NIH **U.S. National Library of Medicine**
National Center for Biotechnology Information

Data and Other Information Related to Controlled Access

Authorized Access Related Information

Where to find the information related to Authorized Access of a study?

The information related to the Authorized Access can be found from the "Authorized Access" section on study page of the dbGaP public website. From the study page (such as this one), there is a link named "Authorized Access" right under the "Variables" tab on top of the page. The link is a shortcut to the section. Please see the screenshot of the section below.

The following information can be found from the section:

1. Email address of Data Access Committee (DAC) of the study
2. Data Use Certificate (DUC) of the study.
3. The name of the consent group.
4. Data use restriction of the consent group.
5. IRB requirement of the consent group.
6. The link to study-report and data manifest for the data available through the Authorized Access System.

Here is a real example of the section. You need mouse over the help ('?') icon to see the consent language of each consent group. (**09/27/2012**)

IRB Requirement

Where can I find the information about IRB requirements of dbGaP datasets?

Please see here for the answer. (**07/03/2012**)

Data Use Restrictions

Where can I find data use limitation information of a study?

The dbGaP data are organized into consent groups which consist of all the data from study participants who have agreed to the same data use as specified in the informed consent for the study. Data access is only approved in unit of consent group, and, therefore data requestors should understand the Data Use Limitations (DULs) of a consent group prior to applying for dbGaP data access.

DULs of a dbGaP study is listed in the "Authorized Access" section on the study page of dbGaP public website. From the study page (such as this one), there is a link named "Authorized Access" right under the

"Variables" tab on top of the page.The link is a shortcut to the section. The DULs of each consent group is listed in the section. A screenshot of the section can be found here. (**07/27/2017**)

Data Use Certification (DUC)

Where can I find the Data Use Certification (DUC) for a particular study?

The terms and conditions of using dbGaP data vary by study. A link to the Data Use Certification (DUC) of a particular study can be found in the "Authorized Access" section on the study page of dbGaP. From the study page (such as this one), there is a link named "Authorized Access" right under the "Variables" tab on top of the page. The link is a shortcut to the section. A link to the DUC is included in the section. A screenshot of the section can be found from here. (**09/24/2017**)

Data Manifest

Can I get a list of data files available for download prior to submitting a Data Access Request (DAR)?

A detailed description of distributed files can be found from the Study Configuration Report and Data Manifest available on the dbGaP public FTP site such as here. The FTP folder can be found through the link named "List of components" in the "Authorized Access" section of the study page. A screenshot of the section can be found here.

Please note that the information of SRA (Short Read Archive) data files is currently not included in the study report and data manifest. The SRA data here is a general term for the Next-Gen sequence raw data, BAM files, or other high throughput data. You may have to directly check the SRA or Biosample websites for the information. (**08/30/2012**))

Data Manifest of Sequencing (BAM, SRA, etc.) Data

I don't see BAM and SRA data in the data manifest available from the dbGaP FTP site. Where can I find the information?

The sequencing data (including Next-Gen sequence raw data, BAM files, or other high throughput data.) distributed through the dbGaP are subject to frequent real time updates, the information of which, therefore, is not included in the study report and manifest available on the dbGaP FTP site.

The information of all data files available for download through the dbGaP Authorized Access System, including that of sequencing data files, however, can be found from a manifest file inside of the Authorized Access System. It can only be accessed through an approved Data Access Request (DAR). The following is how to locate it:

Login to the dbGaP account, go to "My Request" tab, and click on the "Request data" link in the "Actions" column of respective table row. There is a link named "Dataset Manifest" right above the "Create n ew data request for download" section. Clicking on the link allows saving or opening of a spreadsheet file. All the phenotype, genotype, and sequencing files are listed along with their sample ID information in the file.

(**07/14/2017**)