

## 4 DATA NETWORK

### 4.1 INCLUSION OF DATA IN DATA NETWORK

Institutions may choose to submit data to the DCC for inclusion into the data network, or retain data locally and after appropriate data quality analyses are done, make their data available to the distributed Data Network under the conditions established by the PEDSnet governance and regulatory structures.

The core PEDSnet data resource meets the HIPAA definition of a limited data set (includes dates). For the purposes of inclusion in the PEDSnet data network, a patient is defined as an individual with at least one clinician face-to-face visit in any outpatient or inpatient setting and at least one diagnosis code recorded since January 1, 2009.

---

#### 4.1.1 SUBMISSION OF DATA EXTRACTS TO DCC

Institutions that choose to submit data extracts to the DCC have the following responsibilities and expectations:

- Submit data on a quarterly basis;
- Under the guidance of the DCC, extract data from institutional data sources, transform those data according to the PEDSnet Extract-Transform-Load conventions maintained by the DCC, and securely upload the data extract to the DCC; and
- In response to findings based on DCC data quality assessments, remediate the ETL processes to address data quality issues as much as possible, and/or annotate issues that cannot be resolved and resubmit revised data to the DCC.

Once data are received, the DCC will:

- Destroy any data extract that includes protected health information, other than dates or zip codes, as defined by HIPAA;
- Run data characterization analyses to ensure acceptable levels of data quality;
- Communicate to institutions any problems detected in data quality, and institutions will remediate modifiable problems;
- Create any derived variables or data profiles required by data network procedures;
- Load the data extract into the PEDSnet Data Network once the data meet acceptable levels of data quality;
- Load the data extract into the PCORnet Data Network (via PopMedNet as dictated by PCORnet DCC); and,
- Mediate access to the data for usage approved by the PEDSnet Executive Committee and/or as required by PCORnet.

### 4.1.2 RETENTION OF DATA EXTRACTS WITHIN THE INSTITUTION

Institutions that choose to retain data extracts have the following responsibilities and expectations:

- Incorporate data extracts within the Institution on a quarterly basis;
- Under the guidance of the DCC, institutions that retain their data locally will extract data from institutional data sources, transform those data according to the PEDSnet Extract-Transform-Load conventions maintained by the DCC, and load the data extract to a local PEDSnet data mart staging area;
- The institutional informatics team is responsible for creating any derived variables or data profiles required by data network procedures, under the guidance of the DCC;
- Before loading the data extract into the PEDSnet data network, the institutional informatics team will run data characterization analyses, provided by the DCC, to ensure acceptable levels of data quality, and will report results to the DCC;
- Once the data meet acceptable levels of data quality, they will become part of the PEDSnet Data Network; and,
- Respond to data queries from PEDnet investigators and the PCORnet Data Coordinating Center in accordance with PEDSnet and PCORnet data governance and regulatory requirements.

## 4.2 DATA QUALITY

The Informatics Director will submit data quality reports on a quarterly basis to the Executive Committee. Once that Committee reviews and approves it, the report will be sent to the Steering Board. The report will provide a clear assessment of the quality of PEDSnet data for research. The DCC will communicate to institutions any problems detected in data quality, and institutions will remediate modifiable problems.

## 4.3 COMMON DATA MODEL

PEDSnet data will be stored in both the PEDSnet and PCORnet Common Data Model formats, either at the PEDSnet DCC for sites that provide data (Section 4.2.1) or at the local site for sites that elect to manage data access locally (Section 4.2.2).

### 4.3.1 PEDSNET COMMON DATA MODEL

The DCC will maintain a pediatric-specific Common Data Model (PEDSnet CDM) for storage of PEDSnet data. The PEDSnet CDM will be used for studies and data queries that are done in PEDSnet only. The PEDSnet CDM will be expanded on an annual basis, with expansions proposed by the Data Committee and approved by both the Executive Committee and Steering Board.

### 4.3.2 PCORNET COMMON DATA MODEL

The PEDSnet Data Network will also support the PCORnet Common Data Model. The PCORnet CDM will be used for PCORnet studies and data queries. As per PCORnet policy, the PCORnet CDM will be expanded on an annual basis.

## 4.4 PATIENT REIDENTIFICATION

All Institutions contributing data to the PEDSnet Data Network regardless of whether data are submitted to the PEDSnet DCC or kept locally will retain direct patient identifiers locally, within each institution, and will not share this information with the DCC. Patients will be assigned a PEDSnet Patient ID that has no internal meaning. Sites will retain the mapping between the PEDSnet Patient ID and local identifiers (such as a medical record number).

Certain research applications may require patients of interest to be identified in order to contact them for participation in study. All studies requiring reidentification will have IRB oversight. Patient reidentification will be done by providing institutions with the PEDSnet patient identifiers of interest, and the institutions will perform the reidentification.

## 4.5 AUTHORIZED USERS

The DCC will develop and maintain a process and procedures for staff members at member institutions to become authorized users of the PEDSnet data network applications. The Executive Committee will review and approve the process and procedures. The DCC will provide a list to the Executive Committee on a periodic basis of all individuals granted authorized user status during the reporting period.

## 4.6 REQUESTS FOR DATA

Requests for data will be made to the PEDSnet Coordinating Center, which will do an administrative review. Additional data and network governance oversight and review is dictated by the nature of the data request as outlined below.

### 4.6.1 FEASIBILITY AND DESCRIPTIVE ANALYSES

The Data Coordinating Center will provide a service to requesters who would like to obtain counts or cross-tabulations of patients or events in preparation for development of a research project proposal or use in a descriptive study.

These requests will generally not require Institutional Review Board oversight, because they are not considered human subjects research.

All requests will be approved by a Site Principal Investigator or the Executive Committee.

### 4.6.2 PATIENT-LEVEL DATA-SET REQUESTS

Investigators who require either a de-identified or limited (as defined by HIPAA) patient-level data-set will be required to obtain PEDSnet Study Approval (see Section 5.3.2 for definition of “PEDSnet Study”). Once the request is approved, the DCC will provide the minimum data necessary to answer study questions.

## 4.7 REMOVAL OF DATA FROM THE DATA NETWORK

Institutions at any time can request that data stored in either the PEDSnet or PCORnet CDM within the DCC be removed from the data network. These requests must be made in writing and submitted to the PEDSnet Executive Director. Data will be destroyed within 30 days of receipt of this request. Data previously released or archived as part of a research study is not included in the scope of this section.