**Conventions for Population of OMOP CDM V5.0**

**to Support PCORnet Requirements**

Revision Date: January 22, 2016

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Author** | **Changes** |
| 1.0 | 3/3/2016 | Don Torok | Initial document for OMOP v5 to PCORnet v3 |

Contents

[Purpose 2](#_Toc425337826)

[General Conventions 2](#_Toc425337827)

[1. FACT\_RELATIONSHIP 3](#_Toc425337828)

[2. PERSON 4](#_Toc425337829)

[3. DEATH 6](#_Toc425337830)

[4. LOCATION 7](#_Toc425337831)

[5. CARE\_SITE 8](#_Toc425337832)

[6. PROVIDER 9](#_Toc425337833)

[7. OBSERVATION PERIOD 11](#_Toc425337834)

[8. VISIT\_OCCURRENCE 12](#_Toc425337835)

[9. CONDITION\_OCCURRENCE 15](#_Toc425337836)

[10. PROCEDURE\_OCCURRENCE 17](#_Toc425337837)

[11. MEASUREMENT 19](#_Toc425337838)

[12. OBSERVATION 24](#_Toc425337839)

[PCORnet Attributes 26](#_Toc425337840)

[i. Biobank Availability 26](#_Toc425337841)

[ii. Chart Availability 27](#_Toc425337842)

[iii. Hospital Admitting source, Discharge disposition, Discharge status and DRG 27](#_Toc425337843)

[i. Diagnosis Source 29](#_Toc425337844)

[13. OUTSTANDING ISSUES 31](#_Toc425337845)

## Purpose

This document defines a common means of storing information within the OMOP CDM, with the intent that information needed to populate the PCORnet CDM can be obtained from the OMOP CDM using a common set of procedures. Populating OMOP CDMv5 is addressed in the OMOP Common Data Model Specification, Version 5. This document addresses areas where the standards spelled out in the OMOP Common Data Model Specification, Version 5 will not support data elements necessary for the PCORnet CDM or where there is ambiguity in how medical data or observations needed for PCORnet might be recorded in the OMOP CDM.

This is an evolving specification, based in structure on the OMOP Common Data Model with focus on PCORnet requirements.

## General Conventions

1. Concept IDs are taken from OMOP vocabularies v5 or later using the complete (“restricted”) version that includes licensed terminologies such as CPT and others.
2. PCORnet CDM V1.0 requires data elements that are not currently part of the OMOP standard vocabulary. To represent PCORnet concepts that are not represented in the standard OMOP vocabulary, we will be using non-standard concepts from vocabulary\_id = ‘PCORnet’ (former vocabulary\_id = 60). While this violates the OMOP conventions to use only concept\_ids from standard vocabularies, this CDRN-specific convention enables a uniform ETL from OMOP CDM to PCORnet CDM.
3. Representation of “Unknown” flavors.

To support PCORnet conventions for representation of “Unknown” flavors, we will follow these conventions:

|  |  |
| --- | --- |
| **Null Name** | **Definition of each field** |
| A data field is not present in the source system | A corresponding field in the OMOP CDM will be populated with concept\_ID=0. A corresponding record in the OBSERVATION table will not be created. |
| A data field is present in the source system, but the source value is null or blank | A corresponding field in the OMOP CDM will be populated with “No Information” (44814650) from vocabulary\_id = ‘PCORNet’ |
| A data field is present in the source system, but the source value explicitly denotes an unknown value | A corresponding field in the OMOP CDM will be populated with “Unknown”( 44814653) from vocabulary\_id = ‘PCORNet’ |
| A data field is present in the source system, but the source value cannot be mapped to the CDM | A corresponding field in the OMOP CDM will be populated with “Other” (44814649) from vocabulary\_id = ‘PCORNet’ |

## FACT\_RELATIONSHIP

The FACT\_RELATIONSHOP table contains records to detail the relationships between facts within one domain or across two domains, and the nature of the relationship. This table will be used to link Condition\_Occurence and Observation domains and records in Measurement domain.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Field | Type | Required | Description | PCORnet Conventions |
| domain\_concept\_id\_1 | integer | Yes | The concept representing the domain of fact one, from which the corresponding table can be inferred. |  |
| fact\_id\_1 | integer | Yes | The unique identifier in the table corresponding to the domain of fact one. |  |
| domain\_concept\_id\_2 | integer | Yes | The concept representing the domain of fact two, from which the corresponding table can be inferred. |  |
| fact\_id\_2 | integer | Yes | The unique identifier in the table corresponding to the domain of fact two. |  |
| relationship\_concept\_id | integer | Yes | A foreign key identifier to a standard identifier of relationship in the Standardized Vocabularies. |  |

## PERSON

The PERSON table contains records that uniquely identify each patient in the source data who has time at-risk to have clinical events recorded within the source systems.

| **Field** | **Type** | **Required** | **Description** | **PCORnet Conventions** |
| --- | --- | --- | --- | --- |
| person\_id | integer | Yes | A unique system-generated identifier for each person |  |
| gender\_concept\_id | integer | Yes | A foreign key that refers to a standard concept identifier in the Vocabulary for the gender of the person. | Valid OMOP concept\_ids are:   * Female: 8532 * Male: 8507   Allowable concepts have been extended to include the following concepts from vocabulary\_id = ‘PCORNet’:   * Ambiguous: 44814664 * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 * Data field is not present in the source system: 0 |
| year\_of\_birth | integer | Yes | The year of birth of the person. |  |
| month\_of\_birth | integer | No | The month of birth of the person. |  |
| day\_of\_birth | integer | No | The day of the month of birth of the person. |  |
| time\_of\_birth | time | No | The time of birth at the birth day. The format is text: HH:MI:SS military time. |  |
| race\_concept\_id | integer | Yes | A foreign key that refers to a standard concept identifier in the Vocabulary for the race of the person. | Valid concept\_ids are all standard concepts from vocabulary\_id = ‘Race’ plus the following concepts from vocabulary\_id = ‘PCORnet’:   * Multiple Race: 44814659 * Refuse to answer: 44814660 * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 * Data field is not present in the source system: 0   The following standard OMOP concepts have been replaced by PCORnet concepts for uniformity:   * Other Race, 8522, replaced with Other, 44814649 * Unknown, 8552, replaced with Unknown, 44814653   These concepts should not be used. |
| ethnicity\_concept\_id | integer | Yes | A foreign key that refers to the standard concept identifier in the Vocabulary for the ethnicity of the person. | Valid concept\_ids are all standard concepts from vocabulary\_id = ‘Ethnicity’ plus the following concepts from vocabulary\_id = ‘PCORnet’:   * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 * Data field is not present in the source system: 0 |
| location\_id | integer | No | A foreign key to the place of residency or the person in the location table, where the detailed address information is stored. |  |
| provider\_id | integer | No | Foreign key to the primary care provider – the person is seeing in the provider table. |  |
| care\_site\_id | integer | No | A foreign key to the site of primary care in the care\_site table, where the details of the care site are stored |  |
| person\_source\_value | varchar(50) | No | A key derived from the person identifier in the source data. This is necessary when a use case requires a link back to the person data at the source dataset. |  |
| gender\_source\_value | varchar(50) | No | The source code for the gender of the person as it appears in the source data. The size of the field is at least 50. |  |
| race\_source\_value | varchar(50) | No | The source code for the race of the person as it appears in the source data. The size of the field is at least 50. |  |
| ethnicity\_source\_value | varchar(50) | No | The source code for the ethnicity of the person as it appears in the source data. The size of the field is at least 50. |  |
| gender\_source\_concept\_id | integer | No | A foreign key to the gender concept that refers to the code used in the source. |  |
| race\_source\_concept\_id | integer | No | A foreign key to the race concept that refers to the code used in the source. |  |
| ethnicity\_source\_concept\_id | integer | No | A foreign key to the ethnicity concept that refers to the code used in the source. |  |

