



pcornet<sup>SM</sup>

The National Patient-Centered  
Clinical Research Network

## PCORnet Common Data Model Overview

### **Introduction**

This document provides a summary of the PCORnet Common Data Model (CDM).

For more information, you can find the full CDM specifications online at <http://pcornet.org/pcornet-common-data-model>

You can also learn more about PCORnet, the National Patient-Centered Clinical Research Network, by visiting [www.pcornet.org](http://www.pcornet.org)

### **What is the Common Data Model?**

PCORnet created the Common Data Model for use among its participating networks.

The CDM makes it easier for communities (in this case the PCORnet networks) to share information with each other by setting common definitions and organizing data so that:

- 1) PCORnet can analyze data more quickly.
- 2) Different platforms can be used by networks to organize their data.
- 3) Networks can compare their data more easily and efficiently because everyone organizes and defines their data the same way.

PCORnet developed the CDM for wide distribution. The CDM is licensed under [Creative Commons](#) and therefore easy for anyone to access, use, and share. It is based upon an understanding of the data commonly generated by organizations that deliver healthcare, such as hospitals, doctors' offices, and insurance plans.

The CDM is designed to promote multi-site, patient-centered research and is based on work done by many other groups. These include the [FDA Sentinel Initiative](#), the [Health Care Systems Research Network](#), the [Vaccine Safety Datalink](#), various AHRQ Distributed Research Network projects, and the [ONC Standards & Interoperability Framework and Query Health Initiative](#).

The model's standards for how data is organized and defined will make it easier to do collaborative studies. This includes observational research (studies where researchers observe—rather than test—a drug or procedure) and clinical trials (studies that answer specific questions about a drug, procedure, device, etc.).

The CDM will be reviewed and changed to meet stakeholder needs over time.

### **How does the PCORnet Common Data Model help my network?**

The CDM allows PCORnet networks to compare their data with each other, which is powerful and essential when doing research. Not all networks will be able to collect the

same data elements. Using the CDM, organizations can contribute to overall knowledge with their available data sources, even when they are different than other networks.

The CDM provides guidance on how to build and manage your PCORnet analysis database. Your organization's database programmers can access these guides, including standardized keys (codes). The CDM provides a way to map between the format in which you entered data and the CDM representation. This preserves the values and variables within your data sets while allowing for data comparison *across* organizations and networks.

Data handling can be costly and often takes a lot of time and effort. The CDM is based upon standardized codes that are routinely used in healthcare settings. This means that data can be compared more reliably.

For example, date of birth is a common piece of patient information that is collected. Most studies require it, but organizations use different codes for it. This can quickly become confusing and hard to manage, unless there are standards like those in the CDM. A clinic might code birthdate as "Date\_of\_Birth," a health system might call it "Birth\_DT," and a health registry might use the code "DOB." In this example, the CDM offers a universal variable: "BIRTH\_DATE." Each organization's system can map its own birth date, no matter how it is labeled, to this variable. By using the CDM, your organization keeps control over your members' data, but you can also collaborate with other groups or researchers to run more effective observational research and clinical trials.

### **What is the history of the Common Data Model?**

The CDM has already released several versions and continues to evolve.

- Version 1.0 (released in May 2014) focused on a foundation of data elements that tend to be readily available.
- Version 2.0 (released February 2015) added data elements in new domains of data.
- Version 3.0 (released in June 2015) continues to expand domains, including key information to support pragmatic clinical trial execution.
- Version 3.1 (released in November 2016) makes additional improvements to Version 3.0.

Implementation of each new version incorporates lessons learned during the phases before it. Each organization that uses the CDM must document which version it uses and how it is used. This helps PCORnet track progress and continue to improve the CDM.

### **What should I know about version 3.0?**

The CDM version 3.0 includes several upgrades that make it easier to organize data and manage the database:

1. Additional data elements. Ten new fields in existing tables and five new tables have been added to expand the amount and types of information being represented.

2. Modifications to relational integrity specifications and to date formatting practices. This means that data tables have more rules around null values (when no information is entered) and foreign keys (relationships between tables).
3. New specifications to allow the CDM to run with SAS, a computer program used for statistical analyses.
4. A new method for representing dates consistently across organizations.
5. Additional PCORnet trial tables to connect and filter CDM data inside of a clinical trial's protocol.
6. Written guidance with additional instructions and descriptions.

