DCCPS NEW GRANTEES WORKSHOP

NIH GENOMIC DATA SHARING POLICY & GRANTEE REQUIREMENTS

Charlisse Caga-anan JD
Genomic Program Administrator
DCCPS, NCI

NIH GENOMIC DATA SHARING POLICY (NOT-OD-14-124)

Update to the NIH GWAS Data Sharing Policy (NOT-OD-07-088)

- Other NIH data sharing policies may also apply, e.g. for model organisms, $500k/year direct costs

Effective date: January 25, 2015

Applies to NIH-funded research that generates large-scale human or non-human genomic data

Data from funded research falling under the policy must be submitted to an NIH-designated repository, in compliance with policy requirements

- Once data is released, controlled access can be requested by scientific community via database of Genotypes and Phenotypes (dbGaP)

dbGaP GENOTYPES and PHENOTYPES

KEY QUESTIONS FOR GRANTEES AND GDS POLICY

Does my project fall under the NIH Genomic Data Sharing Policy?

What samples am I using?

- What type of informed consent was obtained for the future use and sharing of these samples?
- Has the IRB reviewed plan to submit data to NIH-designated repository, along with other policy points (e.g. risks of data sharing, plans for de-identification)

What do I need to include in my grant application? **Genomic Data Sharing Plan**

What do I need to provide prior to award? **Institutional Certification**

DOES MY PROJECT FALL UNDER THE GDS POLICY?

Large scale data meeting NCI thresholds for

1. Technology/platform used
2. Sample size
3. Human or non-human

OR

Smaller scale high priority data depending on

- State of the science
- Programmatic priorities of NCI
- Utility of the data for the research community

E.g. projects examining rare cancers, rare-cancer-related outcomes, or rare cancer subtypes

https://www.cancer.gov/gds/about-policy#ui-id-3

Example NCI Thresholds

https://www.cancer.gov/gds/about-policy

Type of Data	Data Examples	# of Human Specimens	# of Model Organisms
SNP array data from > 500k single nucleotide polymorphisms (SNPS)	GWAS Data	1,000	500
DNA or RNA sequence data from < 100 genes or regions of interest, or transcripts	Targeted Sequencing	1,000	500
DNA sequence data from ≥ 100 genes or regions of interest	Targeted Sequencing, Whole Exome Sequencing, Whole Genome Sequencing	100	50
Genome-wide RNA sequencing data	Transcriptomic Data	100	50

WHAT DO I NEED TO INCLUDE IN MY GRANT APPLICATION? **GENOMIC DATA SHARING PLAN**

Genomic Data Sharing Plan outlines sharing expectations under the GDS Policy

Investigators should include genomic data sharing plan in grant application "**Resource Sharing Plan,**" along with other required data sharing plans