## DEATH

The death table contains the clinical event for how and when a person dies. Living patients should not contain any information in the death table.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Field | Type | Required | Description | PCORnet Conventions |
| person\_id | integer | Yes | A foreign key identifier to the deceased person. |  |
| death\_date | date | Yes | The date the person was deceased. |  |
| death\_type\_concept\_id | integer | Yes | A foreign key referring to the predefined concept identifier in the Vocabulary reflecting how the death was represented in the source data. | Valid concept\_ids are from vocabulary\_id = ‘Death type’ (IMEDS Death Type), otherwise concept\_id 0 |
| cause\_concept\_id | integer | No | A foreign referring to a standard concept identifier in the Vocabulary for conditions. |  |
| cause\_source\_value | varchar(50) | No | The source code for the cause of death as it appears in the source. The size of the field is at least 50. |  |
| cause\_source\_concept\_id | integer | No | A foreign key to the concept that refers to the code used in the source. Note, this variable name is abbreviated to ensure it will be allowable across database platforms. |  |

**Conventions**

1. There should be only one death record per person.
2. One of the special cases when a death record is created is when Observation table is populated with a record containing concept\_id 44813951 (“Discharge Details”) and value\_as\_concept\_id 4216643 (“Patient Died”) or concept\_id 4137274 (“Discharged to Establishment”) and value\_as\_concept\_id 4216643 (“Patient Died”). In this case. death\_type\_concept\_id is 44818516 (“EHR discharge status "Expired").

## LOCATION

The Location table represents a generic way to capture physical location or address information. Locations are used to define the addresses for Persons and Care Sites.

| **Field** | **Type** | **Required** | **Description** | **PCORnet Conventions** |
| --- | --- | --- | --- | --- |
| location\_id | integer | Yes | A unique system-generated identifier for each geographic location. |  |
| State | varchar(2) | No | The state field as it appears in the source data. |  |
| Zip | varchar(9) | No | The zip code. For US addresses, valid zip codes can be 3, 5 or 9 digits long, depending on the source data. |  |
| location\_source\_value | varchar(50) | No | The verbatim information that is used to uniquely identify the location as it appears in the source data. The size of the field is at least 50. |  |
| address\_1 | varchar(50) | No | The address field 1, typically used for the street address, as it appears in the source data. |  |
| address\_2 | varchar(50) | No | The address field 2, typically used for additional detail such as buildings, suites, floors, as it appears in the source data. |  |
| City | varchar(50) | No | The city field as it appears in the source data. |  |
| County | varchar(20) | No | The county. The county information is necessary because not all zip codes fall into one and the same county. |  |

## 

## CARE\_SITE

The Care\_Site table contains a list of uniquely identified physical or organizational units where healthcare delivery is practiced (offices, wards, hospitals, clinics, etc.).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Field | Type | Required | Description | PCORnet Conventions |
| care\_site\_id | Integer | Yes | A unique system-generated identifier for each defined location of care within an organization. |  |
| care\_site\_name | varchar(255) | No | The description of the care site |  |
| place\_of\_service\_concept\_id | Integer | No | A foreign key that refers to a place of service concept identifier in the Vocabulary | The allowable concepts are limited to the following standard concepts (vocabulary\_id=’Place of Service’)  plus the following concepts from vocabulary\_id = ‘PCORnet’:   * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 * Information not available at the source - 0 |
| location\_id | Integer | No | A foreign key to the geographic location of the administrative offices of the organization in the location table, where the detailed address information is stored. |  |
| care\_site\_source\_value | varchar(50) | No | The identifier for the organization in the source data, stored here for reference. The size of the field is at least 50. |  |
| place\_of\_service\_source\_value | varchar(50) | No | The source code for the place of service as it appears in the source data, stored here for reference. The size of the field is at least 50. |  |

## PROVIDER

The Provider table contains a list of uniquely identified health care providers. These are typically physicians, nurses, etc.

| **Field** | **Type** | **Required** | **Description** | **PCORnet Conventions** |
| --- | --- | --- | --- | --- |
| provider\_id | integer | Yes | A unique system-generated identifier for each provider. |  |
| provider\_name | varchar(50) | No | A description of the provider |  |
| specialty\_concept\_id | integer | No | A foreign key to a standard provider's specialty concept identifier in the Vocabulary. | The allowable concepts are limited to the following standard concepts (vocabulary\_id= ‘Specialty’)  plus the following concepts from vocabulary\_id = ‘PCORnet’:   * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 * Information not available at the source - 0 |
| care\_site\_id | integer | Yes | A foreign key to the main care site where the provider is practicing. |  |
| year\_of\_birth | integer | No |  |  |
| gender\_concept\_id | integer | No |  |  |
| NPI | varchar(20) | No | Optional - Do not transmit to DCC  The National Provider Identifier (NPI) of the provider. |  |
| DEA | varchar(20) | No | Optional - Do not transmit to DCC  The Drug Enforcement Administration (DEA) number of the provider. |  |
| provider\_source\_value | varchar(50) | No | The identifier used for the provider in the source data, stored here for reference. The size of the field is at least 50. |  |
| specialty\_source\_value | varchar(50) | No | The source code for the provider specialty as it appears in the source data, stored here for reference. The size of the field is at least 50. |  |
| specialty\_source\_concept\_id | integer | No | A foreign key to a concept that refers to the code used in the source. |  |
| gender\_source\_value | varchar(50) | No | The source code for the provider gender as it appears in the source data, stored here for reference. The size of the field is at least 50. |  |
| gender\_source\_concept\_id | integer | No | A foreign key to a concept that refers to the code used in the source. |  |

## OBSERVATION PERIOD

The observation\_period table is designed to capture the time intervals in which data are being recorded for the person.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Field | Type | Required | Description | PCORnet Conventions |
| Observation\_period\_id | integer | Yes | A system-generate unique identifier for each observation period |  |
| person\_id | integer | Yes | A foreign key identifier to the person who is experiencing the condition. The demographic details of that person are stored in the person table. |  |
| observation\_period\_start\_date | date | Yes | The start date of the observation period for which data are available from the data source |  |
| observation\_period\_end\_date | date | Yes | The end date of the observation period for which data are available from the source. |  |
| period\_type\_concept\_id | integer | Yes | A foreign key identifier to the predefined concept in the Standardized Vocabularies reflecting the source of the observation period information | The allowable concepts are limited to the following standard concepts from vocabulary\_id = ‘Obs Period Type’:   * Insurance: 44814722 * Geography: 44814723 * Algorithmic: 44814725 * Encounter-based: 44814724 |

**Conventions**

According to PCORnet requirements, “Enrollment” is an insurance-based concept that defines a period during which all medically-attended events are expected to be observed. For partners that do not have enrollment information for some of their patients, other approaches for identifying periods during which complete medical capture is expected can be used.

The Enrollment data can be loaded from OMOP Payer\_Plan\_period table; which in turn is built based on patients' encounters ('E' – encounter based). For Claims based source data this ENR\_BASIS is 'I' – Insurance based.

In the absence of claims data, Encounter-based (44814724) method will be used: where observation period start and end date correspond to the start date of the earliest and end date of the latest available patient visit occurrence respectively.

