IMPLEMENTATION OF THE NIH GENOMIC DATA SHARING POLICY IN THE INTRAMURAL HUMAN SUBJECTS RESEARCH PROGRAM

1. Purpose of the NIH Genomic Data Sharing Policy

The NIH Genomic Data Sharing Policy, also known as the GDS Policy, establishes the expectations and responsibilities that ensure the broad and responsible sharing of large-scale genomic research data generated from NIH-funded research. This policy helps facilitate NIH's mission to share research data that leads to translation of research results into knowledge, products, and procedures that improve human health. For the purpose of the GDS Policy, the "genome" is defined as the entire set of genetic instructions found in a cell.

2. Implementation Date of the NIH GDS Policy

The original effective date for the NIH Genomic Data Sharing Policy for both Extramural and Intramural research was January 25, 2015. After the policy was announced the effective date for NIH intramural research projects was changed to **August 31, 2015**.

3. Applicable Research under the GDS Policy

The discussions and decision about whether the Policy should be applied to a protocol should occur at the stage of Scientific Review, so that the investigator is able to take all the steps to allow for genomic data sharing before the research starts.

The GDS Policy applies to all NIH-funded research generating large-scale human or non-human genomic data and the use of these data for subsequent research. The policy applies to all NIH Intramural Research Program (IRP) research projects generating genomic data **on or after August 31, 2015**, whether the research projects are new or were already ongoing by this date. Examples of large-scale genomic data include, but are not limited to, genome-wide association study (GWAS), Single Nucleotide Polymorphism (SNP) arrays, and genome sequence, transcriptomic, epigenomic, and gene expression data. For more information, please see Does GDS Apply to My Research.

Based on the state of the science, utility of the data for the research community, and the programmatic priorities of the NIH Institute or Center (IC) funding the research**,** individual ICs may choose to deposit data into an NIH-designated genomic repository from a project which is generating data on a smaller scale (i.e., one that does not meet the threshold defined in the Policy). If the investigator still wishes to share data in a genomic repository (e.g., because the journal is requesting it), they should check with the planned repository about the requirements. **If the IRB is being asked to provide assurances listed in the Institutional Certification, the investigator will be expected to share the completed Institutional Certification(s), the original research protocol, as applicable, and the plan for de-identifying the data prior to deposition, outlined in the Policy.** The IRB expects that the same human subjects protections requirements under the GDS policy, i.e., explicit consent for sharing genomic data, as applicable, will have been followed.

If an investigator wishes to deposit data from smaller scale projects or is unsure whether their research falls under the GDS Policy, they should consult with their Intramural Scientific Director (SD) or Genomic Program Administrator (GPA). GPAs are senior Federal employees of the NIH who serve as staff leads for the implementation of the GDS Policy within their ICs. They also facilitate the study registration and data submission processes for human NIH-designated controlled-access data repositories.

4. Applying for an Exception to Share Genomic Data Per the GDS Policy

In cases where it is anticipated that Institutional Certification criteria cannot be met (e.g. sharing might not protect the interests of individuals who are the subject of the data or consent was not obtained), and data cannot be shared as expected by the GDS Policy, investigators should state this in their Data Management and Sharing (DMS) Plan and indicate what data, if any, can be shared and how. In such cases, the DMS Plan and its elements will serve as the alternative data sharing plan as described in the GDS Policy. In some instances, the funding NIH Institute, Center, or Office (ICO) may need to determine whether to grant an exception to the data submission expectation under the GDS Policy. Studies provided exceptions to data submission under the GDS Policy will still be expected to be registered in the NIH Database of Genotypes and Phenotypes (dbGaP) for transparency.

5. NIH-Designated Genomic Repositories

IRP investigators must submit human genomic data and associated data (e.g., phenotype and exposure data) that are subject to the GDS Policy to an NIH-designated repository, meaning a repository maintained/managed or supported by the NIH either directly or through collaboration, unless the IC grants an exception. Examples of appropriate repositories include dbGaP, GEO, SRA, the Cancer Genomics Hub, etc. For more examples, please see Where to Submit Genomic Data. NIH-designated data repositories need not be the exclusive source for facilitating the sharing of genomic data. Investigators may also elect to submit data to a non-NIH-designated data repository in addition to an NIH-designated data repository. However, investigators should ensure that appropriate data security measures are in place and that confidentiality, privacy, and data use measures are consistent with the GDS Policy.

