

**Attachment 1: Documentation of the dbGaP  
registration system, including changes since 2013 PRA  
approval**

# **Study Registration Web Forms**

Public reporting burden for this collection of information is estimated to vary from 30 to 60 minutes per response for initial registration of the required fields. **An agency may not conduct or sponsor, and a person is not required to respond to, a collection of information unless it displays a currently valid OMB control number.** Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to: NIH, Project Clearance Branch, 6705 Rockledge Drive, MSC 7974, Bethesda, MD 20892-7974, ATTN: PRA (0925-xxxx). Do not return the completed form to this address.

### Register New Study

*\*(asterisk symbol) indicates required fields.*

**Study Registration Information**

\* Study name  Study Accession  
phs  .v  .p

\* Study type:  Normal  Collection

\* Institutes or Centers supporting the study:

<input type="checkbox"/> NIGMS	<input type="checkbox"/> NEI	<input type="checkbox"/> NCRR	<input type="checkbox"/> NCI	<input type="checkbox"/> NIAID	<input type="checkbox"/> NIBIB	<input type="checkbox"/> NINDS
<input type="checkbox"/> NIMH	<input type="checkbox"/> NIEHS	<input type="checkbox"/> NIDCR	<input type="checkbox"/> NIDDK	<input type="checkbox"/> NHLBI	<input type="checkbox"/> NIAAA	<input type="checkbox"/> NHGRI
<input type="checkbox"/> NIA	<input type="checkbox"/> NIDA	<input type="checkbox"/> NIAMS	<input type="checkbox"/> NIMHD	<input type="checkbox"/> NIDCD	<input type="checkbox"/> NLM	<input type="checkbox"/> OD
<input type="checkbox"/> NICHD	<input type="checkbox"/> NCATS	<input type="checkbox"/> NINR	<input type="checkbox"/> NCCAM	<input type="text"/>		

Admin. IC:  \* Primary GPA   
NLM  Jane Ye

Target data delivery date  (mm/dd/yyyy) (YYYY-MM-DD) Target public release date  (mm/dd/yyyy) (YYYY-MM-DD) Estimated number of study participants

Please check all data types expected for this study:

<b>General</b> <ul style="list-style-type: none"><li><input type="checkbox"/> Individual Phenotype</li><li><input type="checkbox"/> Individual Genotype</li><li><input type="checkbox"/> Individual Sequencing</li><li><input type="checkbox"/> Supporting Documents</li><li><input type="checkbox"/> Metagenomic</li><li><input type="checkbox"/> Protomic/Metabolomic</li><li><input type="checkbox"/> Images</li></ul>	<b>Sample Types</b> <ul style="list-style-type: none"><li><input type="checkbox"/> Germline</li><li><input type="checkbox"/> Tumor/Normal</li><li><input type="checkbox"/> DNA</li><li><input type="checkbox"/> RNA</li><li><input type="checkbox"/> Mitochondria</li><li><input type="checkbox"/> Microbiome</li><li><input type="checkbox"/> From Repository</li></ul>	<b>Array Data</b> <ul style="list-style-type: none"><li><input type="checkbox"/> SNP Array</li><li><input type="checkbox"/> Expression Array</li><li><input type="checkbox"/> Methylation Array</li></ul>	<b>Genotypes</b> <ul style="list-style-type: none"><li><input type="checkbox"/> Array derived Genotypes</li><li><input type="checkbox"/> CNV calls from microarray</li><li><input type="checkbox"/> CNV calls derived from Sequencing</li><li><input type="checkbox"/> Genotype calls derived from Sequence</li><li><input type="checkbox"/> Somatic SNV (.MAF)</li><li><input type="checkbox"/> Array CGH CNVs</li></ul>	<b>Sequencing</b> <ul style="list-style-type: none"><li><input type="checkbox"/> Whole Genome</li><li><input type="checkbox"/> Whole Exome</li><li><input type="checkbox"/> Targeted Genome</li><li><input type="checkbox"/> Targeted Exome</li><li><input type="checkbox"/> Whole Transcriptome</li><li><input type="checkbox"/> Targeted Transcriptome</li><li><input type="checkbox"/> Epigenomic Marks</li><li><input type="checkbox"/> Sanger</li><li><input type="checkbox"/> 16S rRNA</li></ul>	<b>Analyses</b> <ul style="list-style-type: none"><li><input type="checkbox"/> Association/Linkage Results</li><li><input type="checkbox"/> Array derived Expression</li><li><input type="checkbox"/> RNA Seq derived Expression</li><li><input type="checkbox"/> Array derived Methylation</li></ul>
---	--	---	---	--	---

New categorization of data types for ease of completion and greater consistency among submissions.

## Sequence Read Archive (SRA)

\* Sequence Read Archive submission is expected:

Choose a value ▾

## dbGaP exception

This study was granted an exception from broad data sharing expected in the GWAS/GDS policy due to:

Legal Restrictions

Consent Limitations

Other:

File type	File name
GWAS exception	<input type="button" value="Choose File"/> No file chosen

## Principal Investigator (PI)

First Name  Last Name  Email  NIH grant or Contract #    Account Type

PI institution

PI Assistant/Submitter:

First Name  Last Name  Email    Account Type

New section for inclusion of PI assistant contact information, so all correspondence do not have to go through the PI.