### **What should I know about version 3.1?**

CDM v3.1 is a minor release with small, yet important, improvements made over the 3.0 version. The v3.1 minor release includes:

1. Four new fields (sexual orientation, gender identity, diagnosis origin, and medication prescribing unit)
2. An expanded "encounter types" value set to include observation stays and institutional professional consults
3. A collapsed value set of procedure terminologies so that current procedural terminology (CPT) and healthcare common procedure coding system (HCPCS) medical codes are grouped into a single category
4. Clarification around the expected number of digits for relational database management system (RDBMS) number formatting
5. Removal of the date of death requirement field for death table
6. An enrollment table that now includes drug coverage

### **What influenced the development of the CDM?**

The PCORnet CDM is based on the [Sentinel Common Data Model](#), an FDA surveillance platform that incorporates EHR and health plan data. The Sentinel Initiative was built to support rapid data analysis.

### **What are the technical specifications of the CDM?**

The PCORnet Common Data Model:

- Protects patient confidentiality by using the "minimum necessary" of patient data. The CDM also specifies that each network does not use medical record numbers (MRNs) to identify patients, but instead creates an arbitrary identifier number.
- References standard terminologies such as CPT, HCPCS, the International Classification of Diseases (ICD), Systematized Nomenclature of Medicine (SNOMED), and Logical Observation Identifiers Names and Codes (LOINC®).
- Is covered under a [Creative Commons Attribution 4.0 International License](#), so anyone interested can freely obtain and use it.

## Overview Diagram

The below overview diagram shows high-level domains included in CDM version 3.0 and version 3.1.

## PCORnet Common Data Model Domains, v3.0 and v3.1

### DEMOGRAPHIC v1.0

Demographics record the direct attributes of individual patients.

### ENROLLMENT v1.0

Enrollment is a concept that defines a period of time during which a person is expected to have complete data capture. This concept is often insurance-based, but other methods of defining enrollment are possible.

### ENCOUNTER v1.0

Encounters are interactions between patients and providers within the context of healthcare delivery.

### DIAGNOSIS v1.0

Diagnosis codes indicate the results of diagnostic processes and medical coding within healthcare delivery. Data in this table are expected to be from healthcare-mediated processes and reimbursement drivers.

### PROCEDURES v1.0

Procedure codes indicate the discreet medical interventions and diagnostic testing, such as surgical procedures and lab orders, delivered within a healthcare context.

### VITAL v1.0

Vital signs (such as height, weight, and blood pressure) directly measure an individual's current state of attributes.

### LAB\_RESULT\_CM v2.0

Laboratory result Common Measures (CM) use specific types of quantitative and qualitative measurements from blood and other body specimens. The common measures are defined in the same way across all PCORnet networks, but this table can also include other types of lab results.

### CONDITION v2.0

A condition represents a patient's diagnosed and self-reported health conditions and diseases. The patient's medical history and current state may both be represented.

### PRO\_CM v2.0

Patient-Reported Outcome (PRO) Common Measures (CM) are standardized measures that are defined in the same way across all PCORnet networks. Each measure is recorded at the individual item level: an individual question/statement, paired with its standardized response options.

### DISPENSING v2.0

Outpatient pharmacy dispensing, such as prescriptions filled through a neighborhood pharmacy with a claim paid by an insurer. Outpatient dispensing may not be directly captured within healthcare systems.

### PRESCRIBING v3.0

Provider orders for medication dispensing and/or administration. These orders may take place in any setting, including the inpatient or outpatient basis.

### PCORNET\_TRIAL v3.0

Patients who are enrolled in PCORnet clinical trials.

### DEATH v3.0

Reported mortality information for patients.

### DEATH\_CAUSE v3.0

The individual causes associated with a reported death.

### HARVEST v3.0

Attributes associated with the specific PCORnet datamart implementation, including data refreshes.

## Acknowledgements

The PCORnet Common Data Model is a product of shared expertise, feedback, effort, and collaboration from many people across PCORnet. We gratefully acknowledge all work and individual contributions.