Guidance: Genomic Data Sharing (GDS)

> Introduction

The NIH issued the Genomic Data Sharing (GDS) Policy effective January 25, 2015. The NIH policy provides an overview of the responsibilities for submitting, accessing, sharing, and securing genomic data and includes the expectations for institutional certification and IRB review.

This guidance applies to any research that is supported or funded by an ongoing or new NIH funding mechanisms (grants, cooperative agreements, contracts, Other Transactions, and intramural support) regardless of NIH funding level and also applies to research that is not NIH-supported, when the data sharing plan is to submit genotype/phenotype data to one of the following NIH supported repositories:

1. Databases of Genotypes and Phenotypes (dbGaP)
2. Gene Expression Omnibus (GEO)
3. Sequence Read Archive (SRA)
4. Cancer Genomics Hub (CaHUB)

See GDS website for further guidance, FAQs, updates, templates, etc.

> Institutional Certificate

NIH applications for submission of large-scale genomic data into NIH-designated repositories and databases for future subsequent research must include a data sharing plan which is developed by the principal investigator, reviewed by the IRB, and includes an Institutional Certificate assuring that the data are appropriate for sharing and consistent with the informed consents of the study participants from which the data was obtained.

The principal investigator will submit an Institutional Certificate to the Office of Research and Sponsored Programs (ORSP) to be reviewed during the Just-in-Time pre-award process (or the start of research for NIH intramural investigators) when genomic data generation is proposed. The Institutional Certificate should stipulate the appropriate secondary uses of data submitted to an NIH designated repository. The purpose is to assure that submission of data to an NIH-designated data repository is consistent with the NIH GDS Policy and with the informed consent of the original study participants.

The principal investigator will also submit an initial submission to the IRB for review.

> IRB Review

The IRB will:

1. Verify the protocol for the collection of genomic and phenotypic data is consistent with Department of Health and Human Services (DHHS) regulations on protection of human subjects;

2. Determine if data submission described in the principal investigator's data sharing plan is consistent with informed consent of study participants from which data was or will be obtained;
3. Consider the need for data use limitations necessary to minimize the potential for harm to individuals and their families, groups, or populations from disclosure of secondary research results. Consideration will include whether the proposed sharing of genomic data involves groups, communities, or populations with cultural or tribal considerations and whether additional data use limitations or consultation are warranted to minimize potential group harms or breaches of cultural expectations. and
4. Ensure that the principal investigator's plan for de-identifying datasets is consistent with the standards outlined in the IRB application and is consistent with HIPAA regulations.

IRB review may be conducted in a manner consistent with the expedited review procedure described by 45 CFR 46.110.11.

> IRB Submission

The principal investigator will submit an initial application and Protocol with the appropriate genomic data sharing plan to the IRB for review and include:

1. Intent to contribute data to databases of Genotypes and Phenotypes (dbGaP) or specified NIH-designated repositories;
2. Genotypic and phenotypic data that will be provided (as applicable);
3. Sources of genotypic and/or phenotypic data (e.g., all participants, a subset of individuals).
    a. In cases where data submission to an NIH-designated repository is not appropriate, provide a justification for an exception to the data sharing;
4. A written data use limitations statement for any data that should be excluded from sharing based on standard limitations;
5. Assessment of potential harm or risks to rights and welfare of individual participants, their families, and groups or populations (if applicable) as well as any safeguards to mitigate risks (e.g., Certificate of Confidentiality);
6. Plan for de-identifying datasets per IRB and HIPAA regulations; and
    a. All genomic datasets must be de-identified prior to public sharing or external disclosure, in accordance with IRB and HIPAA regulations and includes the removal of all 18 HIPAA identifiers and the assignment of a random, unique code retained securely at the Medical University of South Carolina.
    b. Data use and disclosure must adhere to the "minimum necessary" standard. Strict access controls and encryption protocols must be implemented for all genomic databases to safeguard participant privacy.
7. Process for obtaining informed consent for future research use and broad sharing of genomic and phenotypic data generated from the participant's specimen or cell line.
    a. The consent describes the type of repositories to which the data will be submitted (open/unrestricted, registered, controlled, or mixed) and

includes the process for participants to withdraw their data from the repositories should they choose to do so (data that has been distributed for approved research cannot be retrieved).

