GUIDANCE ON GENOMIC DATA SHARING

Overview

This guidance applies to the NIH Genomic Data Sharing (GDS) Policy. Additional information can be found at the NIH Genomic Data Sharing (GDS) website and at NIH GDS FAQs. The NIH GDS affects investigators who are conducting NIH-funded large scale (identifiable and de-identified) genomic research and investigators who deposit genomic data or tissue into NIH repositories (required by some collaborators, publications and some funders including NIH). NIH has strict standards for IRB review and informed consent for the human genomic data they will accept for inclusion in public data repositories whether or not your project has NIH funding.

Even if you do not currently anticipate depositing genomic data or human tissue into NIH repositories please consider the FAQ Sections on Submission of Data to Controlled-Access Repositories and Consent for Broad Sharing found at: https://osp.od.nih.gov/scientific-sharing/genomic-data-sharing-faqs/, and the broad sharing guidelines discussed later in this guidance.

Your plans for depositing may change in the future and we strongly suggest maximizing the downstream use of tissue and/or data by including all of the required elements initially.

For existing research using or generating genomic data, the IRB is required to review investigators' submissions to NIH data repositories. The primary focus of the review is on whether informed consent is obtained from subjects in a manner that is consistent with NIH requirements for sharing genomic data and the data sharing plan is consistent with GDS policy. An Institutional Certification is needed in order to deposit data or tissue into NIH repositories. Please refer to the Institutional Certification section below for more information on what this is and how to obtain one.

For all NEW grant applications that request NIH funding for genomic research, a genomic data sharing plan that is consistent with GDS policy must be incorporated as part of the NIH application. If genomic data is being generated, the NIH GDS policy requires an Institutional Certification as part of the Just in Time submission, as well as a Certification at the time of data submission to a data repository. The GDS Policy requires the researcher to address the following elements prior to award:

- Genomic Sharing Plan
- Data Submission Expectation and Timeline
- Data Repository designation
- Access Level (Controlled or Unrestricted)
- Informed Consent
- Institutional Certification

To what does the GDS Policy apply?

Effective 1/25/15, the GDS Policy applies to the following:

- NIH-funded research that generates **large-scale** genomic data (e.g. SNP arrays, genome sequencing, RNA sequencing, transcriptomic, metagenomics, epigenomic and gene expression data, GWAS studies) from more than 100 individuals. The policy also applies to subsequent research studies that use this type of data (secondary use).

Does the NIH GDS policy apply in other situations?

- Yes, portions of the policy apply if you plan to submit genotype/phenotype data to the Database of Genotypes and Phenotypes (dbGaP) and other central repositories.

  ➔ Note that if the GDS Policy does not apply to your grant proposal, if and when you deposit sequencing data into national repositories the consent requirements of the Policy may still apply. Consent requirements may be found at this link on the NIH GDS Policy site.

  ➔ Many journals require broad sharing of genomic data with NIH or other central repositories. A link to a partial list of journal requirements may be found on this NIH site.

  ➔ NIH may expect submission of data from smaller scale research projects based on the state of the science, the programmatic priorities of the Institute funding the research, or the utility of the data for the research community.

Examples of large-scale genomic research projects that are subject to the GDS Policy are available from NIH in Supplemental Information to the National Institutes of Health Genomic Data Sharing Policy.

The GDS Policy does not apply:

- When the genomic data is generated without NIH funds and will never be deposited into a central database.
  - When NIH-funded projects involve instrument calibration exercises, statistical or technical methods development, or the use of genomic data for control purposes, such as for assay development.
  - Smaller studies (e.g., sequencing the genomes of fewer than 100 human research participants) are generally not subject to this Policy.

How do I determine whether my project involves Large Scale Genomic Sequencing?

The following tool can help determine whether your project is large scale genomic sequencing that is subject to the NIH Genomic Data Sharing Policy. Your Program Officer may request that you justify your decision.

If your project does falls into the categories described in the chart, we strongly recommend that investigators incorporate the guidelines below related to consent and broad sharing of genomic data. (Information about broad consent can be found on page 6 of this document).

Tool for Determining Large Scale Genomic Status

Type of Data	*From*	*From*
Human		
>300,000 variant sites	Genotyping, methylation, RNA	>1000 individuals
DNA Sequence	>1 gene or similar region	>1000 individuals
DNA Sequence	>100 genes or regions of similar size	>100 individuals
Sequence	>100 metagenomes or metatranscriptomes	Human microbiome
Animal		
>100,000 SNPs	Genotyping	>1 model organism species or strain
DNA Sequence	Whole exome or whole genome	>1 model organism species or strain
Gene expression	Transcriptome	1 or more model organism species or strain
Sequence	>100 metagenomes or metatranscriptomes	Model organism microbiome
Microbial		
Sequence	DNA or RNA	>100 isolates of infectious organisms
Cells		
DNA methylation	Comparison of genomewide methylated sites	>10 cell types
Other		
DNA methylation	Comparison of genomewide differential methylation at single-base resolution	Within an individual or across cell types within the same subject

What are the requirements at proposal stage if my NIH proposal will involve genomic data sharing?

