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.

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	Choose File No file chosen

Principal Investigator (PI)

First Name | Last Name | Email | NIH grant or Contract # | Look up | Remove | Account Type
Any

PI institution

PI Assistant/Submitter:

First Name | Last Name | Email | Look up | Remove | Account Type
Any

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

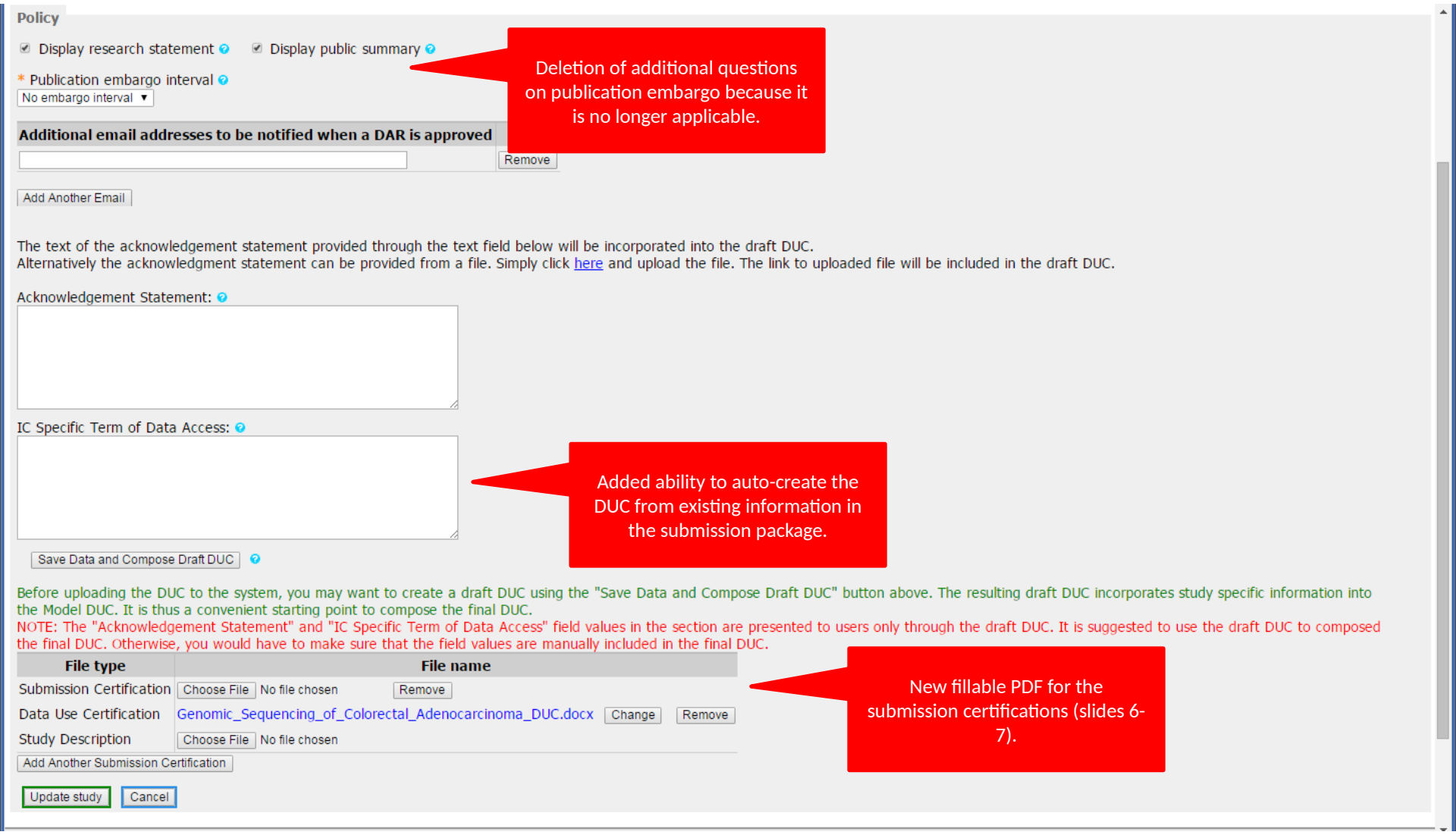
No Submission Certification Yet
Provisional Submission Certification
Submission Certification verified by IC
Study is prior to the GWAS policy or is cleared for submission by the sponsoring IC

Program Officer (PO) assigned to this study

No PO is assigned to this study

First Name | Last Name | Email | NED# | Look up | Remove

Register Study | Back to studies list


Policy
Display research statement
Display public summary
* Publication embargo interval
No embargo interval
Additional email addresses to be notified when a DAR is approved
Remove
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:
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
Choose File
No file chosen
Remove
Data Use Certification
Genomic_Sequencing_of_Colorectal_Adenocarcinoma_DUC.docx
Change
Remove
Study Description
Choose File
No file chosen
Add Another Submission Certification
Update study
Cancel
Deletion of additional questions on publication embargo because it is no longer applicable.
Added ability to auto-create the DUC from existing information in the submission package.
New fillable PDF for the submission certifications (slides 6-7).

New fillable PDF for the submission certifications.

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[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 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[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.

Sincerely,
Investigator:
Name: _____ Title: _____

Signature: _____ Date: _____

Authorized Institutional Official:[8]
Name: _____ Title: _____

Signature: _____ Date: _____

[1] 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).

[2] 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.