6. NIH Principal Investigator's Responsibilities Under the 2023 NIH Data Management and Sharing (DMS) Policy and the GDS Policy

The NIH Principal Investigator's responsibilities under the 2023 NIH DMS Policy and the GDS Policy include the following:

A. Develop a Data Management & Sharing (DMS) Plan and obtain approval of the Plan from the IC Scientific Director (SD) or his or her designee.
B. Complete and submit the applicable IRP Institutional Certification(s).
    i. Ensure that the type of access and data limitations you choose in the Institutional Certification(s) are consistent with the language in your protocol and consent form(s).
    ii. Share your DMS Plan, protocol, consent form(s) and Institutional Certification(s) with the IRB (or equivalent) so that they can determine whether the assurances in the Certification(s) are met.
    iii. Obtain SD or designee approval (and signature) of the Institutional Certification(s).
    iv. Provide the completed and signed Certification(s) to your IC GPA.
C. Before or at the time that data cleaning and quality control measures begin, register the study in dbGaP.
D. Make a copy of the genomic data and remove all identifiers from the data per the expectation in the GDS Policy, code the data and maintain a code key.
E. Submit the data to an NIH-designated repository per the applicable submission requirements. For more information, please see Submitting Genomic Data.

For IRP-specific information about implementation of the 2023 NIH DMS Policy, please see Intramural Data Sharing: 2023 Data Management and Sharing (DMS) Policy.

7. Developing and Obtaining Approval of Your Data Management & Sharing Plan

Under the 2023 Data Management and Sharing (DMS) Policy, NIH expects researchers to maximize the appropriate sharing of scientific data, taking into account factors such as legal, ethical, or technical issues that may limit the extent of data sharing and preservation. NIH requires all PIs planning to generate scientific data to prepare a DMS Plan that describes how the scientific data will be managed and shared. For more on what constitutes scientific data, see Research Covered Under the Data Management & Sharing Policy.

Recognizing the duplicative data sharing plans expected for an investigator subject to both the GDS Policy and the DMS Policy, NIH established a single Plan submission requirement for research subject to both (Implementation Changes for Genomic Data Sharing Plans Included with Applications Due on or after January 25, 2023). Beginning on and after January 25, 2023, NIH will expect one Data Management and Sharing (DMS) Plan from investigators subject to both the GDS and DMS Policies. Protocols subject to the GDS Policy should address GDS-specific considerations within the elements of the DMS Plan which are described in additional detail in section below. (Writing a Data Management & Sharing Plan).

As a result of the DMS Policy, even protocols not subject to the GDS Policy, that are submitted for initial or annual scientific review on or after January 25, 2023, must also include a separate DMS Plan. All research projects that do not require scientific review will still require a DMS Plan. In those cases, the plan should be submitted to the NIH Intramural Database (NIDB), which will be programmed to accept the DMS Plan materials. These plans must be submitted to NIDB, prior to the study being implemented.

Elements to Include in a Data Management and Sharing Plan When Subject to the GDS Policy:

A. **Data Type:** Genomic data expected to be shared under the GDS Policy should be described in this element, whether or not it meets the definition of scientific data.
B. **Data Preservation, Access, and Associated Timelines:**
    i. The name of the repository(ies) where the genomic data arising from the project will be archived. Investigators are expected to submit data to a repository acceptable under the Genomic Data Sharing Policy: See **Section 5** in this document and Where to Submit Genomic Data.
    ii. When the scientific data will be made available to other users and for how long. Identify any differences in timelines for different subsets of data to be shared. Human genomic data is expected to be shared according to NIH's Data Submission and Release Expectations, but no later than the end of the performance period (either at the time of first publication or at the end of the research project), whichever comes first.
C. **Access, Distribution, or Reuse Considerations**
    i. Informed consent expectations: See **Section 12** of this document.
    ii. Institutional Certifications and Data Sharing Limitation Expectations: In cases where it is anticipated that Institutional Certification criteria cannot be met (i.e., data cannot be shared as expected by the GDS Policy), investigators should state the institutional Certification criteria in their DMS Plan, explaining why the element cannot be met, and indicating what data, if any, can be shared and how to enable sharing to the maximal extent possible (for example, sharing data in a summary format). In some instances, the