Unless the Funding Opportunity Announcement states otherwise, applicants preparing NIH grant applications are expected to:

- Contact the appropriate NIH Institute or Center (IC) Program Official or Project Officer as early as possible to discuss Genome Data Sharing expectations and timelines that would apply to their proposed research.
- Develop, at minimum, a basic Genomic Data Sharing plan in the Resource Sharing Plan section of the funding application or proposal. Examples are available in the NIH Guidance for Investigators in Developing Genomic Data Sharing Plans. Outline in the budget section of the funding application the resources needed to prepare the data for submission to appropriate repositories.

- Please note that there are additional *Just in Time* requirements for these projects and you should initiate the IRB review of your genomic data sharing once your proposal has received a score in the fundable range.

Note: In situations in which the sharing of human data is not possible, applicants should provide a justification explaining why they cannot share these data and provide an alternative data sharing plan. Exceptions to NIH expectations for data submission to an NIH-designated data database will be considered on a case-by-case basis by the NIH.

What is an Institutional Certification and what is the process for securing one at the Pre-Award stage?

- Institutional certification is a document signed by an Institutional Official that certifies that the data sharing plan is consistent with GDS requirements. An Institutional Certification must accompany the submission of all large-scale human data to the NIH Database of Genotypes and Phenotypes (dbGaP). The Institutional Certification (for sharing human data), should also be provided to the funding NIH Institute or Center prior to award, along with any other Just in Time information (for extramural researchers).
- The process for getting an institutional certification includes:
    - PI must submit a request for Institutional Certification to the PHS IRB either in the initial application or as an amendment. (More information related to the IRB submission can be found later in this Guidance in the section: "**What information should I submit to the IRB when an Institutional Certification is required?").**
        - The IRB will review the data submission proposal and assure: that collection of genomic and phenotypic data is consistent with 45 CFR 46 (Common Rule); data submission and sharing are consistent with the informed consent; risks of data sharing to participants and their families were considered; risks of data sharing to groups or populations associated with the data were considered; the investigator's plan for de-identification meets HIPAA and DHHS standards.
        - Once review by the Partners IRB is complete, the certification will be sent to the Institutional Official for signature and the IRB will forward the completed certification to the PI. If this request is related to a *Just in Time* request for an NIH proposal, forward the certification to your Pre-Award Grant Administrator for submission to the NIH.

When is Institutional Certification required?

Institutional Certification is required, as noted above at the Pre-Award time period. In addition, Institutional Certification is required at the time of data submission (for NIH funded and non-NIH funded research). Some Institutional Certifications may require two letters if data submission is from specimens collected before and after the NIH GDS Policy (1/25/2015). The Institutional Certification templates may be found at: NIH GDS website.

What information should I submit to the IRB when an Institutional Certification is required?

- Link to the InfoEd grant proposal in your Insight eIRB protocol application.

➔ On the Forms Tab of the protocol or amendment: 1) add a New funding section (don't just edit a previous funding section), 2) use the Search function (search by InfoEd number, fund number, grant PI) to find the proposal, and 3) click the Add button to make the link. Once the amendment is approved, Research Management will be able to see the link.

- A copy of your Genomic Data Sharing plan is required for an NIH *Just in Time* request. If you anticipate that it will differ from the plan included in your grant proposal (linked via InfoEd to your protocol application), attach a copy of the revised plan. If you have an explicit deadline for an Institutional Certification, please contact the IRB Help Desk after you have submitted your protocol amendment or submission with the subject line "Genomic Institutional Certification."
- Copies of all consent forms for samples that will be sequenced or have been sequenced for this project. This includes all versions, if revised over the course of a study.
- A copy of the Consent Form Checklist for each consent group (cohort/source of tissue). The Consent Form Checklist is posted on Research Navigator.
- Your protocol/protocol summary should include the following:
    - State explicitly that you plan to deposit data and/or samples into an NIH data database; describe the genomic and phenotypic data that will be submitted
    - Indicate whether data will be collected PROSPECTIVELY or, if the PI is submitting a modification to an ongoing study, whether the investigator wishes to submit data that has ALREADY been collected. NIH expects that ALL data collected after 1/25/15 was consented according to the GDS-required criteria.
    - Specify if data shared with the NIH for GDS Policy will be for broad use or limited to specific diseases or conditions. Specify if you will be depositing data into a controlled-access database and/or an open-access database. If known, name the databases you anticipate depositing data in.
- Attach a completed copy of the appropriate Institutional Certification letter(s). These letters can be accessed on the NIH GDS Policy Institutional Certification page.