GDSP does not replace other required data sharing plans

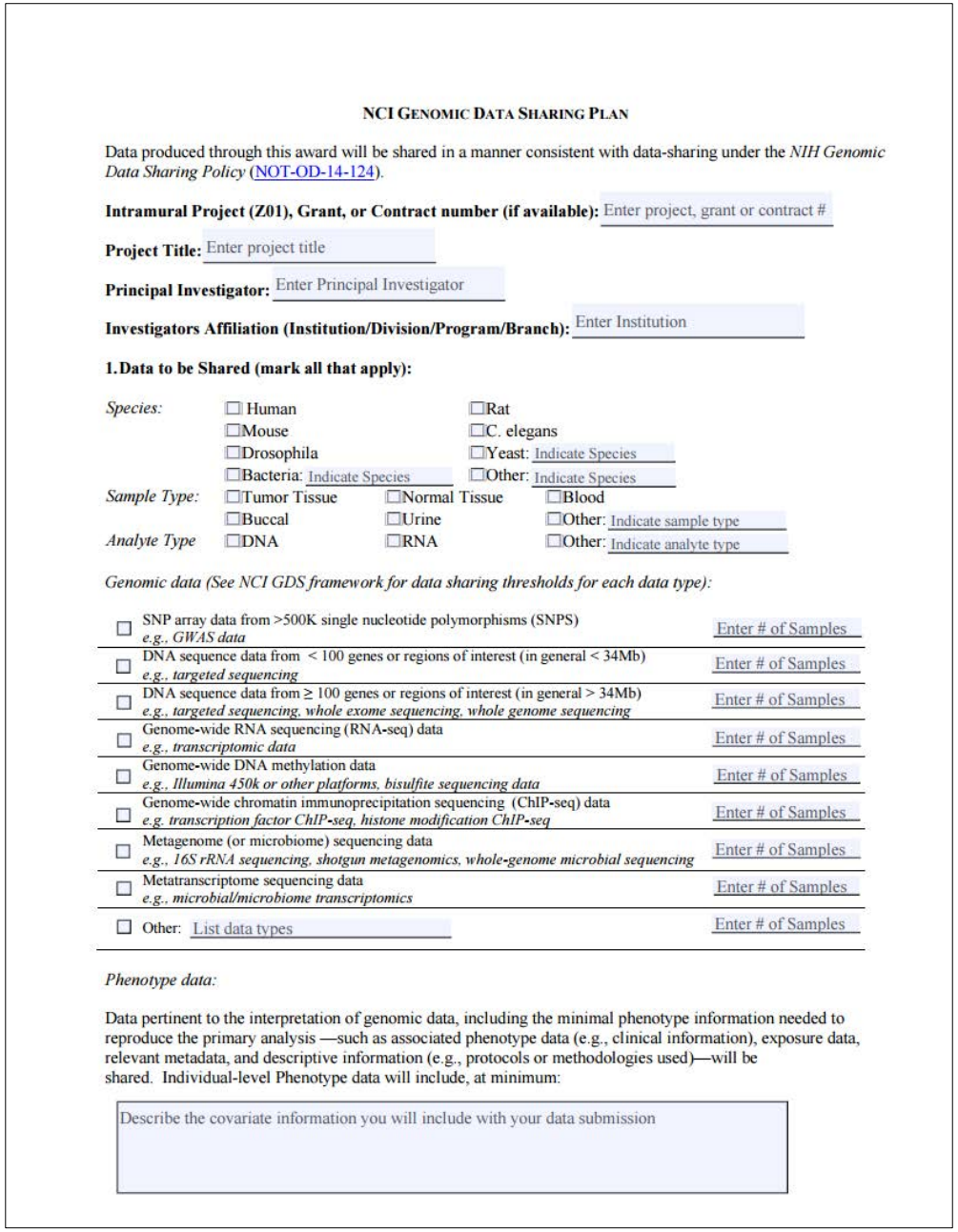
Elements of Genomic Data Sharing Plan:

1. *Data to be shared (data type)*
2. *Data Repository*
3. *Data Submission Timeline*
4. *IRB Assurance of the Genomic Data Sharing Plan*
5. *Appropriate Uses of the Data*

NCI GENOMIC DATA SHARING PLAN TEMPLATE

https://www.cancer.gov/grants-training/policies-process/nci-policies/genomic-data/submission/nci-dsp.pdf

NOTE: Genomic data sharing plan in application may take different form, as long as all elements are covered



NCI GENOMIC DATA SHARING PLAN

Data produced through this award will be shared in a manner consistent with data-sharing under the *NIH Genomic Data Sharing Policy* (NOT-OD-14-124).