NIH IC may need to determine whether to grant an exception to the data submission expectation under the GDS Policy. See **Sections 4 & 8** of this document.

iii. Genomic Summary Results: Investigators conducting research subject to the GDS Policy should indicate in their DMS Plan if a study should be designated as "sensitive" for the purposes of access to Genomic Summary Results (GSR). See **Section 8** in this document.

8. Completing the Institutional Certification

Within the IRP, investigators are required to complete and submit the applicable NIH Intramural Institutional Certification, in consultation with their GPA and obtain approval **prior to starting the study. If a project is being amended to add a genetic component and now must comply with the Policy, this step should be completed before the start of the new research activities**. The approved Data Management and Sharing (DMS) Plan should contain the information required to be able to complete the Institutional Certification.

Please carefully review Completing an Institutional Certification Form and Points to Consider for IRBs and Institutions in Submission and Secondary Use of Human Genomic Data before completing an Institutional Certification.

Type of Access to Data and Genomic Summary Results

The Intramural Institutional Certification states how individual-level data and genomic summary results (GSR) will be shared, i.e., controlled-access (data are available to an investigator for a specific project only if certain stipulations are met) or unrestricted access (data are accessible to anyone via public website).

Controlled Access

When controlled access is selected, the Certification must address whether the data can be shared for general research use; health/medical/biomedical purposes or use of the data must be related to the specified disease. For more information about data use limitations, please review Standard Data Use Limitations.

Genomic Summary Results

Effective November 1, 2018, the NIH Genomic Data Sharing Policy was expanded to enable unrestricted access to Genomic Summary Results (GSR). See Update to NIH Management of Genomic Summary Results Access. 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. The investigator should propose whether unrestricted access to GSR is planned or not, as part of the DMS Plan and the Certification Form.

Sensitive Genomic Summary Results

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. When possible, consultation with communities and study populations that are involved in the research may be appropriate to determine their perspectives about the balance of privacy concerns relative to the priority that many communities may have to support broad data sharing to advance research. If you

have questions about GSR or the sensitive designation, please see Access to Genomic Summary Results (GSR) or contact your NIH GPA or the GDS mailbox at GDS@mail.nih.gov.

Differences in the Requirements to Allow Sharing

The requirements vary depending on when the original specimens were collected, i.e., before or on or after the implementation date of the GDS Policy. If the specimens were collected before the implementation date, the requirements also vary depending on whether the specimens were collected with or without informed consent.

One Institutional Certification must be completed if any of the specimens were originally collected prior to August 31, 2015, and another must be completed, if specimens were collected after August 31, 2015 (i.e., when inclusion of the GDS language in a consent form became mandatory). Specifically, one Institutional Certification must be completed for each scenario described below:

1. Specimens were collected prior to August 31, 2015, with consent.
2. Specimens were collected prior to August 31, 2015, without consent.
3. Specimens were collected after August 31, 2015, with consent.

Also, if the investigator is sharing data from more than one protocol at the same time, he or she should **complete a separate Institutional Certification for each of the involved protocols**.

The specific requirements associated with each scenario are listed within each type of NIH Intramural Institutional Certification.

Role of the SD (or his or her Designee) and the IRB (or equivalent body)

As part of the Certification process, the Scientific Director (or his or her designee), in consultation with its IRB or an equivalent body, must provide assurances about appropriateness of the data being submitted as well as the secondary use of that data, based on the NIH GDS Policy and the consent of the original study participants, as applicable. The IRB (or equivalent body) will also determine whether the proposed access to the genomic summary results (GSR) is appropriate or not.

After review by the IRB/ designee, the IRP Institutional Certification must be approved and signed off on by your SD/designee, as the signing official, and the project's lead investigator. By signing the form, the SD/designee is providing certification for deposition of the data into an NIH-designated repository.

For more information about Institutional Certifications, please see About Institutional Certifications.

