POLICY AND PROCEDURES FOR DATA SUBMISSION TO NIH DESIGNATED DATA REPOSITORIES

POLICY STATEMENT

The NIH Genomic Data Sharing ("GDS") Policy requires the sharing of large-scale human and non-human genomic data generated from NIH-funded research, as well as the use of these data, to an NIH-designated data repository when such sharing is compatible with the consent provided by the participant. The Policy applies to all competing NIH grant applications, proposals, and contracts submitted to NIH for the January 25, 2015 deadline and thereafter. All submissions to an NIH designated data repository, whether required or voluntary, must include a certification by the responsible Institutional Official(s) of the submitting institution that the submission of data to the repository is appropriate and consistent with the NIH GDS Policy. At Harvard, the Institutional Official of the relevant IRB provides this certification. This policy and procedures guidance outlines the responsibilities of the Harvard investigator, the relevant IRB, and the Institutional Officials in submitting, reviewing and providing the required Institutional Certification for a submission to an NIH–designated data repository.

APPLICABILITY

The NIH GDS Policy applies to all Harvard investigators who:

- Submit NIH grant applications or proposals after the January 25, 2015 deadline to conduct research that will generate large-scale human or non-human genomic data or will use these data for subsequent research; or
- Plan to voluntarily submit any large-scale genomic data into one of the NIH-designated data repositories, even if the research itself is not NIH-supported.

The NIH GDS Policy also applies to the relevant Harvard IRBs and Institutional Officials.

DEFINITIONS

NIH GENOMIC DATA SHARING POLICY: ("GDS Policy") is the policy that large-scale human and non-human genomic data generated from NIH-funded research should be shared broadly and responsibly, through a central repository. In the case of human genomic data, the GDS policy requires that the IRB of a submitting institution review the informed consent under which human material was collected to determine whether it is appropriate for the data to be shared for secondary research use. Further information on the GDS Policy can be found at: http://gds.nih.gov

Under the GDS Policy, an **NIH-DESIGNATED DATA REPOSITORY** is any of the following:

- Database of Genotypes and Phenotypes ("dbGaP"),
- Gene Expression Omnibus ("GEO"),
- Sequence Read Archive ("SRA"), or the
- Cancer Genomics Hub

The human **genome** is all of the DNA contained in an organism or a cell, including both the DNA comprising chromosomes within the nucleus and the DNA in mitochondria.

Institutional Officials: At Harvard, the Institutional Officials include the University Chief Research Compliance Officer for the University Area IRB and the Harvard Medical School Chief of Research and Administrative Operations and the Harvard T.H. Chan School of Public Health Associate Dean for Regulatory Affairs and Research Compliance for the Harvard LMA IRBs.

Large-scale genomic data is defined by the NIH GDS policy as including genome-wide association studies ("GWAS"), single nucleotide polymorphisms ("SNP") arrays, and genome sequence, transcriptomic, metagenomics, epigenomic, and gene expression data.

Phenotype data are data on health conditions, behavioral characteristics, or measurable observable traits (such as blood pressure, alcohol consumption, cholesterol, or eye color) that are obtained during physical or psychological examinations and maintained in a medical or research record. Phenotype data may also include information about medical treatments, drug tolerance, and family medical history as well as responses to questionnaires.

Relevant IRB: The Harvard IRB that reviews research proposals from the PI's School. The Committee on the Use of Human Subjects (CUHS) serves as the Institutional Review Board for Schools located on the Cambridge and Allston campuses at Harvard. The two Harvard LMA IRBs serve as the Institutional Review Boards for research studies submitted by Harvard T.H. Chan School of Public Health (HSPH), Harvard Medical School (HMS), and Harvard School of Dental Medicine (HSDM) faculty and students. The submitting office for the Harvard LMA IRBs is the Office of Human Research Administration (OHRA) at Harvard Longwood Medical Area.

PROCEDURES

- **For Institutional Certification of a data submission to an NIH-designated data repository, a Harvard investigator must first submit to the relevant IRB a research summary that includes:**

    - A description of all data fields (genotype and phenotype) being submitted to the NIH-designated data repository

    - A copy of the consent form(s) used to enroll participants and collect their samples and phenotype data.
        - For data submissions where the Harvard investigator is submitting data to an NIH-designated data repository where Harvard was not the IRB that reviewed the protocol under which the original samples and phenotype data were collected:
            - The relevant IRB may accept certification from the IRB of record, in lieu of the original consents, or
            - The relevant IRB may request the original consents from the IRB of record and review the consent forms in accordance with the procedures outlined below, in the section entitled ***IRB Review and Verification,*** or
            - The relevant IRB may request documentation that the original collection

was consistent with 45 CFR 46.
    - If no consent is available, the Harvard investigator must consult with the relevant IRB to determine whether data may be submitted to an NIH designated repository.