## VISIT\_OCCURRENCE

The VISIT\_OCCURRENCE table contains the spans of time a person continuously receives medical services from one or more providers at a facility in a given setting within the health care system.

| **Field** | **Type** | **Required** | **Description** | **PCORnet Conventions** |
| --- | --- | --- | --- | --- |
| visit\_occurrence\_id | integer | Yes | A unique system-generated identifier for each person’s visits or encounter at a healthcare provider. |  |
| person\_id | integer | Yes | A foreign key identifier to the person for whom the visit is recorded. |  |
| visit\_start\_date | Date | Yes | The start date of the visit. |  |
| visit\_start\_time | Time | No | The time the visit started. The format is text: HH:MI:SS military time. |  |
| visit\_end\_date | Date | Yes | The end date of the visit. If this is a one-day visit the end date should match the start date. | According to PCORnet requirements, visit\_end\_date should be populated for all Inpatient Visits and Long Term Care Visits. |
| visit\_end\_time | Time | No | The time the visit ended. The format is text: HH:MI:SS military time. |  |
| provider\_id | integer | No | A foreign key to the provider in the provider table who was associated with the visit. |  |
| care\_site\_id | integer | No | A foreign key to the care site in the care site table that was visited. |  |
| visit\_concept\_id | integer | Yes | A foreign key that refers to a place of service concept identifier in the vocabulary. | The allowable concepts are limited to the following standard concepts (vocabulary\_id=’Visit’):   * Inpatient Visit: 9201 * Outpatient Visit: 9202 * Emergency Room Visit: 9203 * Long Term Care Visit: 42898160   plus the following concepts from vocabulary\_id = ‘PCORnet’:   * Non-Acute Institutional Stay: 44814710 * Other ambulatory visit: 44814711 * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 |
| visit\_type\_concept\_id | integer | Yes | A foreign key to the predefined concept identifier in the Standardized Vocabularies reflecting the type of source data from which the visit record is derived. | The allowable concepts are limited to the following standard concepts (vocabulary\_id=’Visit Type’):  44818519 Clinical Study visit  44818518 Visit derived from EHR record  44818517 Visit derived from encounter on claim |
| visit\_source\_value | varchar(50) | No | The source code for the visit as it appears in the source data. The size of the field is at least 50. | This column holds the source code value that was used to determine the visit type, although visit type if often determine by context rather than an actual POS |
| visit\_source\_concept\_id | integer | No | A foreign key to a concept that refers to the code used in the source. | Not populated |

**Conventions**

1. PCORnet expects all diagnoses and procedures to have an associated encounter.

In case when there is no real encounter (e.g. nocturnal dialysis, medication refill, etc.), a visit occurrence record is not created. An observation (diagnosis, procedure, medication, etc.) is stored in a respective domain table.

In case when there is a foundation to derive encounter information (e.g. claims data), a derived visit occurrence record is created and assigned an appropriate visit type (visit\_concept\_id).

1. PCORnet expects the following classification of encounters:

* Ambulatory Visit: Includes visits at outpatient clinics, physician offices, same day/ambulatory surgery centers, urgent care facilities, and other same-day ambulatory hospital encounters, but excludes emergency department encounters.
* Emergency Department: Includes ED encounters. Those ED encounter that become inpatient stays (in which case inpatient stays would be a separate encounter) should have Discharge to Establishment equal to IP (see OBSERVATION section).

ED excludes urgent care visits that take place at other than ED urgent care facilities.

ED claims should be pulled before hospitalization claims to ensure that ED with subsequent admission won't be rolled up in the hospital event.

* Inpatient Hospital Stay: Includes all inpatient stays, including: same-day hospital discharges, hospital transfers, and acute hospital care where the discharge is after the admission date.
* Non-Acute Institutional Stay: Non-Acute Institutional Stay: Includes hospice, skilled nursing facility (SNF), rehab center, nursing home, residential, overnight non-hospital dialysis and other non-hospital stays.
* Other Ambulatory Visit: Includes other non-overnight AV encounters such as hospice visits, home health visits, skilled nursing facility visits, other non-hospital visits, as well as telemedicine, telephone and email consultations. May also include "lab only" visits (when a lab is ordered outside of a patient visit), "pharmacy only" (e.g., when a patient has a refill ordered without a face-to-face visit), "imaging only", etc.

These types are represented respectively by OMOP concepts stored in visit\_concept\_id in the table above.

1. For transfers, such as Emergency Room Visit to Inpatient Visit, use FACT\_RELATIONSHIP table to record the link between the two visits. There will be two records created in FACT\_RELATIONSHIP using relationship\_concept\_id 35022490 : ‘Patient moved to’ and 35022489 : ‘Occurs after’. Below is an example:

| **Domain\_concept\_id\_1** | **fact\_id\_1** | **Domain\_concept\_id\_2** | **fact\_id\_2** | **relationship\_concept\_id** |
| --- | --- | --- | --- | --- |
| Visit | 46233680 | Visit | 35022490 | Patient moved to |
| Visit | 35022490 | Visit | 46233680 | Patient moved from |

1. Although PCORnet recommends considering multiple visits to the same provider on the same day as one encounter (especially if defined by a reimbursement basis), it is not OMOP representation requirements. It is recommended to preserve source visit granularity unless there is a compelling reason to do otherwise.
2. According to PCORnet requirements, visit\_end\_date should be populated for all Inpatient, Non-Acute Institutional Stay, and Long Term Care Visits.

Since most of Long Term Care Visits will not have end date, this is an open question for PCORnet.

## CONDITION\_OCCURRENCE

The CONDITION\_OCCURRENCE table captures records of a disease or a medical condition based on evaluation by a provider or reported by a patient.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Field | Type | Required | Description | PCORnet Conventions |
| condition\_occurrence\_id | integer | Yes | A unique system-generated identifier for each condition occurrence event. |  |
| person\_id | integer | Yes | A foreign key identifier to the person who is experiencing the condition. |  |
| condition\_concept\_id | integer | Yes | A foreign key that refers to a standard condition concept identifier in the Vocabulary. |  |
| condition\_start\_date | date | Yes | The date when the instance of the condition is recorded. |  |
| condition\_end\_date | date | No | The date when the instance of the condition is considered to have ended |  |
| condition\_type\_concept\_id | integer | Yes | A foreign key to the predefined concept identifier in the Vocabulary reflecting the source data from which the condition was recorded, the level of standardization, and the type of occurrence. For example, conditions may be defined as primary or secondary diagnoses, problem lists and person statuses. | Only the following two types are relevant to PCORnet Principal discharge diagnosis flag:   * Primary Condition: 44786627 * Secondary Condition: 44786629   All other types will translate to the following concepts from vocabulary\_id = ‘PCORnet’:   * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 * Data field is not present in the source system: 0 |
| stop\_reason | varchar(20) | No | The reason, if available, that the condition was no longer recorded, as indicated in the source data. The reason, if available, that the condition was no longer recorded, as indicated in the source data. Valid values include discharged, resolved, etc. Note that a stop\_reason does not necessarily imply that the condition is no longer occurring. |  |
| provider\_id | integer | No | A foreign key to the provider in the provider table who was responsible for determining (diagnosing) the condition. |  |
| visit\_occurrence\_id | integer | No | A foreign key to the visit in the visit\_occurence table during which the condition was determined (diagnosed). | Required for PCORnet |
| condition\_source\_value | varchar(50) | No | The source code for the condition as it appears in the source data. The size of the field is at least 50. |  |
| condition\_source\_concept\_id | integer | No | The source code for the condition as it appears in the source data. | Concept\_id of ICD-9 or other source code |

**Conventions**

1. PCORnet expects all diagnoses and procedures to have an associated encounter.

In case when there is no real encounter (e.g. nocturnal dialysis, medication refill, etc.), a visit occurrence record is not created.

In case when there is a foundation to derive encounter information from the diagnosis record, a derived visit occurrence record is created and assigned an appropriate visit type (visit\_concept\_id).

1. According to PCORnet requirements, ‘Primary Condition’ (44786627) and ‘Secondary Condition’ (44786629) are relevant only to ‘Inpatient Visit’ (9201) and ‘Long Term Care Visit’ (42898160). In OMOP, however, this attribute may accompany any type of visit.
2. There is no attribute for PCORnet required Diagnosis source (Admitting, Discharge, Final, Interim). This attribute will be stored in the OBSERVATION table and linked to CONDITION\_OCCURRENCE via FACT\_RELATIONSHOP table. The details are described in the OBSERVATION section of this document.