Intramural Project (Z01), Grant, or Contract number (if available): Enter project, grant or contract #

Project Title: Enter project title

Principal Investigator: Enter Principal Investigator

Investigators Affiliation (Institution/Division/Program/Branch): Enter Institution

1. Data to be Shared (mark all that apply):

Species: ☐ Human ☐ Rat ☐ Mouse ☐ C. elegans ☐ Drosophila ☐ Yeast: Indicate Species ☐ Bacteria: Indicate Species ☐ Other: Indicate Species

Sample Type: ☐ Tumor Tissue ☐ Normal Tissue ☐ Blood ☐ Buccal ☐ Urine ☐ Other: Indicate sample type

Analyte Type ☐ DNA ☐ RNA ☐ Other: Indicate analyte type

Genomic data (See NCI GDS framework for data sharing thresholds for each data type):

☐ SNP array data from >500K single nucleotide polymorphisms (SNPS) *e.g., GWAS data*	Enter # of Samples
☐ DNA sequence data from < 100 genes or regions of interest (in general < 34Mb) *e.g., targeted sequencing*	Enter # of Samples
☐ DNA sequence data from ≥ 100 genes or regions of interest (in general > 34Mb) *e.g., targeted sequencing, whole exome sequencing, whole genome sequencing*	Enter # of Samples
☐ Genome-wide RNA sequencing (RNA-seq) data *e.g., transcriptomic data*	Enter # of Samples
☐ Genome-wide DNA methylation data *e.g., Illumina 450k or other platforms, bisulfite sequencing data*	Enter # of Samples
☐ Genome-wide chromatin immunoprecipitation sequencing (ChIP-seq) data *e.g. transcription factor ChIP-seq, histone modification ChIP-seq*	Enter # of Samples
☐ Metagenome (or microbiome) sequencing data *e.g., 16S rRNA sequencing, shotgun metagenomics, whole-genome microbial sequencing*	Enter # of Samples
☐ Metatranscriptome sequencing data *e.g., microbial/microbiome transcriptomics*	Enter # of Samples
☐ Other: List data types	Enter # of Samples

Phenotype data:

Data pertinent to the interpretation of genomic data, including the minimal phenotype information needed to reproduce the primary analysis —such as associated phenotype data (e.g., clinical information), exposure data, relevant metadata, and descriptive information (e.g., protocols or methodologies used)—will be shared. Individual-level Phenotype data will include, at minimum:

Describe the covariate information you will include with your data submission

WHAT DO I NEED TO PROVIDE PRIOR TO AWARD? **INSTITUTIONAL CERTIFICATION**

Institutional Certification provides official institutional assurance that data submission is consistent with requirements of GDS Policy

Institutional Certification(s) must cover all participants whose samples are used to generate data

- Multiple institutional certifications may be submitted, depending on how many sample-providing studies are involved

Elements of Institutional Certification:

- ***Consent***
- ***IRB review***
- *Data submission is consistent with applicable national, tribal, and state law and regulations, and with relevant institutional policies*
- *Identities of research participants will not be disclosed to NIH-designated data repositories*
- ***Data Use Limitation / Consent group:*** *Permitted future research use of the data (as consistent with the informed consent)*

INSTITUTIONAL CERTIFICATION TEMPLATE

Different templates available:

