Instructions to request authorized access to the TCGA project dataset

This document contains the instructions for submitting a Data Access Request for The Cancer Genome Atlas (TCGA) dataset. **Note: If you would like to access additional dataset(s) for use in an existing Project, please use the "request new datasets" function under the appropriate Project Title. Do not file a new Project Request. This will minimize paperwork for you.** A complete request includes the following items:

1. Information about your institution, including contact information for you and your Institutional Signing Official. *Please note that the primary requestor MUST be the Principal Investigator.*

2. Brief description of your research objectives, design, and analysis plan through the Research Use Statement (see "Preparing the Research Use Statement" below for specific details).

3. A non-technical summary of your proposed research to communicate to the general public the purpose and objectives of your research.

4. Full legal names and other information about collaborators at your local institution. Only include collaborators at your institution in the "Collaborators" section of the request (Research & Related Senior/Key Person Profile). Collaborators at different institutions must submit separate Data Access Requests for access to individual-level data.

   For Collaborations: Coordinated requests by collaborating organizations should have the same Project Title, and each should reference the collaboration (either the name and institution of the individual PI(s) for smaller collaborations, or the consortium name for larger efforts) in the Research Use Statement.

5. Information about the IT Director or his/her designee at your local institution should be included in the "Collaborators" section of the request. The IT Director or his/her designee is someone with the authority to vouch for the IT capacities at your institution, or higher-level division of an institution (e.g., the School of Medicine).

6. Review of and agreement to the terms and statements outlined in the TCGA Data Use Certification (appended to the SF424 (R&R)).

7. Review and approval of the Data Access Request by your Institutional Signing Official.

These 7 items are included in the SF424 (R&R) generated by the dbGaP request system.

Note: Prior approval for a TCGA dataset does not cover use of this dataset in a new project. If you plan to combine datasets from a previously approved project with a new set of requested datasets, all datasets that will be used **must** be selected as part of the project, even if you were previously approved to use these datasets under a different project.

Preparing the Research Use Statement [SF 424 (R&R), page 3]

- If, in addition to requesting the TCGA dataset, you are requesting other dataset(s) that are not part of TCGA, your Data Access Request will be reviewed by the Data Access Committee(s) for all NIH Institutes and Centers that are responsible for datasets included within your access request. The responsible Institute and Centers may have additional requirements for requesting data or different terms of access. Be sure to read the Instructions and Data Use Certification document for all requested datasets.

- The TCGA Data Access Committee will compare the Research Use Statement to the ***Use Restrictions*** for the dataset to confirm that the proposed research is consistent with the use restrictions provided by the institution that submitted the dataset.

- Both the Research Use Statement and the non-technical summary, which are the basis by which researchers have been allowed to access genetic and phenotypic data, will be included on the dbGaP website to describe your research project to the public.

Tip: Prepare and save the Research Use Statement and the non-technical summary in a word processing application. NOTE: THERE IS A 2200 CHARACTER LIMIT (INCLUDING SPACES). These statements can be copied and pasted into their respective form boxes during the dbGaP request process. The statements will be inserted into page 3 of the SF424

The Research Use Statement should be 1-2 paragraphs no longer than 2200 characters and include:

1. the research objectives,
2. the study design,
3. the analysis plan (including the phenotypic characteristics that will be tested for association with genetic variants),
4. an explanation of how the proposed research is consistent with *Use Restrictions* for the requested dataset(s)
5. if relevant, a brief description of any planned inter-institutional collaborative research including the name of the collaboration and/or the individual investigators and their institutions.

For Collaborations: Data cannot be shared with these collaborators until they have submitted a similar request from their own institution and have been approved for access to the datasets. All data security practices outlined in the Data Use Certification and the dbGaP Security Best Practices are expected to be followed for the collaborative research project.

If the phenotypic characteristics you plan to evaluate were not the primary focus of the original study (e.g., you propose to create a new outcome variable from existing variables), describe the new outcomes and explain how they are consistent with the Use Restrictions for the dataset being requested.

If you intend to use a dataset or multiple datasets to develop or evaluate performance of new analytical methods, include a sentence indicating how this work is relevant to the specific disease datasets being studied, such as: This methodological work will advance the understanding of the genetic basis of each of the primary conditions included in the requested datasets.

Special care should be taken when requesting datasets from more than one study.

1. Make it clear whether you plan to combine data from the studies for your analyses and focus on outcomes or hypotheses that weren't the primary focus of the original study, or analyze the datasets separately with a focus on the primary study outcomes.
2. If you are focusing on outcomes or hypotheses that weren't the primary focus of the original studies, describe the outcomes you propose to use and state how your proposed use is consistent with any Use Restrictions for each of the requested datasets.
3. If you are combining datasets from more than one study for your analyses, discuss whether this creates any additional risks to participants.

Please note: Requesters may apply for access to additional datasets for use in an existing Project at any time. *Do not submit a new Project Request.* Simply log on to dbGaP and use the "request new datasets" function under the appropriate study on the "My Research Projects" page.

Research & Related Senior/Key Person Profile [SF 424 (R&R), page 4+]

- List all the collaborating investigators and the IT Director or his/her designee at your organization. By submitting an individual's name on the form, requestors and Institutional Signing Officials affirm that the collaborators have read and agreed to the terms and statements within the Data Use Certification.
- At a minimum, you are required to complete the "name", "position/title", and "organization name" information fields for each collaborator listed.