## PROCEDURE\_OCCURRENCE

The PROCEDURE\_OCCURRENCE table contains records of activities or processes ordered by and/or carried out by a healthcare provider on the patient to have a diagnostic and/or therapeutic purpose.

| **Field** | **Type** | **Required** | **Description** | **PCORnet Conventions** |
| --- | --- | --- | --- | --- |
| procedure\_occurrence\_id | integer | Yes | A unique system-generated identifier for each procedure occurrence |  |
| person\_id | integer | Yes | A foreign key identifier to the person who is subjected to the procedure. |  |
| procedure\_concept\_id | integer | Yes | A foreign key that refers to a standard procedure concept identifier in the Vocabulary. |  |
| procedure\_date | date | Yes | The date on which the procedure was performed. |  |
| procedure\_type\_concept\_id | integer | Yes | A foreign key to the predefined concept identifier in the Vocabulary reflecting the type of source data from which the procedure record is derived. |  |
| quantity | integer | No | The quantity of procedures ordered or administered. |  |
| modifier\_concept\_id | integer | No | A foreign key to a standard concept identifier for a modifier to the procedure (e.g. bilateral) |  |
| provider\_id | integer | No | A foreign key to the provider in the provider table who was responsible for carrying out the procedure. |  |
| visit\_occurrence\_id | integer | No | A foreign key to the visit in the visit\_occurence table during which the procedure was carried out. |  |
| procedure\_source \_concept\_id | integer | No | A foreign key to a procedure concept that refers to the code used in the source. | Concept\_id of ICD-9-CM or other source code |
| procedure\_source\_value | varchar(50) | No | The source code for the procedure as it appears in the source data. |  |
| modifier\_source\_value | varchar(50) | No | The source code for the modifier as it appears in the source data. |  |

**Conventions**

PCORnet expects all procedures to have an associated encounter.

In case when there is no real encounter (e.g. nocturnal dialysis), a visit occurrence record is not created.

In case when there is a foundation to derive encounter information from the procedure record, a derived visit occurrence record is created and assigned an appropriate visit type (visit\_concept\_id).

## MEASUREMENT

The MEASUREMENT domain captures measurement orders and measurement results. The measurement domain can contain laboratory results, vital signs, quantitative findings from pathology reports, etc.

Vital signs and labs for PCORnet will be stored in this table.

| **Field** | **Required** | **Type** | **Standard** | **Description** | PCORnet Conventions |
| --- | --- | --- | --- | --- | --- |
| measurement\_id | Yes | integer |  | A system-generated unique identifier for each lab result. |  |
| person\_id | Yes | integer |  | A foreign key identifier to the person about whom the lab result was recorded. The demographic details of that person are stored in the person table. |  |
| measurement\_concept\_id | Yes | integer | LOINC | A foreign key to the standard lab result (lab test really) concept identifier in the vocabulary. | Valid Observation Concepts belong to the "Measurement" domain. |
| measurement\_source\_concept\_id | Yes | integer |  | A foreign key to a measurement concept that refers to the code used in the source. |  |
| measurement\_date | Yes | date |  | The date of the Measurement. |  |
| measure\_time | No | time |  | The time of the Measurement. The format is text: HH:MI:SS military time. |  |
| operator\_concept\_id | No | integer | OMOP | A foreign key identifier to the mathematical operator that is applied to the value\_as\_number. Operators are <, ≤, =, ≥, > |  |
| value\_as\_number | No | float |  | The lab result stored as a number. This is applicable to lab results where the result is expressed as a numeric value. |  |
| value\_as\_concept\_id | No | integer |  | A foreign key to an lab result stored as a concept identifier. This is applicable to lab results where the result can be expressed as a standard concept from the vocabulary (e.g., positive/negative, present/absent, low/high, etc.). | Valid concept\_ids provided in the tables below |
| unit\_concept\_id | No | integer | UCUM | A foreign key to a standard concept identifier of measurement units in the vocabulary. |  |
| range\_low | No | float |  | The lower limit of the normal range of the lab result. It is not applicable if the lab result results are non-numeric or categorical, and are in the same units of measure as the lab result value. |  |
| range\_high | No | float |  | The upper limit of the normal range of the lab result. It is not applicable if the lab result results are non-numeric or categorical, and are in the same units of measure as the lab result value. |  |
| measurement\_type\_concept\_id | Yes | integer | OMOP | A foreign key to the predefined concept identifier in the vocabulary reflecting the type of the lab result. | Valid concept\_ids found in CONCEPT table where vocabulary\_id = ‘Meas Type’:  Possible values are:   * Patient reported: 44814721 * Observation Recorded from EHR: 38000276   Or the following concepts from vocabulary\_id = ‘PCORnet’:   * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 * Data field is not present in the source system: 0 |
| provider\_id | No | integer |  | A foreign key to the provider in the provider table who was responsible for making the lab result. |  |
| visit\_occurrence\_id | No | integer |  | A foreign key to the visit in the visit table during which the lab result was recorded. |  |
| measurement\_source\_value | No | varchar(50) |  | The lab test code as it appears in the source data. This code is mapped to a standard concept in the vocabulary and the original code is, stored here for reference. |  |
| unit\_source\_value | No | varchar(50) |  | The source code for the unit as it appears in the source data. This code is mapped to a standard unit concept in the vocabulary and the original code is, stored here for reference. |  |
| value\_source\_value | No | varchar(50) |  | The source value associated with the structured value stored as numeric or concept. This field can be used in instances where the source data are transformed to produce the structured value. |  |

### Populating Vital Signs

Height, Weight, Body mass index (BMI), Systolic & Diastolic blood pressure will be stored in this table.

PCORnet CDM includes the following VITAL table

| **Field Name** | **Data Type** | **Predefined Value Sets and Descriptive Text for Categorical Fields** | **Definition / Comments** |
| --- | --- | --- | --- |
| PATID | TEXT(x) |  | Arbitrary person-level identifier. Used to link across tables. |
| ENCOUNTERID |  |  | Arbitrary encounter-level identifier. This is an optional relationship; the ENCOUNTERID should be present if the vitals were measured as part of healthcare delivery. |
| MEASURE\_DATE | TEXT(10):  Format as YYYYMM-  DD |  | Date of vitals measure. |
| MEASURE\_TIME | TEXT(5): Format as HH:MI using 24-hour clock and zero-padding for hour and minute |  | Time of vitals measure. |
| VITAL\_SOURCE | TEXT(2) | PR = Patient-reported  HC = Healthcare delivery setting  NI = No information  UN = Unknown  OT = Other | The “Patient-reported” category can include reporting by patient’s family or guardian |
| HT | NUMBER(8) |  | Height (in inches) measured by standing. Only populated if measure was taken on this date. If missing, leave blank. Decimal precision is permissible. |
| WT | NUMBER(8) |  | Weight (in pounds). Only populated if measure was taken on this date. If missing, leave blank. Decimal precision is permissible. |
| DIASTOLIC | NUMBER(4) |  | Diastolic blood pressure (in mmHg). Only populated if measure was taken on this date. If missing, leave blank. Only report 1 reading per encounter. |
| SYSTOLIC | NUMBER(4) |  | Systolic blood pressure (in mmHg). Only populated if measure was taken on this date. If missing, leave blank. Only report 1 reading per encounter. |
| ORIGINAL\_BMI | NUMBER(8) |  | BMI if calculated in the source system.  Important: **Do not calculate BMI during CDM implementation**. This field should only reflect originating source system calculations, if height and weight are notstored in the source. |
| BP\_POSITION | TEXT(2) | 01 = Sitting  02 = Standing  03 = Supine  NI = No information  UN = Unknown  OT = Other | Position for orthostatic blood pressure. Leave blank if blood pressure was not measured. |
| RAW\_VITAL\_SOURCE | TEXT(x) |  | Optional field for originating value of field, prior to mapping into the PCORnet CDM value set. |
| RAW\_ DIASTOLIC | TEXT(x) |  | Optional field for originating value of field, prior to mapping into the PCORnet CDM value set. |
| RAW\_ SYSTOLIC | TEXT(x) |  | Optional field for originating value of field, prior to mapping into the PCORnet CDM value set. |
| RAW\_ BP\_POSITION | TEXT(x) |  | Optional field for originating value of field, prior to mapping into the PCORnet CDM value set. |

Each of these attributes will be represented by a single record in the MEASURUMENT table. Each type of measure will be identified by the measurement\_concept\_id. For example, measurements that record the weight will have measurement\_concept\_id *3025315 (Body Weight)*.