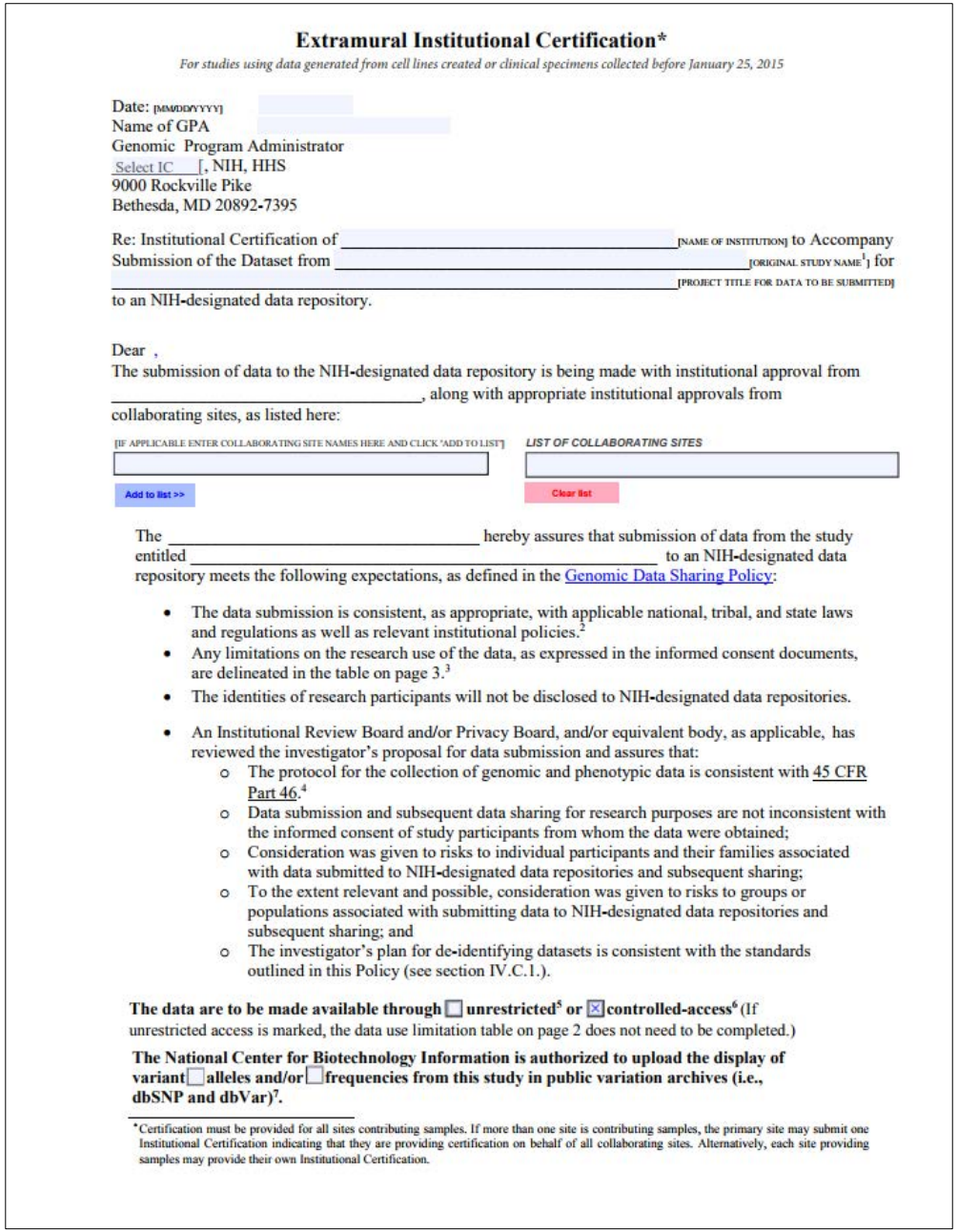
Samples collected Pre-January 2015

https://osp.od.nih.gov/wp-content/uploads/GDS_Extramural_Certification_Pre2015.pdf

Samples collected starting January 2015

https://osp.od.nih.gov/wp-content/uploads/GDS_Extramural_Certification.pdf

Contact GPA for questions on how to complete form



Extramural Institutional Certification*

For studies using data generated from cell lines created or clinical specimens collected before January 25, 2015

Date: [MM/DD/YYYY]
Name of GPA
Genomic Program Administrator
Select IC [, NIH, HHS
9000 Rockville Pike
Bethesda, MD 20892-7395

Re: Institutional Certification of _____ [NAME OF INSTITUTION] to Accompany Submission of the Dataset from _____ [ORIGINAL STUDY NAME[1]] for _____ [PROJECT TITLE FOR DATA TO BE SUBMITTED] to an NIH-designated data repository.