i. Open/Unrestricted: Repositories that store open-access or unrestricted access genomic data and consequently no special credentials are required for downloading data. These datasets are available for the public to access. Investigators that download unrestricted-access data from NIH-designated data repositories should not attempt to identify individual human research participants from whom the data were obtained.

ii. Registered Access: Repositories that allow access to their data only if users are registered with the repository. In addition, the repository might monitor the usage.

iii. Controlled Access: Repositories, such as Database of Genotypes and Phenotypes (dbGaP), that require credentialed users to apply for access to data.

iv. Mixed: Repositories that contain both open and controlled-access datasets.

b. If a participant does not consent to broad sharing of data, the data may not be shared.

> **Existing Studies**

For existing studies, the principal investigator must submit an IRB Amendment to contribute data to dbGaP or specified NIH-designated repository addressing items (1-7) above.

1. For archived specimens, include information on informed consent process and/or ALL versions of the consent form(s) signed by research participants.

> **Future Prospective Studies, Post 2015 collections, and/or Pre-2015 collections**

The IRB will determine if the data sharing plan is acceptable relative to the informed consent for future prospective collections, post 2015 collections, and/or pre-2015 collections:

1. Future Prospective Enrollment
    a. NIH expects researchers who intend to use research, clinical specimens collected, or cell lines created after 1/25/15 to generate human genomic data to obtain participants' consent for their data to be shared broadly for future research.
    b. If participant recruitment is continuing and the existing consent form does not address genomic data sharing, it should be revised and submitted with the Amendment for IRB review.
2. Post 1/25/15 existing collection
    a. For data from specimens collected before the effective date of the GDS Policy (1/25/15), NIH-designated data repositories may accept the submission and subsequent sharing of data, if the IRB finds that

submission of the data is appropriate and meets the criteria specified within the GDS Policy.

b. If consent was obtained for archived collections, the IRB will need to review all versions of the consent form(s) signed by research participants to determine whether submission of the data is appropriate.

3. For pre 1/25/15 collections that lack consent or the consent process/form(s) are not consistent with the proposed sharing plan, the IRB may take one or more of the following actions:

a. Require principal investigator to seek permission to share data by consenting or re-consenting research participants;

b. Apply standard informed consent waiver criteria where consent or re-consent is not practicable (e.g., pre 2015 archived leftover clinical or research specimens with no direct or indirect identifiers to allow donor to be contacted);

c. Require data be excluded from submission for specimens collected without consent or with a consent process/form(s) that prohibited sharing; and/or

d. Determine that the request is not consistent with the NIH GDS Policy, applicable laws (national, tribal, or state), federal regulations, or institutional policies. (See HRPP 4.13 Privacy and Confidentiality)

References:

1. NIH Genomic Data Sharing (GDS) Policy
2. GDS website
3. Institutional Certificate
4. Special Considerations for Genomics Research
5. d104-gds-guidance.pdf
6. NIH Genomic Data Sharing (GDS) Policy and the Genome-Wide Association Studies (GWAS) | Human Research Protection Program (HRPP)
7. NIH Human Genomic Data Sharing Policy FAQs
8. Guidance on NIH Genomic Data Sharing (GDS) Policy | Committee on the Use of Humans as Experimental Subjects
9. NOT-OD-25-159: Required Security and Operational Standards for NIH Controlled-Access Data Repositories
10. NIH Genomic Data Sharing (GDS) Policy and the Genome-Wide Association Studies (GWAS) | Human Research Protection Program (HRPP)