The PCORnet *vital source* is determined by the *measurement\_type\_concept\_id* where the possible values are Patient reported (44814721) or Observation Recorded from EHR (38000276).

The PCORnet attribute, *bp\_position*, is derived from the various concept ids for blood pressure readings, Diastolic Blood Pressure – Sitting (3034703) vs Diastolic Blood Pressure – Standing (3019962).

To synchronize Diastolic and Systolic BP in case of multiple measurements, measurement\_date and measurement\_time of the same measurement should be the same. Additionally, records for the same measurement are linked together via FACT\_RELATIONSHIP table. For each pair of BP measurements, there will be two records in the FACT\_RELATIONSHIP table. The first record will contain: domain\_concept\_id\_1 and domain\_concept\_id\_2 equal to 21 (‘Measurement’), Fact\_id\_1 and Fact\_id\_2 equal to the respective measurement\_id of diastolic and systolic BP records in the Measurement table coming from the same measurement, and relationship\_concept\_id equal to 46233682 (‘ Diastolic to systolic blood pressure measurement’). The second record will contain: domain\_concept\_id\_1 and domain\_concept\_id\_2 equal to 21 (‘Measurement’), Fact\_id\_1 and Fact\_id\_2 equal to the respective measurement\_id of systolic and diastolic BP records in the Measurement table coming from the same measurement, and relationship\_concept\_id equal to 46233683 (‘Systolic to diastolic blood pressure measurement’).

The following table lists the concept ids that should be used as observation concept ids within the OMOP CDM to record the vitals.

|  |  |  |
| --- | --- | --- |
| **Measurement** | **Concept Name** | **Concept Id** |
| Height | Body height | 3036277 |
| Weight | Body weight | 3025315 |
| Body Mass Index | Body mass index (BMI) [Ratio] | 3038553 |
| Diastolic Blood Pressure | Diastolic Blood Pressure - Sitting | 3034703 |
|  | Diastolic Blood Pressure - Standing | 3019962 |
|  | Diastolic Blood Pressure - Supine | 3013940 |
|  | Diastolic BP | 3012888 |
| Systolic Blood Pressure | Systolic Blood Pressure - Sitting | 3018586 |
|  | Systolic Blood Pressure - Standing | 3035856 |
|  | Systolic Blood Pressure - Supine | 3009395 |
|  | Systolic BP | 3004249 |

### Populating Labs

PCORnet lab list contains only 11 lab names. However, each lab may be represented by a variety of lab tests that differ in the specimen source, units of measurement, precision, and other attributes. These lab tests are uniquely represented by LOINC codes. The total number of various LOINC codes corresponding to the lab tests in each source system may reach hundreds. To focus and expedite mapping from source representations to LOINC codes, a list of most prevalent lab tests and corresponding LOINC codes has been created (see the table below). This list contains OMOP concept\_ID for LOINC codes and units of measurements. This list can and will be expanded as contributing institutions detect LOINC codes that best represent their local lab tests.

