UIC CENTER FOR CLINICAL AND TRANSLATIONAL SCIENCE

FAQs: CRDW Data Extract Requests

What are the data request and data set types? *(See corresponding tables on page 2)*

Data Counts* - Assistance in obtaining general counts of clinical data (i.e. the number of patients seen for a specified condition or set of conditions).

** Investigators can also access, QUICKSet/i2b2, to obtain some counts data on their own. Training is required to access this system - http://training.ccts.uic.edu **

Data Extracts - Assistance in obtaining complex sets of clinical data for a study, such as identifiable patient information or detailed queries for specific data elements.

Data Set Types:

1) **Requests that are preparatory for research/aggregate counts – Cohort Identification:**
Provides the number of patients within stipulated parameters. No PHI is given - only a count of patients or events. This type of data is typically requested in preparation for a research proposal/grant submission to determine study feasibility.

2) **Requests for De-identified Data Sets:**
This data set is a result of a positive aggregate count (mentioned in #1 above). Health data is provided without any of the 18 direct identifiers (as defined by HIPAA) and is non-coded. This type of data is typically requested in preparation for a research proposal/grant when more specific information about a potential cohort is needed but individual patients do not need to be identifiable.

3) **Requests for Coded Data Sets**:
This data set is the same as type 2 (also stripped of any direct identifiers), but is provided to the researcher in a uniquely coded rather than a de-identified manner. This allows CRDW staff to re-identify the data and provide an updated data set on a continuous basis. The researcher cannot request re-identification of the dataset. The key to enable linkage of a coded dataset with PHI will remain only with the CRDW staff. This type of data can be either a) retrospective that includes all existing patient data at the time data is requested or b) prospective that includes all new incoming patient data. This can include updates on patient's data that has already been collected.

4) **Requests for a Limited Data Set**:
Similar to dataset type 3 above, this data set contains health data that does not include direct identifiers (such as name and street address). However, limited data sets MAY contain the following indirect identifiers:
    a. Town or City, state, zip code;
    b. Ages in years up to 90 years (must aggregate all ages 90 or older)
    c. Dates directly related to an individual – such as birth date, date of death, admission date, discharge date, visit date, diagnosis date, etc. (Month/year is preferred).

Like data set type 3, the data request could be a snapshot of what is available at the time of the request (retrospective) or updated data could be provided on an ongoing basis (prospective). The researcher may not request additional subject identifiers.

5) **Requests for Information for Recruitment Purposes/Identified Data Sets –**
This type of data set includes contact information of patients within a limited data set. The goal is to promote best practices for recruiting and consenting potential research subjects. This dataset is provided such that potential subjects can be approached to determine their interest in participating in research projects. This data set is typically provided for the purpose of pre-trial screening or possible recruitment for a clinical trial, interventional, or observational study. Data provided may include patient names, medical record number, telephone number, address, clinic schedule (time and location), primary care physician name, and provider.

6) **Requests for Defined Identified Cohort Data Sets –**
This type of dataset contains identified patient data that includes health data for a research study. The researcher defines a group of human subjects for whom a specific set of data is requested. *The identified cohort could be individuals that have **provided consent** in the context of an approved IRB protocol, individuals identified through one of the above processes, or individuals identified through the researcher's clinical activities*. The data request could be a snapshot of what is available at the time of request or updated data that could be provided on an ongoing basis.

Example uses of requested dataset	Dataset Request Type					
	1	2	3	4	5	6
	Aggregate Counts	**De-identified Data Set**	**Coded Data Set**	**Limited Data Set**	**Information for Recruitment Purposes**	**Identified Data Set**
	Study feasibility; Hypothesis testing	Chart review, retrospective only (no updates or supplements)	Chart review, retrospective and prospective	Chart reviews that require certain identifier such as zip code or dates of service	Recruitment for clinical studies, clinical registries	Studies on an identified cohort, clinical trials

Summary of Required Documentation for Each Dataset Type:				
	IDUA/DRA	**Non-Human Subjects Determination**	**IRB Approval (Protocol & Approval Letter)**	**Patient Consent**
1. Aggregate Counts				
1a. Self Service	No	No	No	No
1b. Consult Service	No	No	No	No
2. De-identified Data	Yes	Yes	No	No
3. Coded Data				
3a. Retrospective	Yes	Yes	No	No
3b. Prospective	Yes	N/A	Yes	As determined by OPRS
4. Limited Data Set				
4a. Retrospective	Yes	Yes	No	No
4b. Prospective	Yes	N/A	Yes	As determined by OPRS
5. Identified data for Recruitment Purposes	Yes	N/A	Yes	No*
6. Patient Identified Data	Yes	N/A	Yes	Yes/As determined by OPRS
Data set provided only for recruiting and obtaining consent from potential subjects.				

How do investigators retrieve their data extracts?

In most cases, completed datasets will be loaded into a secure REDCap database. Access can only be granted to the REDCap projects after all proper approvals and steps are completed.

What are the IRB Requirements?

(*See the table on page 2 to determine if your project requires IRB approval)*
Note: After your completed DRA form is reviewed, you will receive a PDF copy with a personalized summary of your requirements and next steps.

For projects *without* approved protocol:

1. The PDF copy of your Data Request Authorization (DRA) form should be submitted to IRB with your protocol.
    1. The CCTS Biomedical Informatics Core Manager will email you a PDF copy of the DRA form after it is reviewed.
    2. Submit the DRA form in OPRS Live under other documents
    3. Update protocol to indicate the source of data requested (i.e. *CCTS - CRDW – Clinical Research Data Warehouse*)
    4. Ensure your protocol document lists all data elements requested – this should match your DRA form
2. Once the IRB application is approved, **submit the approved DRA form, updated protocol document, and a copy of IRB approval letter** to the CCTS Biomedical Informatics Core Manager

For projects *with* prior approved protocol:

1. Complete the Data Request Authorization (DRA) form and submit IRB protocol documents to CCTS Biomedical Informatics Core Manager.
        1. If not already stated, update the protocol to indicate the source of data requested (i.e. *CCTS - CRDW*) and ensure your protocol document lists all data elements requested (this should match your DRA form)
        2. Submit IRB amendment with completed DRA form – you will retrieve a PDF copy of the DRA form that can be uploaded to OPRS Live.
        3. Once the IRB application is approved, submit **the approved DRA form, updated protocol document, and a copy of IRB approval letter** to the CCTS Biomedical Informatics Core Manager

OR

1. If the protocol document already lists all requested data elements and source of data, the CCTS Biomedical Informatics Core Manager will review these items to determine if further IRB approval is required.
    1. If all requirements are met, no additional IRB approval will be required.
    2. If it is determined that additional IRB approval is required:
        1. Submit IRB amendment with the completed DRA form – you will retrieve a PDF copy that can be uploaded to OPRS Live.
        2. Once the IRB application is approved, submit **the approved DRA form, updated protocol document, and a copy of IRB approval letter** to the CCTS Biomedical Informatics Core Manager.

Time estimates: How long does it take for an investigator to receive data?

This depends on the complexity of the request – some requests require calculations not easily obtained and others may require crossing systems to match patient data for a complete dataset. Typically, a request can be completed in 4-6 weeks. However, some requests can be fulfilled more quickly. Investigators will receive an individualized time estimate after their request is evaluated and approved.