My genomic research project involves tissue and data that is not individually identifiable. Do I need to submit an IRB protocol for this research?

Yes**,** the GDS policy applies even if the research is considered exempt or 'not human subject research'. Therefore, an IRB review is needed. The Partners IRB requires a protocol submission in order to provide an 'exempt' or 'not human subject' determination. Refer back to the IRB submission section of this guidance.

This NIH questionnaire may be helpful when you are completing your grant proposal in regard to this issue.

I plan to use existing excess clinical samples and my project is NOT HUMAN SUBJECTS research or the IRB has granted a waiver of consent. Do I need to comply with the consent requirement?

After January 25, 2015, the consent requirements are applicable. You should consult with the IRB and review the NIH GDS Policy FAQs on this topic (Submission of Human Subject Data to Controlled Access Database Section). The Partners IRB needs to make this determination. A new protocol or amendment to an existing protocol is required for the IRB to review this issue.

Is large-scale genomic data generated from microbiomes considered human or non-human genomic data?

The data from microbes and microbiomes are considered non-human, if the human genomic data has been filtered and removed. Large-scale non-human genomic data includes data from microbes, microbiomes, and model organisms. An Institutional Certification is not required for depositing non-human genomic data into NIH data repositories.

What information should be included in consent forms to enable broad sharing of human genomic data?

- For studies initiated after January 25, 2015, the consent form should include a provision for genomic and phenotypic data to be used for future research and shared broadly, and a statement that a participant's individual level data will be shared through a controlled-access database. If you anticipate that data will be deposited into an unrestricted open-access data database, this should be specified in the consent form.
- For studies using genomic data from cell lines or clinical specimens that were created or collected after January 25, 2015, informed consent for future research and broad data sharing must be obtained, even if the cell lines or clinical specimens are de-identified.
- For studies conducted prior to January 25, 2015, the IRB will determine whether data submission is not inconsistent with the informed consent provided by the participant.
- For ongoing studies, or specimens previously collected, the IRB has developed a NIH GDS Policy Consent Checklist that delineates the informed consent elements to meet GDS Policy data submission requirements. For multi-center research, submit a copy of this checklist for each consent form version and also submit a copy of the consent form from each consent group(cohort). Submit this checklist to IRB at the time of your request for a GDS Institutional Certification.
- For new protocol submissions, The Partners IRB Tissue Repository Consent Form templates include all of the language required to meet the GDS Policy. You may find these templates on the Partners Research IRB Navigator at: templates.

My consent form may not be adequate, are there any exceptions to this policy?

If the IRB determines that an Institutional Certification cannot be provided due to inadequate consent, or other reasons, the NIH may grant an exception to the GDS Policy requirements if there is a compelling scientific reason. Please refer to the NIH GDS Policy FAQs section on this topic.

What are the implications for research funded before the effective date of the GDS Policy?

Although the GDS Policy does not apply to research submitted prior to the Policy's effective date, NIH, nonetheless, strongly encourages investigators to comply with the expectations outlined in the Policy. Investigators should provide an updated genomic data sharing plan to the funding IC in the research performance progress report. For studies involving human participants that were initiated before the Policy's effective date and used consent forms that do not meet the expectations of the GDS Policy, investigators are expected to plan to transition to a consent for future research uses and broad sharing, if possible, particularly for new or additional collections of specimens. There will be reasonable accommodation, determined on a case-by-case basis by the funding IC, for long-term projects ongoing at the time of the Policy's effective date to come into alignment with NIH's expectations for consent and data sharing. The goal is to bring these projects into alignment, to the extent possible, in a reasonable timeframe.

Which NIH-designated repositories are approved to deposit data?

The Genomic Data Sharing Policy expects that genomic research data from NIH-supported studies involving human specimens as well as non-human and model organisms will be submitted to an NIH-designated data repository. The NIH GDS website has a webpage dedicated to approved databases. You should check with your program officer regarding requirements of the specific database.

Who should I contact if I have questions?

For grant-related questions, please contact your Pre-Award Grant Administrator.

For Partners IRB process questions, please contact your IRB protocol administrator.

Links to NIH resources:

NIH Genomic Data Sharing (GDS) website.

NIH GDShttps://osp.od.nih.gov/scientific-sharing/genomic-data-sharing-faqs/ FAQs

NIH Guidance on Elements of Consent under the GDS Policy