Target LOINC Codes

| **PCORnet Lab Name** | **LOINC Code** | **OMOP Concept ID** | **LOINC long common name** | **LOINC short name** | **Component** | **Property** | **TM** | **Specimen** | **Scale** | **Method** | **Example Units (UCUM)** | **Unit OMOP Concept ID** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Creatinine | 12190-5 | 3016662 | Creatinine [Mass/volume] in Body fluid | Creat Fld-mCnc | Creatinine | MCnc | Pt | Body fld | Qn |  | mg/dL | 8840 |
| Creatinine | 2160-0 | 3016723 | Creatinine [Mass/volume] in Serum or Plasma | Creat SerPl-mCnc | Creatinine | MCnc | Pt | Ser/Plas | Qn |  | mg/dL | 8840 |
| Creatinine | 38483-4 | 3051825 | Creatinine [Mass/volume] in Blood | Creat Bld-mCnc | Creatinine | MCnc | Pt | Bld | Qn |  | mg/dL | 8840 |
| Creatinine kinase MB | 13969-1 | 3005785 | Creatine kinase.MB [Mass/volume] in Serum or Plasma | CK MB SerPl-mCnc | Creatine kinase.MB | MCnc | Pt | Ser/Plas | Qn |  | ng/mL | 8842 |
| Creatinine kinase MB | 32673-6 | 3029790 | Creatine kinase.MB [Enzymatic activity/volume] in Serum or Plasma | CK MB SerPl-cCnc | Creatine kinase.MB | CCnc | Pt | Ser/Plas | Qn |  | U/L | ? |
| Creatinine kinase MB | 5912-1 | 3017761 | Creatine kinase isoenzymes [interpretation] in Serum or Plasma | CK Isos SerPl-Imp | Creatine kinase isoenzymes | Imp | Pt | Ser/Plas | Nom |  |  |  |
| Creatinine kinase MB/creatinine kinase total | 12187-1 | 3007150 | Creatine kinase.MB/Creatine kinase.total in Serum or Plasma by Electrophoresis | CK MB CFr SerPl Elph | Creatine kinase.MB/Creatine kinase.total | CFr | Pt | Ser/Plas | Qn | Electrophoresis | % | 8554 |
| Creatinine kinase MB/creatinine kinase total | 20569-0 | 3016311 | Creatine kinase.MB/Creatine kinase.total in Serum or Plasma | CK MB CFr SerPl | Creatine kinase.MB/Creatine kinase.total | CFr | Pt | Ser/Plas | Qn |  | % | 8554 |
| Creatinine kinase MB/creatinine kinase total | 49136-5 | 3048863 | Creatine kinase.MB/​Creatine kinase.total [Ratio] in Serum or Plasma | CK MB SerPl-Rto | Creatine kinase.MB/Creatine kinase.total | Ratio | Pt | Ser/Plas | Qn |  |  |  |
| Creatinine kinase total | 2157-6 | 3007220 | Creatine kinase [Enzymatic activity/volume] in Serum or Plasma | CK SerPl-cCnc | Creatine kinase | CCnc | Pt | Ser/Plas | Qn |  | U/L | ? |
| Hemoglobin | 718-7 | 3000963 | Hemoglobin [Mass/volume] in Blood | Hgb Bld-mCnc | Hemoglobin | MCnc | Pt | Bld | Qn |  | g/dL | 8713 |
| Hemoglobin A1c | 4548-4 | 3004410 | Hemoglobin A1c/Hemoglobin.total in Blood | Hgb A1c MFr Bld | Hemoglobin A1c/Hemoglobin.total | MFr | Pt | Bld | Qn |  | % | 8554 |
| International normalized ratio | 6301-6 | 3022217 | INR in Platelet poor plasma by Coagulation assay |  | Coagulation tissue factor induced.INR | RelTime | Pt | PPP | Qn | Coag |  |  |
| Low-density lipoprotein | 13457-7 | 3028288 | Cholesterol in LDL [Mass/​volume] in Serum or Plasma by calculation | LDLc SerPl Calc-mCnc | Cholesterol.in LDL | MCnc | Pt | Ser/Plas | Qn | Calculated |  |  |
| Low-density lipoprotein | 18262-6 | 3009966 | Cholesterol in LDL [Mass/volume] in Serum or Plasma by Direct assay | LDLc SerPl Direct Assay-mCnc | Cholesterol.in LDL | MCnc | Pt | Ser/Plas | Qn | Direct Assay | mg/dL | 8840 |
| Low-density lipoprotein | 2089-1 | 3028437 | Cholesterol in LDL [Mass/​volume] in Serum or Plasma | LDLc SerPl-mCnc | Cholesterol.in LDL | MCnc | Pt | Ser/Plas | Qn |  | mg/dL | 8840 |
| Low-density lipoprotein | 22748-8 | 3001308 | Cholesterol in LDL [Moles/​volume] in Serum or Plasma |  | Cholesterol.in LDL | SCnc | Pt | Ser/Plas | Qn |  |  |  |
| Low-density lipoprotein | 43727-7 | 3046549 | Lipoprotein.beta.subparticle.small [Moles/volume] in Serum or Plasma | LDL Small SerPl-sCnc | Lipoprotein.beta.subparticle | SCnc | Pt | Ser/Plas | Qn |  | nmol/L | 8736 |
| Low-density lipoprotein | 47213-4 | 3053190 | Cholesterol in LDL real size pattern [Identifier] in Serum or Plasma | LDLc real size Pat SerPl | Cholesterol.in LDL real size pattern | Prid | Pt | Ser/Plas | Nom |  |  |  |
| Low-density lipoprotein | 54434-6 | 40757565 | Lipoprotein.beta.subparticle [Moles/volume] in Serum or Plasma | LDL SerPl-sCnc | Lipoprotein.beta.subparticle | SCnc | Pt | Ser/Plas | Qn |  | nmol/L | 8736 |
| Low-density lipoprotein | 55440-2 | 40758569 | Cholesterol.in LDL (real) [Mass/volume] in Serum or Plasma by VAP | LDLc-R SerPl VAP-mCnc | Cholesterol.in LDL | MCnc | Pt | Ser/Plas | Qn | VAP | mg/dL | 8840 |
| Troponin I cardiac | 10839-9 | 3021337 | Troponin I.cardiac [Mass/volume] in Serum or Plasma | Troponin I SerPl-mCnc | Troponin I.cardiac | MCnc | Pt | Ser/Plas | Qn |  | ng/mL | 8842 |
| Troponin I cardiac | 42757-5 | 3033745 | Troponin I.cardiac [Mass/volume] in Blood | Troponin I Bld-mCnc | Troponin I.cardiac | MCnc | Pt | Bld | Qn |  | ng/mL | 8842 |
| Troponin T cardiac (qualitative) | 33204-9 | 3042837 | Troponin T.cardiac [Presence] in Serum or Plasma | Troponin T SerPl Ql | Troponin T.cardiac | ACnc | Pt | Ser/Plas | Ord |  |  |  |
| Troponin T cardiac (qualitative) | 48426-1 | 3052931 | Troponin T.cardiac [Presence] in Blood | Troponin T Bld Ql | Troponin T.cardiac | ACnc | Pt | Bld | Ord |  |  |  |
| Troponin T cardiac (quantitative) | 48425-3 | 3048529 | Troponin T.cardiac [Mass/volume] in Blood | Troponin T Bld-mCnc | Troponin T.cardiac | MCnc | Pt | Bld | Qn |  | ug/L | 8748 |
| Troponin T cardiac (quantitative) | 6597-9 | 3019572 | Troponin T.cardiac [Mass/volume] in Venous blood | Troponin T BldV-mCnc | Troponin T.cardiac | MCnc | Pt | BldV | Qn |  | ug/L | 8748 |
| Troponin T cardiac (quantitative) | 6598-7 | 3019800 | Troponin T.cardiac [Mass/volume] in Serum or Plasma | Troponin T SerPl-mCnc | Troponin T.cardiac | MCnc | Pt | Ser/Plas | Qn |  | ug/L | 8748 |

Lab test results will be stored in the Measurement table. Only those tests that produce numeric results are stored presently. At this point, no qualitative results are stored.

| **Field** | **Required** | **Type** | **Standard** | **Description** | PCORnet Conventions |
| --- | --- | --- | --- | --- | --- |
| measurement\_id | Yes | integer |  | A system-generated unique identifier for each lab result. |  |
| person\_id | Yes | integer |  | A foreign key identifier to the person about whom the lab result was recorded. The demographic details of that person are stored in the person table. |  |
| measurement\_concept\_id | Yes | integer | LOINC | A foreign key to the standard lab result (lab test really) concept identifier in the vocabulary. |  |
| measurement\_source\_concept\_id | Yes | integer |  | A foreign key to a measurement concept that refers to the code used in the source. |  |
| measurement\_date | Yes | date |  | The date of the Measurement. | This is specimen collection date corresponding to PCORnet SPECIMEN\_DATE |
| measure\_time | No | time |  | The time of the Measurement. The format is text: HH:MI:SS military time. | This is specimen collection time corresponding to PCORnet SPECIMEN\_TIME |
| operator\_concept\_id | No | integer | SNOMED | A foreign key identifier to the mathematical operator that is applied to the value\_as\_number. Operators are <, ≤, =, ≥, > | Valid concept\_ids found in CONCEPT table are:  4171756 <  4171754 <=  4172703 =  4172704 >  4171755 >= |
| value\_as\_number | No | float |  | The lab result stored as a number. This is applicable to lab results where the result is expressed as a numeric value. |  |
| value\_as\_concept\_id | No | integer |  | A foreign key to an lab result stored as a concept identifier. This is applicable to lab results where the result can be expressed as a standard concept from the vocabulary (e.g., positive/negative, present/absent, low/high, etc.). | Not populated for the selected LOINC codes |
| unit\_concept\_id | No | integer | UCUM | A foreign key to a standard concept identifier of measurement units in the vocabulary. | Unit vocabulary is given in the above LOINC code table. |
| range\_low | No | float |  | The lower limit of the normal range of the lab result. It is not applicable if the lab result results are non-numeric or categorical, and are in the same units of measure as the lab result value. | Not populated presently |
| range\_high | No | float |  | The upper limit of the normal range of the lab result. It is not applicable if the lab result results are non-numeric or categorical, and are in the same units of measure as the lab result value. | Not populated presently |
| measurement\_type\_concept\_id | Yes | integer | OMOP | A foreign key to the predefined concept identifier in the vocabulary reflecting the type of the lab result. | 44818702 Lab result |
| provider\_id | No | integer |  | A foreign key to the provider in the provider table who was responsible for making the lab result. |  |
| visit\_occurrence\_id | No | integer |  | A foreign key to the visit in the visit table during which the lab result was recorded. |  |
| measurement\_source\_value | No | varchar(50) |  | The lab test code as it appears in the source data. This code is mapped to a standard concept in the vocabulary and the original code is, stored here for reference. |  |
| unit\_source\_value | No | varchar(50) |  | The source code for the unit as it appears in the source data. This code is mapped to a standard unit concept in the vocabulary and the original code is, stored here for reference. |  |
| value\_source\_value | No | varchar(50) |  | The source value associated with the structured value stored as numeric or concept. This field can be used in instances where the source data are transformed to produce the structured value. |  |

