

National Institutes of Health (nih.gov): NIH-GDS: Genomic Data Sharing

Data type: human genomic data and non-human genomic data

The GDS Policy applies to all NIH-funded 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, irrespective of funding level and funding mechanisms (e.g. grant, contract, cooperative agreement, or intramural support). NIH Institute or Centers (IC) may expect submission of data from smaller scale research projects based on the state of the science, the programmatic priorities of the IC funding the research, and the utility of the data for the research community.

Guidance:

- [NIH Genomic Data Sharing Policy](#)
- [NIH Genomic Data Sharing FAQ](#)
- [Supplemental Information to the NIH Genomic Data Sharing Policy](#)
- [NIH Office of Science Policy. NIH Genomic Data Sharing.](#)
- [NIH Office of Science Policy. Scientific Data Management.](#)

Genomic research advances our understanding of factors that influence health and disease, and sharing genomic data provides opportunities to accelerate that research through the power of combining large and information-rich datasets. <https://osp.od.nih.gov/scientific-sharing/policies/>

Guidance:

To promote robust sharing of human and non-human data from a wide range of genomic research and to provide appropriate protections for research involving human data, the National Institutes of Health (NIH) issued the *NIH Genomic Data Sharing Policy* (GDS Policy) on August 27, 2014 in the *NIH Guide Grants and Contracts* (available at <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-14-124.html>), and in the *Federal Register* (available at <https://federalregister.gov/a/2014-20385>) on August 28, 2014. The GDS Policy and related documents are available at:

- [Preamble to the NIH GDS Policy](#)
- [NIH GDS Policy PDF](#)
- [Supplemental Information to the NIH GDS Policy](#)
- [NIH Intramural Investigator Responsibilities under the NIH GDS Policy](#)
- [NIH Guide Notice on Implementation of the GDS Policy for NIH Grant Applications and Awards](#)
- [NIH Guide Notice on Development of Data Sharing Policy for Sequence and Related Genomic Data](#)
- [GDS Policy Overview](#)

Other NIH Data Sharing Policies

2007 Genome-wide Association Studies Policy

- [Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies \(GWAS\)](#)
- [NIH Modifications to GWAS Data Access](#)

Data repositories

Identify the data repositories to which the data will be submitted, and for human data, whether the data will be available through unrestricted or controlled-access.

Investigators should register all studies with human genomic data that fall within the scope of the GDS Policy in dbGaP by the time that data cleaning and quality control measures begin. After registration in dbGaP, investigators should submit the data to the relevant NIH-designated data repository (e.g., dbGaP, GEO, SRA, the Cancer Genomics Hub). NIH-designated data repositories need not be the exclusive source for facilitating the sharing of genomic data, that is, 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.

Non-human data may be made available through any widely used data repository, whether NIH- funded or not, such as GEO, SRA, Trace Archive, Array Express, Mouse Genome Informatics, WormBase, the Zebrafish Model Organism Database, GenBank, European Nucleotide Archive, or DNA Data Bank of

Japan.

Guidance:

- [NIH Genomic Data Sharing Policy](#)
- [NIH Genomic Data Sharing FAQ](#)
- [Supplemental Information to the NIH Genomic Data Sharing Policy](#)
- [NIH Office of Science Policy. NIH Genomic Data Sharing.](#)
- [NIH Office of Science Policy. Scientific Data Management.](#)
- [Database of Genotypes and Phenotypes \(dbGaP\)](#)
- [NCBI GEO Gene Expression Omnibus](#)
- [NCBI SRA Sequence Read Archive](#)
- [NIH National Cancer Institute. Genomic Data Commons.](#)

Data submission expectations and timeline

Investigators should submit large-scale genomic data as well as relevant associated data (e.g. phenotype and exposure data) to an NIH-designated data repository in a timely manner. Investigators should also submit any information necessary to interpret the submitted genomic data, such as study protocols, data instruments and survey tools. Genomic data undergo different levels of data processing, which provides the basis for NIH's expectations for data submission and timelines for the release of the data for access by investigators. These expectations and timelines are provided in the Supplemental Information. In general, NIH will release data submitted to NIH-designated data repositories no later than six months after the initial data submissions begins, or at the time of acceptance of the first publication, whichever occurs first, without restrictions on publication or other dissemination.