## Certification

- Controlled Access  or  Unrestricted
- No Submission Certification Yet
- Provisional Submission Certification
- Submission Certification verified by IC
- Study is prior to the GWAS policy or is cleared for submission by the sponsoring IC

## Program Officer (PO) assigned to this study

No PO is assigned to this study

First Name  Last Name  Email  NED#

[Register Study](#)

[Back to studies list](#)

Policy

- Display research statement
- Display public summary
- \* Publication embargo interval
  - No embargo interval

Deletion of additional questions on publication embargo because it is no longer applicable.

Additional email addresses to be notified when a DAR is approved

Add Another Email

The text of the acknowledgement statement provided through the text field below will be incorporated into the draft DUC. Alternatively the acknowledgment statement can be provided from a file. Simply click [here](#) and upload the file. The link to uploaded file will be included in the draft DUC.

Acknowledgement Statement:

IC Specific Term of Data Access:

Added ability to auto-create the DUC from existing information in the submission package.

Save Data and Compose Draft DUC

Before uploading the DUC to the system, you may want to create a draft DUC using the "Save Data and Compose Draft DUC" button above. The resulting draft DUC incorporates study specific information into the Model DUC. It is thus a convenient starting point to compose the final DUC.

NOTE: The "Acknowledgement Statement" and "IC Specific Term of Data Access" field values in the section are presented to users only through the draft DUC. It is suggested to use the draft DUC to composed the final DUC. Otherwise, you would have to make sure that the field values are manually included in the final DUC.

File type	File name
Submission Certification	<input type="button" value="Choose File"/> No file chosen <input type="button" value="Remove"/>
Data Use Certification	<a href="#">Genomic_Sequencing_of_Colorectal_Adenocarcinoma_DUC.docx</a> <input type="button" value="Change"/> <input type="button" value="Remove"/>
Study Description	<input type="button" value="Choose File"/> No file chosen

New fillable PDF for the submission certifications (slides 6-7).

Add Another Submission Certification

## Extramural Institutional Certification\*

For studies using data generated from cell lines created or clinical specimens collected after 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<sup>1</sup>] 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.<sup>2</sup>
- Any limitations on the research use of the data, as expressed in the informed consent documents, are delineated in the table on page 3.<sup>3</sup>
- 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](#).<sup>4</sup>
  - Data submission and subsequent data sharing for research purposes are consistent 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<sup>5</sup> or  controlled-access<sup>6</sup> (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)<sup>7</sup>.

\* 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.

New fillable PDF for the  
submission certifications.

Sincerely,  
Investigator:  
Name: \_\_\_\_\_ Title: \_\_\_\_\_

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

Authorized Institutional Official:<sup>8</sup>  
Name: \_\_\_\_\_ Title: \_\_\_\_\_

Signature: \_\_\_\_\_ Date: \_\_\_\_\_

<sup>1</sup> Original Study Name should reflect the name of the original IRB-approved study (e.g. cohort or case-control study, clinical trial) under which participants provided informed consent and biospecimens were collected (e.g. Nurses' Health Study, Framingham Heart Study).

<sup>2</sup> For the submission of data derived from cell lines or clinical specimens lacking research consent that were created or collected before the effective date of this Policy, the Institutional Certification needs to address only this item.

<sup>3</sup> For guidance on clearly communicating inappropriate data uses, see NIH Points to Consider in Drafting Effective Data Use Limitation Statements, [http://www.nih.gov/pd/NIH\\_PTC\\_in\\_Drafting\\_DUL\\_Statements.pdf](http://www.nih.gov/pd/NIH_PTC_in_Drafting_DUL_Statements.pdf)

<sup>4</sup> 45 CFR Part 46. Protection of Human Subjects. See <http://www.gpo.gov/fdsys/pkg/CFR-2013-title45-vol1/html/CFR-2013-title45-vol1-part46.xml>

<sup>5</sup> Data made publicly available to anyone.



<sup>6</sup> Data made available for secondary research only after investigators have obtained approval from NIH to use the requested data for a particular project.

<sup>7</sup> The Single Nucleotide Polymorphism Database (dbSNP) is a public archive for genetic variation (apparently neutral polymorphisms, polymorphisms corresponding to known phenotypes, and regions of no variation) within and across species. The Database of Genomic Structural Variants (dbVar) is a collection of genomic structural variation data, typically 50 nucleotides in length or longer, for different organisms. For more information, see: [http://www.ncbi.nlm.nih.gov/variation/dbSNP\\_dbVar\\_FAQ/](http://www.ncbi.nlm.nih.gov/variation/dbSNP_dbVar_FAQ/)

<sup>8</sup> A senior official at an institution who is authorized to enter the institution into a legally binding contract and sign on behalf of an investigator who plans to submit data to NIH, e.g., Dean, Vice President for Research.



Consent Groups

Consent Group Type  Select One 

Does this consent group require IRB approval? Choose a value 

- Publication required
- Collaboration required
- Not-for-profit use only
- Methods
- Genetic Studies Only
- Related disorders

Consent Group's Abbreviation   Consent Group's Title

Data Use Limitations

DAC  To add new DAC please go to [DACs section](#) (all unsaved data will be lost)

Categories for consent groups now provided in lieu of free text box for ease of completion and greater consistency among submissions.

Data Use Limitations are autopopulated based on the consent groups selected.