**Steps for dbGaP Study Registration, Submission, and release of Data**

NIH Principal Investigator (PI) contacts his/her Program Officer/Director (PO/PD) and/or Institute or Center (IC) Genomic Program Administrator (GPA) to discuss the project, data sharing plan*, and data certification process (non-NIH-funded projects, should contact GDS@mail.nih.gov )

As early as possible in the research planning process, the PI discusses his/her data sharing plan and data use limitations with an Institutional Review Board

The sponsoring IC's GPA registers the study with minimal information in the dbGaP registration system and invites the PI to continue the registration by clicking on the "Invite PI" button

The sponsoring IC's GPA completes the registration of the study in the dbGaP registration system and invites the PI to view the study by clicking on the "Invite PI" button

dbGaP registration system sends an email to the PI or designee (cc to PI, GPA, and PO) with instructions on how to access and continue the study registration

dbGaP registration system sends an email to the PI or designee (cc to PI, GPA, and PO) with instructions on how to access and review the study

The PI completes the study registration by entering basic study information (e.g. target delivery and release dates, number of participants, data type, secondary contact information, consent groups, and an acknowledgement statement)

The PI logs into the dbGaP registration system to verify the accuracy of the study submission information

GPA works in collaboration with the PI to upload a signed Institutional Certification# and create a Data Use Certification

GPA verifies the study registration information using the "GPA review" section. Upon this verification, the registration system:
   a) Sends an automated email to the PI on how to submit files using the submission portal
   b) Changes the study status to "Completed by GPA" and locks the study from any further changes until it is released

PI or designee uploads the required study data to dbGaP submission portal

dbGaP staff starts processing the submitted study data in coordination with other NCBI teams

After submission to dbGaP is completed by the PI, the PI receives a study accession number for the study and a password-protected preview site to review the study prior to release

After the PI's approval of the preview site, dbGaP publishes a public display of summary-level data and controlled-access release of individual-level data

* In limited cases, following consultation with the GPA, an exception to deposit data in dbGaP may be granted by the NIH Institutional Official. Reasons for the exception will be made publicly available. An acceptable alternative data sharing plan is expected. Exceptions should be registered in dbGaP.

# The PI submits the certification for signature to the authorized institutional official who assures that the submission of the data to dbGaP meets conditions defined in the NIH GDS Policy.

# 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 files, please see the dbGaP Submission guide packet www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/GetZip.cgi?zip_name=dbGaP_SubmissionPackage.zip. 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:
1_dbGaP_StudyConfig
2a_dbGaP_SubjectPhenotypesDS
2b_dbGaP_SubjectPhenotypesDD
3a_dbGaP_SampleAttributesDS
3b_dbGaP_SampleAttributesDD
4a_dbGaP_SubjectDS
4b_dbGaP_SubjectDD
5a_dbGaP_SubjectSampleMappingDS
5b_dbGaP_SubjectSampleMappingDD
6a_dbGaP_PedigreeDS (only required if there are related individuals) 6b_dbGaP_PedigreeDD

Submitters can contact dbGaP at: dbgap-sp-help@ncbi.nlm.nih.gov