CDS New Request Questionnaire

OMB	**No.**:	0925-7775
Expiration	Date:	06/30/2025

Collection of this information is authorized by The Public Health Service Act, Section 411 (42 USC 285a). Rights of participants are protected by The Privacy Act of 1974. Participation is voluntary, and there are no penalties for not participating or withdrawing at any time. Refusal to participate will not affect your benefits in any way. The information collected will be kept private to the extent provided by law. Names and other identifiers will not appear in any report. Information provided will be combined for all participants and reported as summaries. You are being contacted by email to complete this form so that NCI can consider your study for submission into the Cancer Data Service.

Public reporting burden for this collection of information is estimated to average 60 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. 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 (0925-7775). Do not return the completed form to this address.

The following sets of high-level questions are intended to provide an insight to CDS, into the data storage, access and secondary sharing needs and requirements of data submitters. It is requested that the submitters answer as many questions as they can. It is not required to answer all questions.

Data Characteristics

1. Is your study funded by NCI?
2. What is the name of the study the data being submitted to CDS belongs to?
3. What are the principal types of data the program will be submitting (e.g., genomic, clinical, imaging, clinical trial)?

   Note: Imaging data being submitted to CDS
   - needs to be de-identified and does not contain PII and PHI information

- should not include Head, Neck, and Brain images
- should be data which does not fit the criteria to get into IDC

4. Are there additional data types associated with this principal data type (from questions one)? For ex: Proteomics, Imaging, along with Genomics.
5. Do you anticipate other additional data types to be submitted to CDS in future? For example, data type that does not fit the submission criteria of other CRDC nodes.
6. Is the data from Humans or human tumor models, if so is it adult or pediatric?

   NOTE: CDS accepts only Human, human model data at his point.
7. What additional associated data would you be providing? For ex: Clinical/Phenomics data from study participants and/or any other study associated metadata/searchable variables. Describe the format for each.

   NOTE: CDS at this point will accept all metadata submitted.
8. What is the total number of samples and participants per study, being submitted?

Data Storage and Management

1. Who is the PI on the study?
2. Which NCI program/initiative/Grant is the data generated under and who is the Program Officer?
3. How much data are you planning to submit to CDS?
   a. By data type (if known)?
4. Will your program fund the storage and download of the data?
5. What is the reason you are looking for storage with CDS? What are your challenges related to the storage of data?
6. Does any of the data come from international institutions?
   a. If yes, please provide the name of the institution and/or program.
   b. Please provide the funding mechanism for the international component for this study.

Data Submission

1. Who will submit the data, the PI (or the PI's team) or a collaborator?
2. Is there a program timeline associated with the data Submission?

   NOTE: The CDS metadata template needs to be completed and submitted before the data submission can start.

3. When do you plan to start submitting data to CDS?
4. Who is the primary point of contact for data submission?
5. Will there be multiple uploaders (ex: by data type or working groups)?
6. Do you plan one or multiple submissions to CDS? For example, multiple studies or newer versions of the data for the same study, new data types.
    a. If yes, do you have a timeline for the successive submissions?
    b. If this submission has data from a newer version of the study already submitted to CDS, do you want to retain data from the older version/s at CDS?

Data Sharing

1. Is your data being released to broader research community for secondary sharing? If so, when?
2. Is your data sensitive, i.e., require controlled access?
   NOTE: If you have sensitive data, the study needs to be registered with dbGaP before coming to CDS.
3. Has the data already been registered with dbGaP (or any other public data sharing repository)?
    a. If registered with dbGaP, share the dbGaP PHS number and the name of the Genomics Program Administrator (GPA), for any other data sharing repository, please provide information.
    b. Is the study RELEASED by dbGaP?
        a. If in progress what is a plausible timeline?
           NOTE: CDS would like to be notified prior to the dbGaP release so that the release dates are aligned.
4. Has the data already been submitted to any data repository? For Ex: SRA
5. CDS does not allow downloads. Given this, does CDS meet your data sharing needs?
   NOTE: Data can be made downloadable if funded by the program/initiative/study.
6. Are there any data access limitations?
    a. Is the data embargoed?
7. Is any part of this data “open-access” ? for ex: VCFs from Genomics studies.

NOTE: Open access files can be made downloadable if submitted separately from the controlled access data files.

8. How can you assure the data does not contain PII and PHI and/or identifiable data elements?
9. What de-identification methods were used?

Data Users/Access

1. How do the users access the data presently?
    a. How well are these methods working today?
2. Will your data be made accessible through any other repository?

Data Analysis

1. For what purpose(s) do the approved users access the data?
    a. Conduct analyses / computations?
    b. Cite in a publication?

    NOTE: Please provide details about the publication.
2. Do they need to link this data to other data types in other repositories/CRDC nodes for analysis? If yes, please provide a summary of the data associations and any data identifiers.
3. Do you have custom tools you would like to import to Velsera CGC for analyzing your data in CDS?
   NOTE: CDS data is accessed through Velsera CGC where you can analyze the data using various tools and workflows. More information can be explored here.

Data Post CDS Destination

1. Is there a plan to move data out of CDS in the future to other CRDC nodes such as GDC, PDC, IDC etc.

Is there any other information you would like to share about your data?