## OBSERVATION

The OBSERVATION table captures any clinical facts about a patient obtained in the context of examination, questioning or a procedure. The observation domain supports capture of data not represented by other domains, including unstructured measurements, medical history and family history.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Field | Type | Required | Description | PCORnet Conventions |
| observation\_id | integer | Yes | A unique system-generated identifier for each observation. |  |
| person\_id | integer | Yes | A foreign key identifier to the person about whom the observation was recorded. |  |
| observation\_concept\_id | integer | Yes | A foreign key to the standard observation concept identifier in the Vocabulary. | Valid Observation Concepts belong to the "Observation" domain. |
| observation\_date | date | Yes | The date of the observation |  |
| observation\_time | time | No | The time of the observation. The format is text: HH:MI:SS military time. |  |
| observation\_type\_concept\_id | integer | Yes | A foreign key to the predefined concept identifier in the Vocabulary reflecting the type of the observation. | Valid concept\_ids found in CONCEPT table where vocabulary\_id = ‘Observation Type’. |
| value\_as\_number | float | No\*  (see convention) | The observation result stored as a number. This is applicable to observations where the result is expressed as a numeric value. |  |
| value\_as\_string | varchar(60) | No\*  (see convention) | The observation result stored as a string. This is applicable to observations where the result is expressed as verbatim text. |  |
| value\_as\_concept\_id | integer | No\*  (see convention) | A foreign key to an observation result stored as a concept identifier. This is applicable to observations where the result can be expressed as a standard concept from the Vocabulary (e.g., positive/negative, present/absent, low/high, etc.). | Valid concept\_ids provided in the tables below |
| unit\_concept\_id | integer | No | A foreign key to a standard concept identifier of measurement units in the Vocabulary. |  |
| provider\_id | integer | No | A foreign key to the provider in the provider table who was responsible for making the observation. |  |
| visit\_occurrence\_id | integer | No | A foreign key to the visit in the visit table during which the observation was recorded. |  |
| observation\_source\_concept\_id |  | No |  |  |
| qualifier\_source\_value | varchar(50) | No | The qualifier code as it appears in the source data. This code is mapped to a standard concept in the Vocabulary and the original code is, stored here for reference. |  |
| observation\_source\_value | varchar(50) | No | The observation code as it appears in the source data. This code is mapped to a standard concept in the Vocabulary and the original code is, stored here for reference. |  |
| unit\_source\_value | varchar(50) | No | The source code for the unit as it appears in the source data. This code is mapped to a standard unit concept in the Vocabulary and the original code is, stored here for reference. |  |

### PCORnet Attributes

There are a number of attributes that are needed to populate the PCORnet V1 Common Data Model which are not available in respective domains of the OMOP CDM v5. These attributes will be stored in the OMOP CDM as observations and fact relationships. This section describes conventions for storing these items.

Items needed by PCORnet that are not explicitly defined in OMOP CDMv5.

* Biobank availability
* Chart availability
* Encounter Admitting source, Discharge disposition, Discharge status and DRG
* Diagnosis source

### Biobank Availability

The PCORnet Demographic table has the attribute, *biobank\_flag*, with the possible values of ‘Y’ or ‘N’.

There are two places in OMOP CDM where information regarding biobank information is available:

1. Observation record for that person with *observation\_concept\_id* equal to Biobank flag (4001345) and the Observation *value\_as\_concept\_id* set to concept Yes (4188539). Biobank records may come from multiple sources. The convention is to have only one record per source.
2. At least one Specimen record for that person exists.

Either one of those two conditions indicates existence of biobank information (biobank\_flag=Y). The absence of Specimen record for the patient and no record in Observation table with *observation\_concept\_id* equal to Biobank flag (4001345) and the Observation *value\_as\_concept\_id* set to concept Yes (4188539) will set biobank\_flag=N.

### Chart Availability

The PCORnet Enrollment table has the attribute *chart* with the possible values of ‘Y’, ‘N’. Then PCORnet Enrollment table corresponds with the OMOP CDM Observation\_Period table. For each person/enrollment period combination, if you can review or requests charts for this person, the will need to be an observation record created. The observation date should be the same as the enrollment period start date. The *observation\_concept\_id* should be Chart availability (4030450) and the *value\_as\_concept\_id* should be set to either Yes (4188539) or No (4188540). The absence of an Observation record for a person for an Observation Period will be interpreted as No. There should be only one record for chart availability per observation period.

### Hospital Admitting source, Discharge disposition, Discharge status and DRG

The PCORnet Encounter table has Admitting source, Discharge Disposition, Discharge Status, and DRG. The OMOP CDM table corresponding to Encounter is VISIT\_OCCURRENCE, it does not have these specific attributes. The observation records with these values need to be associated with hospital visits recorded in VISIT\_OCCURRENCE. There should be only one record for each attribute per visit occurrence.

According to PCORnet requirements, Admitting source, Discharge Disposition, Discharge Status, and DRG should be populated for Inpatient and Long-term visits, and may be populated for Emergency Room visits.

The Hospital Admitting source will be in an observation record for the person that has the same date as the Visit Occurrence *visit\_occurrence\_start\_date* and the *observation\_concept\_id* equal to Admission from Establishment (4145666). The *value\_as\_concept\_id* should contain the OMOP concept that best represents the source data admission source and the *observation\_source\_value* should hold the code or description used to determine the concept id. The ETL from OMOP CDM to PCORnet will need to map these various values into PCORnet defined values. Following the rule of *Unknown Flavors* given earlier in this document, an admitting source record should be created unless the source data does NOT contain admitting source.

The Discharge Disposition (Discharge Details), Discharge Status (Discharge to Establishment) and Hospital Discharge DRG are similarly recorded, only the observation date will correspond to the *visit\_occurrence\_end\_date*, which is assumed to be the discharge date. The following table defines the concept id’s that should be used to identify these records.

If visit type is “Emergency Room Visit” (9203) and it becomes an inpatient stay, Discharge to Establishment should be “Inpatient Hospital” (8717).

The following table contains allowable concept\_id values for each attribute:

| **Concept Name** | **Concept Id** | **Notes** |
| --- | --- | --- |
| Admission from establishment | 4145666 | Possible standard value\_as\_concept\_id:   * Agencies, Foster Care Agency: 38004205 * Nursing & Custodial Care Facilities, Assisted Living Facility: 38004301 * Ambulatory Health Care Facilities, Clinic/Center: 38004207 * Emergency Room – Hospital: 8870 * Agencies, Home Health: 38004195 * Home: 8536 * Hospice: 8546 * Hospitals, General Acute Care Hospital: 38004279 * Nursing Facility: 8676 * Comprehensive Inpatient Rehabilitation Facility: 8920 * Residential Facility: 44814680 * Skilled Nursing Facility: 8863   plus the following concepts from vocabulary\_id = ‘PCORnet’:   * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 |
| Discharge details | 44813951 | Possible standard *value\_as\_concept\_id*:   * Discharged alive: 4161979 * Patient died: 4216643   plus the following concepts from vocabulary\_id = ‘PCORnet’:   * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 |
| Discharge to establishment | 4137274 | * Agencies, Foster Care Agency: 38004205 * Nursing & Custodial Care Facilities, Assisted Living Facility: 38004301 * Patient self-discharge against medical advice: 4021968 * Absent without leave: 44814693 * Patient died: 4216643 * Agencies, Home Health: 38004195 * Home: 8536 * Hospice: 8546 * Hospitals, General Acute Care Hospital: 38004279 * Nursing Facility: 8676 * Comprehensive Inpatient Rehabilitation Facility: 8920 * Residential Facility: 44814680 * Inpatient Hospital: 8717 * Skilled Nursing Facility: 8863   plus the following concepts from vocabulary\_id = ‘PCORnet’:   * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 |
| Hospital discharge DRG | 3040464 | *Value\_as\_concept\_id* is the result of looking up the DRG code using OMOP Vocabulary ‘DRG’.  In case DRG code is not found in the vocabulary or missing in the source system use the following concepts from vocabulary\_id = ‘PCORnet’:   * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 |

### Diagnosis Source

The PCORnet Diagnosis table has the attribute *DX\_SOURCE* (diagnosis source) with the possible values of ‘Admitting’, ‘Final’,’Interim’, and ‘Discharge’. Corresponding OMOP CDM table Condition\_Occurrence does not have this attribute. To record diagnosis source in the OMOP CDM, two records will be created: one in the OBSERVATION and one in FACT\_RELATIONSHIP table. The OBSERVATION record should have observation\_date equal to the corresponding condition\_start\_date. The *observation\_concept\_id* should contain 4021918 (‘Qualifier for type of diagnosis’) and the *value\_as\_concept\_id* should be set to one of the concepts in the table below. The FACT\_RELATIONSHIP record should have domain\_concept\_id\_1 = 19 (‘Condition’) and domain\_concept\_id\_2 = 27 (‘Observation’). Fact\_id\_1 should be equal to the respective condition\_occurence\_id and fact\_id\_2 to the respective observation\_ id. Concept ID for relationship\_concept\_id has been requested.