[3] For guidance on clearly communicating inappropriate data uses, see NIH Points to Consider in Drafting Effective Data Use Limitation Statements, http://gwas.nih.gov/pdf/NIH_PTC_in_Drafting_DUL_Statements.pdf

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

[5] Data made publicly available to anyone.

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

[7] 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 Varian (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/.

[8] 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.

Extramural Institutional Certification

For guidance on drafting data use limitations, please refer to the NIH Points to Consider in Drafting Effective Data Use Limitation Statements found at: http://gds.nih.gov/pdf/nih_ptc_in_drafting_dul_statements.pdf. Data use limitations are developed based on the original informed consent from the participant.

Data Use Limitations (will be used in dbGaP to create Consent Groups)

General Research Use	GRU	Use of the data is limited only by the terms of the Data Use Certification:these data will be added to the dbGaP Collection.
Health/Medical/Biomedical	HMB	Use of the data is limited to health/medical/biomedical purposes, does not include the study of population origins or ancestry.
Disease-specific [list disease]	DS	Use of the data must be related to the specified disease.
Other		[ENTER CUSTOMIZED TEXT, IF APPLICABLE]

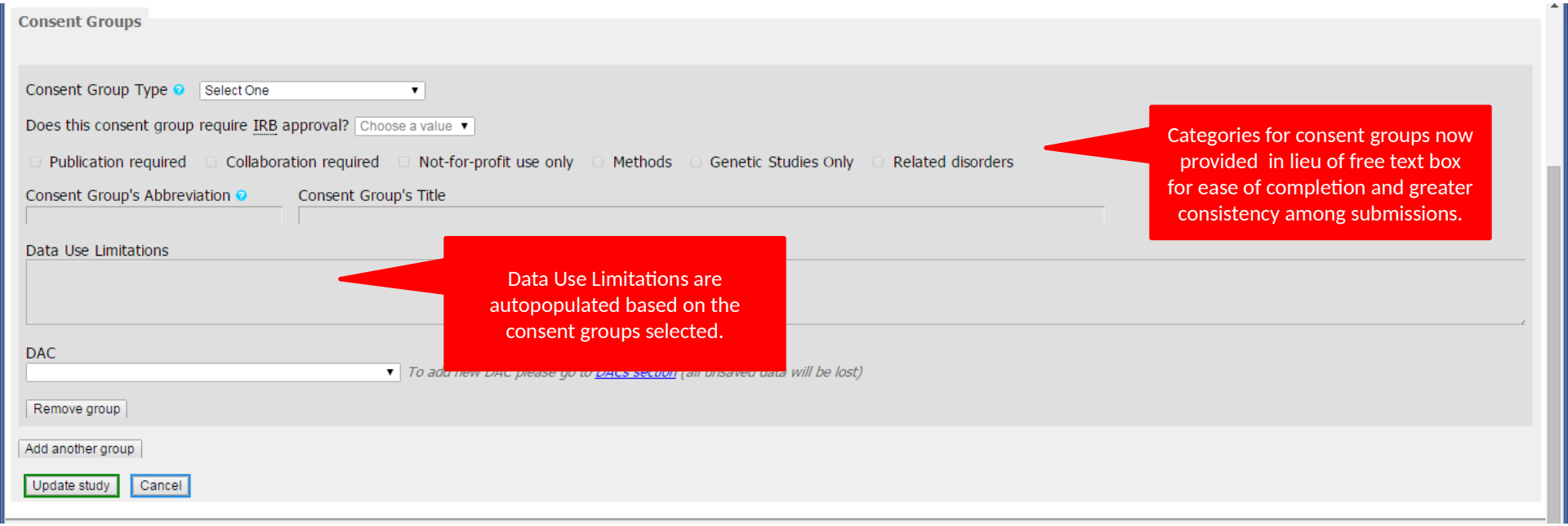
New standard data use limitations as part of the fillable PDF.

Data Use Limitation Modifiers

IRB approval required	IRB	Requestor must provide documentation of local IRB approval.
Publication required	PUB	Requestor agrees to make results of studies using the data available to the larger scientific community.
Collaboration required	COL	Requestor must provide a letter of collaboration with the primary study investigator(s).
Not-for-profit use only	NPU	Use of the data is limited to not-for-profit organizations.
Methods	MDS	Use of the data includes methods development research (e.g., development of software or algorithms)
Genetic studies only	GSO	Use of the data is limited to genetic studies only.

Using the tables above, please indicate in the form below the consent group(s) for each collaborating study site. Use one row per consent group.

Collaborating Site Name	Data Use Limitation	Data Use Limitation Modifiers
Eg: Cold Cohort Study	*Health/Medical/Biomedical*	IRB☐ PUB☐ COL☐ NPU☐ MDS☐ GSO☐
Eg: Cold Cohort Study	*Disease Specific Research [Lung Cancer]*	IRB☒ PUB☐ COL☐ NPU☒ MDS☐ GSO☐
	Select consent group title	IRB☐ PUB☐ COL☐ NPU☐ MDS GSO
	Select consent group title	IRB☐ PUB☐ COL☐ NPU☐ MDS GSO
	Select consent group title	IRB☐ PUB☐ COL☐ NPU☐ MDS GSO
	Select consent group title	IRB☐ PUB☐ COL☐ NPU☐ MDS GSO
	Select consent group title	IRB☐ PUB☐ COL☐ NPU☐ MDS GSO
	Select consent group title	IRB☐ PUB☐ COL☐ NPU☐ MDS GSO


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
Remove group
Add another group
Update study
Cancel
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.