

[Overview 2](#_Toc8222527)

[Requestor Information 3](#_Toc8222528)

[Principal Investigator/Requestor Information 3](#_Toc8222529)

[Project Overview 3](#_Toc8222530)

[Query Description 4](#_Toc8222531)

[Variable Identification 5](#_Toc8222532)

[Logic Rules for Query 5](#_Toc8222533)

[Output Structure 6](#_Toc8222534)

[Datasets 6](#_Toc8222535)

[Aggregate Results 7](#_Toc8222536)

[Codeset Identification 8](#_Toc8222537)

[INDEX 9](#_Toc8222538)

[Variable Identification 9](#_Toc8222539)

## Overview

This document is designed to facilitate collaboration between the study team and the PEDSnet Data Coordinating Center (DCC) and to specify the data requirements for a research study.   
  
The '***Requestor Information'*** section elicits general information from the study investigator, including contact information, study title, status of project, etc. The '***Query Description'*** section allows the investigator to describe the study and list the main variables of interest. These two sections are designed to be completed by the Investigator.   
  
The ***‘Variable Identification'*** section assists study teams in mapping the study variables to domains within the PEDSnet data model. Most datasets returned to investigators will be formatted similarly to the PEDSnet data model. The 'Cohort Identification' page allows the investigator to become better acquainted with how the data is organized in the database. This page is designed to be completed by the Investigator in collaboration with either their local site Informatics Analyst or one of the DCC Data Scientists.   
  
The '***Codeset Identification'*** section assists study teams in creating and/or translating specific codesets for their study or project. Examples of codesets include diagnostic criteria, procedure codes, medication prescriptions, etc. These codesets are critical in ensuring that the information the study team wishes to capture in the study is adequately reflected in the data that the PEDSnet DCC generates. This page is designed to be a collaborative effort between the PEDSnet DCC and the study team.

## Requestor Information

|  |  |
| --- | --- |
| Request Date |  |
| Principal Investigator/Requestor Information | |
|
| First Name |  |
| Last Name |  |
| Title |  |
| Institution |  |
| Email Address |  |
| [Optional] Administrative Assistant Name |  |
| [Optional] Administrative Assistant Email |  |
| [Optional] Additional Research Contact (Coordinator, Project Manager etc.) |  |
| [Optional] Additional Research Contact Email |  |
|  |  |
| Project Overview | |
| Are you affiliated with a PEDSnet site? |  |
| If Yes, please state your PEDSnet member institution affiliation |  |
| Project Title |  |
| Status of Project |  |
| Please provide a brief overview of your project aims and goals. |  |

## Query Description

|  |  |  |
| --- | --- | --- |
| Characteristics for cohort inclusion | Inclusion Criteria |  |
| Exclusion Criteria |  |
| Notes (please include time constraints here and any other notes you wish to include) |  |
| Other variables | Criteria |  |
| Notes (please include time constraints here and any other notes you wish to include) |  |

## Variable Identification

**Instructions:** The table below maps the specifications in the 'Query Description' to specific tables in the PEDSnet data model.

Please complete the following:

* Identify your study variables and the corresponding domain in the PEDSnet data model that contains your variable. An example of a completed table is provided in the index for you.
* Provide any additional logic to assist with writing the code for the query (e.g., years of inclusion, conditional inclusion, temporal associations, etc)

Please do not specify your codesets here. For example, if you have identified ICD-9 codes that you wish to use, that list will go on the ‘Codeset Identification’ page. The purpose of this page is just to map your study variables to the PEDSnet data model.

Please reference the PEDSnet website for detailed information on our data:

<https://pedsnet.org/data/common-data-model/>

|  |  |  |
| --- | --- | --- |
|  | **Cohort Identification** | **Other Variables** |
| **Demographics** |  |  |
| **Visits** |  |  |
| **Diagnoses** |  |  |
| **Procedures** |  |  |
| **Medications** |  |  |
| **Immunizations** |  |  |
| **Labs/Measurement** |  |  |
| **Death** |  |  |
| **Provider** |  |  |
| **Location** |  |  |
| **Care Site** |  |  |
| **Other** |  |  |

## Logic Rules for Query

**Instructions:** Provide the logic of the query based on the variables identified above. Examples include rules for inclusion, exclusion, confounding, variable definitions, etc. An example of logic rules can be found in the index.