9. Obtaining NIH IRB (or equivalent) Review and Assurances for the Institutional Certification

Once the Institutional Certification(s) have been completed, it is the IRB's role to attest to various elements of the Certification in order for the investigator to be able to share the genomic data with an NIH-designated repository(ies). As part of this process, an IRB (or equivalent) representative must review the approved DMS Plan, all versions of the informed consent documents signed by the subjects whose data is being shared, and the data sharing information included in the Institutional Certification. The elements that the IRB must provide assurances for are listed in the Institutional Certification document.

In order to request review and assurances from the NIH IRB, the investigator should submit a separate request via email for each protocol involved in the planned sharing. The PI or Associate Investigator (AI)

should email this request to the IRB@od.nih.gov and attach all the documentation identified below. **In the body of the email, they should provide the IRB # for each protocol; note whether the original specimens were collected prior to the implementation date, after the implementation date or at both time periods; and name the repository that will be used for the sharing.** Please work with your GPA to make sure that you have all the necessary materials before emailing the IRB.

For each request, the investigator should attach the following documents to the email:

A. The appropriate Investigator Certification(s) that has been completed by the investigator and GPA for each protocol which covers the data you intend to submit. (Please note that Certification can remain unsigned until after the IRB review process is complete.)
B. A copy of the IRB-approved protocol, as applicable
C. A copy of the **approved** DMS Plan associated with the protocol
D. When applicable, copies of all versions of the consent documents for all protocols (associated with the data that you wish to share) with the language about sharing and future research involving genomic data highlighted. (Please note that the Intramural Research Program requirements are different depending on whether the specimens were originally collected **before or after the effective date, August 31, 2015**.)
    i. The investigator should first review all versions of the consent forms associated with the data that they plan to share to ensure they think that the consent requirements have been met to enable sharing the data.
    ii. If the appropriate consent language was not included in the consent forms signed by the applicable subjects, the investigator should instead contact IC leadership to determine if the importance of the data is such that they wish to waive this requirement and allow the data to be deposited. **This decision is not up to the IRB**.

A helpful flowchart, entitled NIH GDS Policy - Flowchart for Informed Consent Expectations, can be found on the NIH OHSRP website.

10. Submitting Data that Results from Multi-Site Research

An institutional assurance of the data submission must be provided for all data generated from contributing specimens, whether the data is generated from multiple study sites or a single site. The primary study site may submit one Institutional Certification indicating that it is providing certification on behalf of all collaborating sites. Alternatively, each site (which collected the specimens) may complete their own single-site Institutional Certification (see the Extramural Institutional Certification Form for non-NIH sites) and provide it to the lead investigator. The lead investigator should work with GPA from the funding NIH IC to discuss the appropriate certification, register the study, and submit data to an NIH-designated repository.

If an institution chooses to submit one Institutional Certification on behalf of multiple sites, it is certifying that, based on either its own review or assurance from other institutions, the expectations and conditions of the Institutional Certification(s) are met for the entire multi-site dataset. Further, the submitting institution should explicitly identify within the certification any data use limitations that apply to the submitted dataset or subsets of such data collected at all sites. In obtaining assurance from other sites in a multi-site study, the site (who will be submitting the data) should retain copies of any documents or information that it receives from the other sites.

Though the Institutional Certification allows for multiple collaborating institutions to be added to the same form, the NIH IRB will not provide assurances for genomic data that was generated from specimens collected under one or more protocols that were approved by external IRBs. In this case, the lead NIH investigator should request an Institutional Certification from each external site. The external institution will be responsible for obtaining assurances from the IRB of record (or equivalent body). After that point, it is up to the NIH IC whether they wish to submit one multi-site Institutional Certification or submit all single-site Institutional Certifications to the NIH-designated repository.

11. Addressing the GDS Policy in the IRB-Approved Protocol and Consent Form

All investigators doing research covered by GDS Policy must develop and have in place an approved DMS Plan **prior to start of the research** that is generating genomic data. NIH protocol templates for various types of studies (e.g., interventional clinical trials, behavioral and social science research, natural history and observational trials and secondary research protocol) can be found in the Protocol Templates and Forms page of the NIH OHSRP website. These templates include explanations of the relevant content that should be included in the sections of the protocol that address GDS compliance.