Dear ,
The submission of data to the NIH-designated data repository is being made with institutional approval from _____, along with appropriate institutional approvals from collaborating sites, as listed here:

[IF APPLICABLE ENTER COLLABORATING SITE NAMES HERE AND CLICK 'ADD TO LIST'] | LIST OF COLLABORATING SITES
Add to list >> | Clear list

The _____ hereby assures that submission of data from the study entitled _____ to an NIH-designated data repository meets the following expectations, as defined in the Genomic Data Sharing Policy:

- The data submission is consistent, as appropriate, with applicable national, tribal, and state laws and regulations as well as relevant institutional policies.[2]
- Any limitations on the research use of the data, as expressed in the informed consent documents, are delineated in the table on page 3.[3]
- The identities of research participants will not be disclosed to NIH-designated data repositories.
- An Institutional Review Board and/or Privacy Board, and/or equivalent body, as applicable, has reviewed the investigator's proposal for data submission and assures that:
  - The protocol for the collection of genomic and phenotypic data is consistent with 45 CFR Part 46.[4]
  - Data submission and subsequent data sharing for research purposes are not inconsistent with the informed consent of study participants from whom the data were obtained;
  - 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 this Policy (see section IV.C.1.).

The data are to be made available through ☐ unrestricted[5] or ☒ controlled-access[6] (If unrestricted access is marked, the data use limitation table on page 2 does not need to be completed.)

The National Center for Biotechnology Information is authorized to upload the display of variant ☐ alleles and/or ☐ frequencies from this study in public variation archives (i.e., dbSNP and dbVar)[7].

*Certification must be provided for all sites contributing samples. If more than one site is contributing samples, the primary site may submit one Institutional Certification indicating that they are providing certification on behalf of all collaborating sites. Alternatively, each site providing samples may provide their own Institutional Certification.

KEY PERSONNEL FOR GDS POLICY

Program Director (Program Official)

- Works with grantee investigators throughout life of grant to ensure data submission in compliance with policy, including collecting required documents

Genomic Program Administrator (Charlisse Caga-anan for DCCPS)

- Point of contact for GDS policy within division
- Provides guidance to Program Directors, provides final approval for institutional certifications, registers projects in dbGaP in preparation for data submission
- Liaison between the investigator and dbGaP staff

dbGaP Staff

- Works with grantee investigators to process submitted data

RESOURCES

NCI Genomic Data Sharing Policy website: https://cancer.gov/gds

- Data submission: https://www.cancer.gov/gds/submission

NCI DCCPS Genomic Program Administrator

- Charlisse Caga-anan, charlisse.caga-anan@nih.gov
- Jennifer Lee (Analyst), jennifer.lee7@nih.gov

Template Institutional Certifications

- Pre-2105 samples: https://osp.od.nih.gov/wp-content/uploads/GDS_Extramural_Certification_Pre2015.pdf
- Samples from January 2015 and after: https://osp.od.nih.gov/wp-content/uploads/GDS_Extramural_Certification.pdf

NIH Genomic Data Sharing Policy: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-14-124.html

QUESTIONS

SAMPLES: POLICY REQUIREMENTS FOR INFORMED CONSENT & IRB REVIEW

Informed Consent

Samples obtained prior to January 25, 2015

- Consent for future research and broad data sharing may vary across studies
- Submitting data and sharing for future research are not inconsistent with consent

Samples obtained on or after January 25, 2015

- Specific consent for future research use and broad data sharing is expected

IRB Review

Consent

Consider the risks to the participants and their families

Consider risks to groups or populations

Investigator's plan for de-identifying datasets is consistent with GDS Policy

- HIPAA and 45 CFR Part 46 standard

Collection of genomic and phenotypic data consistent with 45 CFR Part 46

WHO CERTIFIES?: OPTIONS

Grantee institution certifies for all

Grantee institution provides a multi-site certification on behalf of all participating studies

- Grantee will need to coordinate with participating studies to assure all points were complied with, provide data use limitation for each study

Participating studies provide institutional certifications

Each institution that provides samples used also provides its own certification

- Original institution's IRB oversaw required points in the institutional certification: sample collection, consent for future research uses of the samples/data
- Existing multi-site studies or consortia that collected samples may choose to provide a multi-site IC to cover entire study/consortia

Combination: Grantee institution may certify for some sites/studies, while others may certify for themselves