Rules for Cohort Inclusion



Rules for Cohort Exclusion



Rules for Dataset Creation



Other Rules



## Output Structure

### Datasets

When the output of the query is a dataset that will be returned to the investigator, the PEDSnet data team uses the following general table structures:

* **Patient Cohort Table**:
  + All of the patients identified in the cohort
  + Demographic information
  + Derivations (if desired) – e.g., flag for main outcome of study
* **Patient Fact Tables:** 
  + Separate tables for each medical domain – e.g., diagnoses, drugs, labs, etc.
  + Generally follows the structure of the PEDSnet CDM
  + Derived variables here are less common but can be discussed in advance
* **Code Lookup Table:**
  + Contains all codes and their definitions

|  |  |  |
| --- | --- | --- |
|  | **PEDSnet Variables** | **Derived Variables** |
| **Patient Cohort Table** |  |  |
| **Visits** |  |  |
| **Diagnoses** |  |  |
| **Procedures** |  |  |
| **Medications** |  |  |
| **Immunizations** |  |  |
| **Labs/Measurement** |  |  |
| **Death** |  |  |
| **Provider** |  |  |
| **Visit Payer** |  |  |
| **Care Site** |  |  |
| **Code Lookup Table** |  |  |
| **Other** |  |  |

### Aggregate Results

When the output of the query is aggregate level data, the DCC will work with the investigator to create table shells that will be filled with the output of the query. The investigator will include the table shells in this section of this document.

## Codeset Identification

**Instructions:**

The purpose of this table is to develop codesets that map to standard concepts in the PEDSnet vocabulary table. The table consists of the following columns:

* **Domain** provides the PEDSnet domain the codes map to (e.g., diagnoses, procedures, etc)
* **Vocabulary** specifies the vocabulary of the codes (e.g., ICD-9, SNOMED, etc)
* **Concept** defines the concept that the codeset attempts to capture (e.g., asthma)
* **Codeset** lists all the applicable codes
* **Notes** provides space to comment or make note of any information not captured in the other fields

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Domain** | **Vocabulary** | **Concept** | **Codeset** | **Notes** |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

## INDEX

### Variable Identification

The table and logic below serves as an example of how to complete the Variable Identification Logic Rules for Query portions of this specifications form. In this example, the investigator wishes to understand potential factors contributing to obesity in patients with asthma.

|  |  |  |
| --- | --- | --- |
|  | **Cohort Identification** | **Other Variables** |
| **Demographics** | * Children between the ages of 5-10 years old | * Race, ethnicity, and gender |
| **Visits** | * Outpatient, inpatient, and ED. Only face to face visits. | * All visits for the identified cohort |
| **Diagnoses** | * Asthma | * Medically complex diagnoses (for flagging) * All inpatient discharge diagnoses * All respiratory-related outpatient and ED visits |
| **Procedures** | * None | * Chest x-rays |
| **Medications** | * None | * Bronchodilators, steroids * Antipsychotics, mood stabilizers, antidepressants |
| **Immunizations** | * None | * None |
| **Labs/Measurement** | * None | * All heights, weights, BMI, and BMI z-scores available for patient * PFT measurements * Blood tests for viral and bacterial infections |
| **Death** | * None | * None |
| **Provider** | * None | * All interactions with a pulmonary specialist |
| **Visit Payer** | * None | * None |
| **Care Site** | * None | * All interactions with a pulmonary specialist |
| **Other** | * None | * None |

**Logic**

Rules for Cohort Inclusion

* Children should receive 2+ visits of asthma between the ages of 5-10 years old.
* Here, a visit is counted as a face to face visits associated with an asthma diagnosis. Once a patient is included, we pull all their face to face visits outside of the age window for inclusion.
* Pull both office visits and problem list diagnoses, as long as these do not occur on the same day
* The index visit for a patient is the date of their first asthma diagnosis for inclusion.

Rules for Cohort Exclusion

* If patients have multiple visits for asthma but do not meet the requirement of 2+ face to face visits, they should still be excluded.
* A patient is excluded if 2+ asthma diagnoses are outside of the age window of 5-10 years, defined as the day of a patient’s fifth birthday until the first day of the patient’s eleventh birthday.

Rules for Dataset Creation

* Create a complexity flag for each patient for each body system. The flag should indicate if a patient has 2+ visits for that body system in the window of interest. A yearly complexity grid should be created for each year a patient is in the cohort.

Other Rules

* Diagnoses for HIV and pregnancy-related diagnoses will need to be excluded from the cohort.