The IRB will review the protocol as well as the proposed consent language to determine if the protocol and consent form are consistent. For example, if the protocol, states that the GDS Policy is applicable to the study, any versions of the consent form, used to collect specimens on or after August 31, 2015, should include language about future research and broad sharing of genomic data. Please refer to the "Storage, Sharing and Future Research" sections of the NIH consent templates on the Consent Templates and Guidance page of the NIH OHSRP website. This section includes required consent language that must be included if the protocol involves deposition of genomic data into an NIH-designated repository per the GDS. For additional information, please see NIH Guidance on Consent for Future Research Use and Broad Sharing of Human Genomic and Phenotypic Data Subject to the NIH Genomic Data Sharing Policy.

12. Consent Requirements to Be Able to Submit Genomic Data in an NIH-Designated Repository

Depending on if the original specimens used to generate the genomic data were collected on or after the effective date of the policy, the consent form may or may not have to include specific content to allow submission of the data into a repository. In the NIH IRP, if the specimens were collected prior to August 31, 2015, without informed consent (e.g., not collected under an IRB-approved protocol or informed consent was waived), no consent requirements apply. If the specimens were collected prior to August 31, 2015, using a consent form, the consent form cannot contain any language that conflicts with a plan to submit de-identified genomic data to a genomic repository for broad sharing. If the specimens were collected on or after August 31, 2015, with informed consent, the applicable consent form must include language about future research and broad sharing. Please see the "Storage, Sharing and Future Research" section of the NIH consent templates on the Consent Templates and Guidance page for example language.

Please note that when investigators are conducting secondary research with de-identified/anonymized specimens, there is still a requirement under the policy to consider the language in the original consent form, associated with the collection of the specimens, when applicable. See H.7. under "Consent for Broad Sharing" in the Genomic Data Sharing FAQs. For data from specimens collected

before the effective date of the GDS Policy, there is no explicit requirement for consent. For specimens collected after the effective date of the GDS Policy, informed consent for future research uses and broad sharing of data is expected. If there are compelling scientific reasons that necessitate the use of genomic data from cell lines or clinical specimens that were created or collected after the effective date of the GDS Policy and that lack consent for research use and data sharing, the investigator should provide a justification in the Data Management and Sharing (DMS) Plan. The SD or designee will review the justification and decide whether to make an exception to the consent expectation.

13. Submitting Data to a Repository as Required by a Journal

When the investigators submit a manuscript involving genomic results to a journal, the author may be told that they are required to submit the genomic data to a repository to be able to publish. If the study was not otherwise required to comply with the GDS Policy, there is no expectation that the investigator meet all the requirements of the Policy to be able to share the data. However, if the specimens were collected on or after August 31, 2015, the implementation date of the GDS Policy in the NIH IRP, in most cases, OHSRP will expect that the investigators obtained consent for future research and sharing before the data is shared in a genomic repository. If the authors wish to share the data in dbGaP specifically, dbGaP will require a signed Institutional Certification to be submitted along with the deposition which will likely trigger the need to meet other aspects of the Policy.

14. Registering Your Study in dbGaP

All human genomic data studies covered by the GDS Policy must be first registered in dbGaP. Registration in dbGaP should occur by the time data cleaning and quality control measures begin, regardless of which NIH-designated data repository (or alternative repository if approved by the IC) will ultimately house and distribute the data. For more information, please see How to Register and Submit a Study in dbGaP.

15. Timeline and Requirements for Submitting Data in an NIH-Designated Repository

The Data Submission and Release Expectations addresses the timeline and expectations for data submission and release of data based on the type of data and level of processing that the data have undergone. Before sharing, the data should be de-identified according to expectations of the Common Rule and HIPPA (removing 18 identifiers) and coded, and the key to the code should be retained by the research team.

For a list of NIH-designated repositories, please see Where to Submit Genomic Data.

16. Additional Information about the GDS Policy

Additional information about implementation of the GDS Policy in the IRP can be found on the Genomic Data Sharing page of the NIH IRP Sourcebook. The NIH Office of Science Policy has also developed extensive Genomic Data Sharing FAQs that those generating genomic data should find very helpful to review.