require revision, the plan will be returned to the investigator with instructions from the GPA about the needed modifications.
        - Once the revisions have been made, resubmit the GDS Plan for review and approval by the GPA.
    - The submitter will be notified of GPA approval via an email generated from the submission system.
- Scientific Director review and approval
    - GPA will submit the GPA approved CCR Genomic Data Sharing Plan to the CCR Scientific Director for review and approval.
    - Approval by the Scientific Director completes the GDS Plan submission and approval process. The submitter will be notified, via email, of approval by Scientific Director.
    - For clinical studies, all completed GDS Plans will be available to the Protocol Support Office (PSO) staff. The PSO will upload a copy of completed Genomic Data Sharing Plan to the appropriate protocol specific regulatory file.

- For plans not approved by the Scientific Director:
  - GPA will work with the investigator to fulfill the requirements of the CCR Genomic Data Sharing Plan policy.
  - Once a revised CCR Genomic Data Sharing Plan is in place, the revised plan will require resubmission for approvals.

- Once approved, at any point you wish to generate the GDS Plan, select Export.

Step 5: Institutional Certification Memo(s)

- An Institutional Certification Memo is established by clicking the Institutional Certification ribbon.
  - **Note:** more than one Institutional Certification Memo is required if the research began prior to August 31, 2015 (with or without consent) and continued past the policy implementation date when consent is mandatory. One memo is required for each instance:

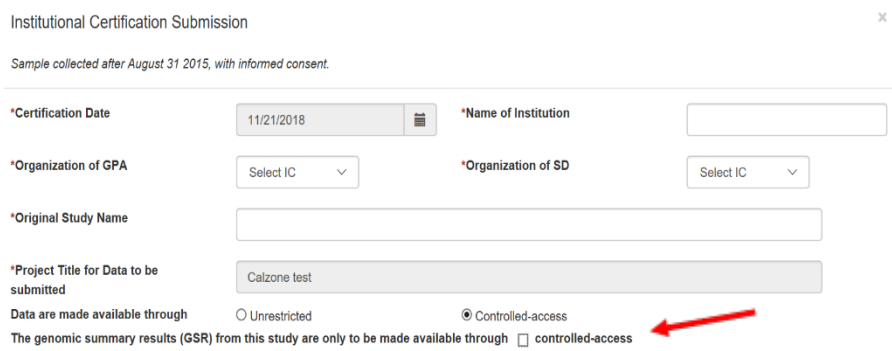
    1) prior to August 31, 2015 with consent

    2) prior to August 31, 2015 without consent; and/or

    3) after August 31, 2015 with consent.
  - **Only ONE Institutional Certification Memo can be generated at a time.** If your project requires more than one memo, you will need to repeat the procedures described below for EACH instance requiring an Institutional Certification Memo.
- Click to expand the ribbon and select NEW to start generating an Institutional Certification Memo.
  - A pop-up box will open
    - Indicate if the samples have been collected before August 31, 2015 (Yes/No)
    - Indicate if the samples were collected with or without consent (Yes/No)
- Complete the header information that is not already pre-populated from the Project Information

  NOTE: This portal may be used by other NIH Institutes so not all information provided is prepopulated for NCI.
  - Name of Institution – NCI
  - Organization of GPA – NCI
  - Organization of SD (Scientific Directors) – NCI
  - Original Study Name
  - Project Title for data to be submitted will be auto-populated from the Project Information entered in Step 1.
  - Select data access, unrestricted or controlled access
    - Human studies are mostly controlled access

- IF the release of Genomic Summary Results (GSR) was determined by the IRB to be sensitive, check the controlled-access box (see figure below).



and provide an explanation in the box provided (see figure below).

NOTE: The GSR are available through unrestricted access **except** when a population is determined to meet sensitive criteria.

- IF APPLICABLE: Add consent groups for each collaborating site
  - Select ADD
  - Collaborating site name (one collaborator per row)
  - Select Data Use Limitations
    - Data Use Limitations are based on the terms of the informed consent of the study participants from whom the genomic data have been generated.
    - Click the **?** to open definitions for each option.
  - Data Use Limitation Modifiers
    - These modifiers are also based on the consent.
    - Click the **?** to open definitions for each option if applicable.
- Select action
  - Save – enables you to return at a later time and edit content
  - Submit – circulates the memo for review and approval/signature
  - Cancel – all content entered will be withdrawn

NOTE: You may have to complete more than one Institutional Certification Memo

- Institutional Memo review and approval
  - Submitted Institutional Certification Memos are first routed to the GPA for review.

        - Memos that require revision will be returned to the investigator with instructions from the GPA of the needed modifications.
            - Once the revisions have been made, resubmit the Institutional Certification Memo for review and approval by the GPA.
    - The submitter will be notified of GPA approval via an email generated from the submission system.
- Scientific Director review and approval
    - GPA will submit the GPA approved Institutional Certification Memo to the CCR Scientific Director for review and approval.
    - Approval by the Scientific Director completes the Institutional Certification Memo submission and approval process. The submitter will be notified, via email, of approval by Scientific Director.
    - For Institutional Certification Memos not approved by the Scientific Director:
        - GPA will work with the investigator to fulfill the requirements of the Institutional Certification Memo.
        - Once a revised Institutional Certification memo is in place, the revised plan will require resubmission for approvals.
- Once approved, at any point you wish to generate the Institutional Certification Memo, select Export.

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