**NIMH Instructions for Submission of Data to NCBI (dbGaP) Databases**

All studies submitting data to dbGaP MUST provide a Genomic Data Sharing Policy [Institutional Certificate](https://osp.od.nih.gov/scientific-sharing/institutional-certifications/) (IC) letter through their AOR to their Program Officer (PO). Non-NIMH funded studies which are registering with dbGaP through NIMH should provide the IC letter to the NIMH Genomics inbox, [**NIMH.genomics.resources@mail.nih.gov**](mailto:NIMH.genomics.resources@mail.nih.gov)**,** with the email **Subject: “dbGaP study registration PI NAME.”** When completing the IC letters, they should be addressed to the NIMH Genomic Program Administrators (GPAs) Thomas Lehner and Anjene Addington.

After the submission of the IC letter, complete the registration of your study with dbGaP by providing the information outlined below to the NIMH Genomics inbox, [**NIMH.genomics.resources@mail.nih.gov**](mailto:NIMH.genomics.resources@mail.nih.gov)**,** with the email **Subject: “dbGaP study registration PI NAME Grant Number.”** For NIH funded studies your Program Officer should be cc’ed on this email.

For any registration issues please contact [NIMH.genomics.resources@mail.nih.gov](mailto:NIMH.genomics.resources@mail.nih.gov).

Once a study is registered, technical concerns with the data submission process should be addressed to the dbGaP help desk at [dbgap-sp-help@ncbi.nlm.nih.gov](mailto:dbgap-sp-help@ncbi.nlm.nih.gov).

**Example of information needed to complete dbGaP study registration:**

1. **Study Name**

* It must be less than 75 characters, if possible.
* It may contain acronyms.
* It should be identifiable as the same study name on the consent form(s) and IRB approval letter(S) used in enrolling the subjects.
* The name provided will be the name listed in dbGaP for the study.

Click or tap here to enter text.

2. **Funding Information** (Please list all relevant sources, both NIH and Non-NIH [e.g. Welcome Trust, Klarman foundation, NIMH, NHGRI, NIA, etc.])

Click or tap here to enter text.

3. **Target data delivery date (mm/dd/yyyy)** 4. **Target public release date(mm/dd/yyyy)**

Click or tap to enter a date. Click or tap to enter a date.

4. **Expected number of subjects**  Click or tap here to enter text.

5. **Expected size of data (in terabytes)** Click or tap here to enter text.

6. **Contents of Data Sets**

**General** **Sample Types**

Individual Phenotype  Germline

Individual Genotype  Tumor/Normal

Individual Sequencing  DNA

Supporting Documents  RNA

Metagenomic  Mitochondria

Protomic/Metabolomic  Microbiome

Images  From Repository: Specify the repository below if not NIMH Repository and Genomics Resource (NRGR)

Click or tap here to enter text.

**Array Data** **Genotypes**

SNP Array  Array derived Genotypes

Expression  Array CNV calls from microarray

Methylation  Array CNV calls derived from Sequencing

Genotype calls derived from Sequence

**Sequencing**  Somatic SNV (.MAF)

Whole Genome  Array CGH CNVs

Whole Exome

Targeted Genome **Analyses**

Targeted Exome  Association/Linkage Results

Whole Transcriptome  Array derived Expression

Targeted Transcriptome  RNA Seq derived Expression

Epigenomic Marks  Array derived Methylation

Sanger

16S rRNA

7. **Sequence Read Archive (SRA) submission expected?**

Yes  No

7a. **Expected storage location for Sequence Data**:

*\*If an SRA submission is expected, please indicate the expected storage location for the sequence data. This information is needed so that requesters and authorized users will know where the data is stored on the cloud and how to use it.*

NCBI  Google Cloud  Amazon Cloud

7b. **Cloud Service Provider:**

\**The person with this role is responsible for implementing the IC's cloud policy, security and billing. Your CIO should be able to define who this person is for this study/project.*

**First & Last Name E-mail Address**

Click or tap here to enter text. Click or tap here to enter text.

7c. **Data** **Steward**:

*\*The data steward is responsible for the data integrity, submission of metadata to SRA and answer questions from users about the data if they arise.*

**First & Last Name E-mail Address**

Click or tap here to enter text. Click or tap here to enter text.

8. **Include analyses for this study in CADA?**

\**Compilation of Aggregate Genomic Data is a collection of analyses across many dbGaP studies that can be accessed with a single Data Access Request.*

Yes  No

9. **Principal Investigator (PI)**

First Name Last Name

Click or tap here to enter text. Click or tap here to enter text.

Email Institution

Click or tap here to enter text. Click or tap here to enter text.

NIH Grant or Contract (If Applicable)

Click or tap here to enter text.

10. **PI Assistant/ Submitter** (This person can act on behalf of the PI to complete the registration and submit the data. Must have eRA account.)

First Name Last Name

Click or tap here to enter text. Click or tap here to enter text.

Email

Click or tap here to enter text.

11.  Controlled Access or  Unrestricted Access

12. **Acknowledgment Statement (for controlled access only)**

Once the data is ready to release all that is needed is a statement of how the investigator(s)

want to be acknowledged in publications that use their data set(s). This acknowledgement

statement will be put into a Data Use Certificate. All investigators who request access to the

data must sign the Data Use Certificate to gain approval.

**Examples of Acknowledgement Statements**:

***Example 1***: This study was conducted as part of a whole genome sequencing collaboration

between Harvard University and the Stanley Center for Psychiatric Research at the Broad

Research Institute. This work was supported by The Harvard Stem Cell Institute, The Howard

Hughes Medical Institute, The National Human Genome Research, and philanthropic gifts made to the Stanley Center

***Example 2***: The study “Joint Genome-Wide Gene Expression and GWAS Mapping in the MGS

Dataset” of dataset(s) used for the analyses described in this manuscript were obtained from

dbGaP found at http://www.ncbi.nlm.nih.gov/gap through dbGaP study accession numbers

phs[will be provided]. This Data Set was provided by Alan R. Sanders, M.D. Alan R. Sanders, M.D. and collaborators request that publications resulting from these data cite their original publication: Sanders AR, Göring HH, Duan J, Drigalenko EI, Moy W, Freda J, He

D, Shi J; MGS, Gejman PV. Transcriptome study of differential expression in schizophrenia.

Hum Mol Genet. 2013 22(24):5001-14. PMID: 23904455. Support for the collection of the data

sets was provided by grant RC2MH090030 awarded to Dr. Alan R. Sanders through NIMH.

Click or tap here to enter text.

11. Study Description (Short description of study aim and the data that will be provided)

Click or tap here to enter text.