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

Background

NIH-funded studies that generate large-scale human genomic data are subject to the NIH Genomic Data Sharing (GDS) Policy.[1] According to the GDS Policy, investigators who intend to use research or clinical specimens collected or cell lines created after January 25, 2015, to generate genomic data may only do so when informed consent processes explicitly discuss future research use and broad data sharing, even if the data are generated from specimens that are de-identified. NIH-designated data repositories will not accept genomic data derived from specimens or cell lines collected or created after January 25, 2015, without this type of consent.[2] NIH strongly encourages the broadest appropriate future use and sharing of genomic and phenotypic data.

NIH also recognizes that in some circumstances broad sharing may not be consistent with the consent of the research participants whose data are included in the dataset.[3] If the research that involves the generation of genomic and phenotypic data is part of a larger study, such as a clinical trial, and a participant declines to consent to future research use and broad sharing of their data, the participant should not be excluded from the larger study on that basis. If future research use and data sharing are intrinsic to the study, investigators may decline to enroll participants who are unwilling to provide consent for future research use and broad data sharing.

This guidance provides information to be tailored to individual studies and conveyed to prospective participants during the consent process in order to meet GDS Policy expectations. This document will be updated, as appropriate.

Guidance for Consent under the GDS Policy

In order to meet the expectations for future research use and broad sharing under the GDS Policy, the consent should capture and convey in language understandable to prospective participants information along the following lines:

- Genomic and phenotypic data, and any other data relevant for the study (such as exposure or disease status) will be generated and may be shared broadly and used for future research in a manner consistent with the participant's informed consent and all applicable federal and state laws and regulations.
- Prior to submitting the data to an NIH-designated data repository, data will be stripped of identifiers such as name, address, account and other identification numbers and will be de-identified by standards consistent with the Common Rule and HIPAA. Safeguards to protect the data according to Federal standards for information protection will be implemented.
- Access to de-identified, individual-level participant data will be controlled, unless participants explicitly consent to allow unrestricted access to and use of their data for any purpose.
- Aggregate study information (including genomic summary results) and study analyses may be shared in the scientific literature or through other public scientific resources, such as data

[1] NIH Genomic Data Sharing Policy, http://gds.nih.gov/03policy2.html

[2] For studies initiated before January 25, 2015, studies vary considerably in how sharing is addressed in consent documents. IRBs should ensure that the proposal for data submission is not inconsistent with the informed consent provided by the research participant. NIH will accept data derived from de-identified cell lines or clinical specimens created before the effective date that lack consent for research use if determined to be appropriate by the submitting institution.

repositories or other data sharing resources that provide broad or unrestricted access to the information.

- Because it may be possible to re-identify de-identified genomic data, even if access to data is controlled and data security standards are met, confidentiality cannot be guaranteed, and re-identified data could potentially be used to discriminate against or stigmatize participants, their families, or groups. In addition, there may be unknown risks due to computational methods, analytic technologies, or techniques (e.g., generation of information that could allow participants' identities to be readily ascertained).
- No direct benefits to participants are expected from any secondary research on de-identified individual-level data or genomic summary results that may be conducted.
- Participants may withdraw consent for research use of genomic or phenotypic data at any time without penalty or loss of benefits to which the participant is otherwise entitled. In this event, data will be withdrawn from any repository, if possible, but data already distributed for research use will not be retrieved.
- The name and contact information of an individual who is affiliated with the institution and familiar with the research and will be available to address participant questions.
- The privacy protections, and limitations of those protections, afforded by a Certificate of Confidentiality to individual-level data do not apply to summary results.

For research projects for which the IRB has granted a waiver of some or all of the required elements of informed consent under the relevant provisions of 45 CFR 46.116, or if consent is not required because the activity is not subject to 45 CFR 46, investigators will still need to seek and document consent for future use and broad sharing of genomic and phenotypic data to meet NIH expectations under the GDS Policy. At minimum, the information described above should be provided to prospective participants. Investigators may request exceptions to the NIH consent expectations for compelling scientific reasons. These policy expectations do not supplant or replace consent provisions in the existing or any future amended provision of 45 CFR 46.

Additional guidance for institutions and investigators can be found on the GDS Policy website at http://gds.nih.gov/index.html.