  - A description of the PI's plan for de-identifying datasets for transmission to the NIH-supported data repository and how the key linking the identity of each study participant will be maintained.
  - A written statement by the Harvard investigator that the key to identifiers of study participants will **never** be shared by the Harvard investigator with NIH.

- **IRB Review and Verification:** For any proposed submission of phenotype data and genotype data to an NIH-designated data repository, the relevant IRB will review the documents provided by the PI to determine whether:

  - The protocol for the collection of genomic and phenotypic data is consistent with 45 CFR Part 46;
  - Data submission and subsequent data sharing for research purposes are consistent with the informed consent of study participants from whom the data were obtained[1];
    - For studies using data generated from human material collected **after** January 25, 2015 (the effective date of the GDS policy), the consent form must:
      - Include a discussion of future research uses and broad sharing of genotypic and phenotypic data.
      - Address whether participants' individual-level data will be shared through unrestricted or controlled-access repositories
      - Describe the NIH designated data repository, broad data sharing, and the risks associated with broad data sharing, and
      - Describe the likelihood of re-identification in the future.
    - For studies using data generated from human material collected **before** January 25, 2015, the IRB should review the consent form, if any, to ensure that the data submission is ***not inconsistent*** with the informed consent provided by the research participant.
      - If there is a consent, the IRB may consider the following questions:

        - Do the consent documents identify future research use of the specimens or data?
        - Do the consent documents address disposal of the specimens or data?
        - Do any of the consent forms contain statements such as "your data/specimens will not be shared" or "will only be seen by the

[1] The consent review portion of the Institutional Certification process may not occur if the data is derived from individuals who are deceased, or if the data is derived from material collected prior to the January 25, 2015 effective date and no consent is available. In such instances, Institutional Certification will be based upon the IRB's determination of whether:
  - The protocol for the collection of genomic and phenotypic data is consistent with 45 CFR Part 46;
  - To the extent relevant and possible, consideration was given to risks to groups or populations associated with submitting data to NIH-designated data repositories and subsequent sharing; and
  - The investigator's plan for de-identifying datasets is consistent with the standards outlined in NIH GDS Policy.

research team?"
- Did the original consent forms limit future use to specific projects, conditions, disease states, or to non-commercial research?

- Consideration was given to risks to individual participants and their families associated with data submitted to NIH-designated data repositories and subsequent sharing;
- To the extent relevant and possible, consideration was given to risks to groups or populations associated with submitting data to NIH-designated data repositories and subsequent sharing; and

- The investigator's plan for de-identifying datasets is consistent with the standards outlined in NIH GDS Policy.

- **IRB determination:** After reviewing the research protocol, data sharing plan, and the relevant consents, the IRB may determine the following:

  - Submission of the specified data to an NIH-designated data repository is appropriate;
  - Submission of the specified data to an NIH-designated data repository is not appropriate
  - Submission of the specified data to an NIH-designated data repository is not appropriate, but that an attempt to cure the deficiency is appropriate. Therefore, as a condition for submission, the IRB may:
    - Require revision of the consent form(s) to be consistent with a submission of data to the NIH-designated repository with re-consent of the research participants;
    - Permit the submission, but subject it to certain restrictions or limitations on use, as the relevant Harvard IRB may specify; or
    - Request additional information, as necessary;

INSTITUTIONAL CERTIFICATION

Once the IRB review is complete, the relevant IRB will forward the Institutional Certification form letter for signature by the IRB's Institutional Official. Signed letters will be sent to the Harvard investigator via email, who must then send it to NIH to satisfy the requirements of the NIH GDS policy for a submission of data to an NIH-designated data repository.

CONTACTS AND SUBJECT MATTER EXPERTS

The Office of the Vice Provost for Research: http:///www.vpr.harvard.edu
CUHS IRB: http://cuhs.harvard.edu/
Harvard LMA IRBs: http://www.hsph.harvard.edu/ohra/irb-operations/