

# Broad Steps to Publishing Data on dbGaP

Determine NIH Institute or Center (IC) sponsoring the study data

- If data was funded by NIH contract or grant, contact your NIH project official
- If data was not funded by the NIH, please contact [dbGaP Help](#)



Sponsoring NIH IC registers study with dbGaP

- Submitter obtains a Submission Certification from their IRB and Returns Certification to IC.
- See [http://grants.nih.gov/grants/gwas/gwas\\_ptc.pdf](http://grants.nih.gov/grants/gwas/gwas_ptc.pdf)
- Sponsoring IC verifies certification in the dbGaP study Submission system



NCBI creates a secure ftp login account for submission

- Submitter reviews dbGaP submission guide packet for study meta, and phenotype data
- Submitter uploads the study meta data , sample/subject, and phenotypic measures



dbGaP processes subject/sample information for coordination with

- Sequence Read Archive (SRA) for Next-Gen Sequence reads
- Gene Expression Omnibus (GEO) for expression or other non- GWAS array based data
- Genotyping Center for GWAS / SNP Chip data.



dbGaP publishes

- Public data to password-protected preview site for submitter review
- Public display of summary level data and
- Controlled Access release of individual level data

# dbGaP Submission Workflow

## Cohort / PI / Sample repository

### **Step 1 - Register Study with dbGaP**

- Submit study details and meta-data
- Obtain Institutional Submission Certification

### **Step 2 - Register Subject and sample ids**

- Submit de-identified subject Id and consent category for each study participant
- Submit sample id and subject id lookup file for each sample

### **Step 3a – Submit Phenotype data (See next page for details)**

- Variables for subjects or samples
- Data dictionaries

Actual samples Labeled with sample IDs



## Sequencing/Genotyping Center

### **Step 3b – Submit Genetic Data**

- GWAS
- Next Gen or Array data

## NCBI/ SRA/ dbGaP

### **Step 4 – Staging and QC of Data**

### **Step 5 – Accessioning and release of data**



Submission of Sequences with sample IDs registered in Step 2



# dbGaP Data File Submission

**WARNING: Do not submit any data files containing individual level data to dbGaP.**

Your study needs to be registered in dbGaP Registration system (please refer to page 1) prior to any data file submission. Then you'll be eligible to acquire a secure submission account for the person who will be uploading the files to dbGaP (if different from you).

Your study data files should meet dbGaP file format requirements in order to be published. To prepare your data files based on this requirements please [click here](#) to download dbGaP Submission guide packet. This packet contains templates and instructions on how to format your data files. You may start with the "Submission\_Guide\_Instructions.doc" file which provides detailed description of this packet.

**Note:** dbGaP required files for each single study are:

- 2\_dbGaP\_StudyConfig
- 5a\_dbGaP\_SubjectDS
- 5b\_dbGaP\_SubjectDD
- 6a\_dbGaP\_SubjectSampleMappingDS
- 6b\_dbGaP\_SubjectSampleMappingDD
- 4a\_dbGaP\_SampleAttributesDS (if applicable)
- 4b\_dbGaP\_SampleAttributesDD
- 7a\_dbGaP\_PedigreeDS (if applicable)
- 7b\_dbGaP\_PedigreeDD

Submitters can contact dbGaP at: [dbgap-help@ncbi.nlm.nih.gov](mailto:dbgap-help@ncbi.nlm.nih.gov)