Guidance:

- [NIH Genomic Data Sharing Policy](#)
- [NIH Genomic Data Sharing FAQ](#)
- [Supplemental Information to the NIH Genomic Data Sharing Policy](#)
- [NIH Office of Science Policy. NIH Genomic Data Sharing.](#)
- [NIH Office of Science Policy. Scientific Data Management.](#)

Informed consent and institutional certification

Respect for, and protection of the interests of, research participants are fundamental to NIH's stewardship of human genomic data. The informed consent under which the data or samples were collected is the basis for the submitting institution to determine the appropriateness of data submission to NIH-designated data repositories, and whether the data should be available through unrestricted or controlled access.

For research that falls within the scope of the GDS Policy, submitting institutions, through their Institutional Review Boards (IRB's), privacy boards, or equivalent bodies, are to review the informed consent materials to determine whether it is appropriate for data to be shared for secondary research use. Specific considerations may vary with the type of study and whether the data are obtained through prospective or retrospective data collections. NIH provides additional information on issues related to the respect for research participant interests its "*Points to Consider for IRB's and Institutions in their Review of Data Submission Plans for Institutional Certifications*" (updated in 2016 to "*Points to Consider for Institutions and Institutional Review Boards in Submission and Secondary Use of Human Genomic Data under the National Institutes of Health Genomic Data Sharing Policy*").

Guidance:

- [NIH Office of Science Policy. Researchers & Institutional Certifications](#)
- [Points to Consider for Institutions and Institutional Review Boards in Submission and Secondary Use of Human Genomic Data under the NIH Genomic Data Sharing Policy](#)
- [NIH Standard Data Use Limitations](#)
- [NIH Genomic Data Sharing Policy for Human and Non-Human Data](#)

Exceptions to data submission expectations

In cases where data submission to an NIH-designated data repository is not appropriate, that is, the Institutional Certification criteria cannot be met, investigators should provide a justification for any data submission exceptions requested in the funding application or proposal. The funding IC may grant an exception to submitting relevant data to NIH, and the investigator would be expected to develop an alternate plan to share data through other mechanisms.

Guidance:

- [NIH Genomic Data Sharing Policy](#)
- [NIH Genomic Data Sharing FAQ](#)
- [Supplemental Information to the NIH Genomic Data Sharing Policy](#)
- [NIH Office of Science Policy. NIH Genomic Data Sharing.](#)
- [NIH Office of Science Policy. Scientific Data Management.](#)

Intellectual Property

NIH encourages patenting of technology suitable for subsequent private investment that may lead to the development of products that address public needs without impeding research. However, it is important to note that naturally occurring DNA sequences are not patentable in the U.S. Therefore, basic sequence data and certain related information (e.g. genotypes, haplotypes, *p*-values, allele frequencies) are pre-competitive. Such data made available through NIH-designated data repositories, and all conclusions derived directly from them, should remain freely available, without any licensing requirements.

NIH encourages broad use of NIH-funded genomic data that is consistent with a responsible approach to management of intellectual property derived from downstream discoveries, as outlined in the NIH *Best Practices for the Licensing of Genomic Inventions* and Section 8.2.3. Sharing Research Resources, of the NIH Grants Policy Statement. NIH discourages the use of patents to prevent the use of or to block access to genomic or genotype-phenotype data developed with NIH support.

Guidance:

- [NIH Genomic Data Sharing Policy](#)
- [NIH Genomic Data Sharing FAQ](#)
- [Supplemental Information to the NIH Genomic Data Sharing Policy](#)
- [NIH Office of Science Policy. NIH Genomic Data Sharing.](#)
- [NIH Office of Science Policy. Scientific Data Management.](#)
- [NIH Best Practices for the Licensing of Genomic Inventions: Final Notice](#)
- [NIH Sharing Research Resources](#)