The absence of an Observation and FACT\_RELATIONSHIP table for the respective CONDITION\_OCCURENCE record will be interpreted as “field is not available at the source”.

There should be only one record for diagnosis source per condition occurrence.

| **Concept Name** | **Concept Id** | **Notes** |
| --- | --- | --- |
| Qualifier for type of diagnosis | 4021918 | Possible standard value\_as\_concept\_id:   * Admitting diagnosis: 4203942 * Final diagnosis: 4230359 * Preliminary diagnosis: 4033240 * Discharge diagnosis: TBD   plus the following concepts from vocabulary\_id = ‘PCORnet’:   * No Information: 44814650 * Unknown: 44814653 * Other: 44814649 |

### Vital Signs

### Tobacco

There are two fields in PCORnet VITAL table, TOBACCO and TOBACCO\_TYPE. TOBACCO field vocabulary reflects current and former smoking status, frequency, and amount of the tobacco smoking. TOBACCO\_TYPE reflects the type of tobacco.

The source for TOBACCO field is in a pair of *Observation.observation\_concept\_id* and *Observation.value\_as\_concept\_id* fields. The value of *Observation.observation\_concept\_id* is 4275495 (‘Tobacco smoking behavior - finding’). Permissible values of *Observation.value\_as\_concept\_id* represent mutually exclusive tobacco smoking status including: current and former smoking status, frequency, and daily consumption, as described in the table below. There should be only one Observation record for a given time point.

| **Concept Name** | **Concept Id** | **Status** | **Frequency** | **Daily consumption** |
| --- | --- | --- | --- | --- |
| Moderate smoker (20 or less per day) | 4209585 | Current smoker | Daily | Moderate smoker (20 or less per day) |
| Heavy smoker (over 20 per day) | 4209006 | Current smoker | Daily | Heavy smoker (over 20 per day) |
| Smokes tobacco daily | 42709996 | Current smoker | Daily | Unknown |
| Occasional tobacco smoker | TBD | Current smoker | Some day | N/A |
| Smoker | 4298794 | Current smoker | Unknown | Unknown |
| Never smoked tobacco | 4144272 | Never smoker | N/A | N/A |
| Non-smoker | 4222303 | Current non-smoker, former status unknown | N/A | N/A |
| Ex-smoker | 4310250 | Former smoker | Unknown | Unknown |
| Unknown | 44814653 | Unknown | N/A | N/A |
| No Information | 44814650 | Unknown | N/A | N/A |
| Other | 44814649 | Status cannot be mapped to any concepts above | N/A | N/A |

The source for TOBACCO\_TYPE field is in a pair of *Observation.observation\_concept\_id* and *Observation.value\_as\_concept\_id* fields. The value of *Observation.observation\_concept\_id* is 4298794 ‘Smoker’*.* The permissible values of *Observation.value\_as\_concept\_id* concepts are given in the table below. Tobacco type record can appear only if tobacco smoking status is or was positive. There should not be any tobacco type records if tobacco status is negative. Permissible combinations of tobacco status and tobacco type concepts are given in the table below. There may be one or more Observation records describing tobacco type for a given time point. This depends on the meaning of the tobacco type concept as described below.

| **Concept Name** | **Concept Id** | **Permissible tobacco status concepts** | **Permissible tobacco type concepts** |
| --- | --- | --- | --- |
| Cigarette smoker | 4276526 | Moderate smoker (20 or less per day): 4209585  Heavy smoker (over 20 per day): 4209006  Chain smoker: 4044778  Smokes tobacco daily: 42709996  Occasional tobacco smoker: TBD  Smoker: 4298794 | Cigar smoker: 4246415  Pipe smoker: 4218917 |
| Cigar smoker | 4246415 | Moderate smoker (20 or less per day): 4209585  Heavy smoker (over 20 per day): 4209006  Chain smoker: 4044778  Smokes tobacco daily: 42709996  Occasional tobacco smoker: TBD  Smoker: 4298794 | Cigarette smoker: 4276526  Pipe smoker: 4218917 |
| Pipe smoker | 4218917 | Moderate smoker (20 or less per day): 4209585  Heavy smoker (over 20 per day): 4209006  Chain smoker: 4044778  Smokes tobacco daily: 42709996  Occasional tobacco smoker: TBD  Smoker: 4298794 | Cigarette smoker: 4276526  Cigar smoker: 4246415 |
| Ex-cigarette smoker | 4298794 | Ex-smoker: 4310250 | Ex-cigar smoker: 4144272  Ex-pipe smoker : 4222303 |
| Ex-cigar smoker | 4144272 | Ex-smoker: 4310250 | Ex-cigarette smoker : 4298794  Ex-pipe smoker : 4222303 |
| Ex-pipe smoker | 4222303 | Ex-smoker: 4310250 | Ex-cigarette smoker : 4298794  Ex-cigar smoker: 4144272 |
| Unknown | 44814653 | Moderate smoker (20 or less per day): 4209585  Heavy smoker (over 20 per day): 4209006  Chain smoker: 4044778  Smokes tobacco daily: 42709996  Occasional tobacco smoker: TBD  Smoker: 4298794  Ex-smoker: 4310250 | None |
| No Information | 44814650 |
| Other | 44814649 |

The two observation records for tobacco status and tobacco type are linked together via FACT\_RELATIONSHIP table. For each pair of observation records, there will be two records in the FACT\_RELATIONSHIP table. The first record will contain: domain\_concept\_id\_1 and domain\_concept\_id\_2 equal to 27 (‘Observation’), Fact\_id\_1 and Fact\_id\_2 equal to the respective observation\_id of tobacco status and tobacco type records in the Observation table, and relationship\_concept\_id equal to TBD. The second record will contain: domain\_concept\_id\_1 and domain\_concept\_id\_2 equal to 27 (‘Observation’), Fact\_id\_1 and Fact\_id\_2 equal to the respective observation\_id of tobacco type and tobacco status records in the Observation table, and relationship\_concept\_id equal to TBD.

## DRUG\_EXPOSURE

The drug exposure domain captures records about the utilization of a Drug when ingested or otherwise introduced into the body. A Drug is a biochemical substance formulated in such a way that when administered to a Person it will exert a certain physiological effect. Drugs include prescription and over-the-counter medicines, vaccines, and large-molecule biologic therapies. Radiological devices ingested or applied locally do not count as Drugs.

Drug Exposure is inferred from clinical events associated with orders, prescriptions written, pharmacy dispensings, procedural administrations, and other patient-reported information, for example:

* The “Prescription” section of an EHR captures prescriptions written by physicians or from electronic ordering systems
* The “Medication list” section of an EHR for both non-prescription products and medications prescribed by other providers
* Prescriptions filled at dispensing providers such as pharmacies, and then captured in reimbursement claim systems
* Drugs administered as part of a Procedure, such as chemotherapy or vaccines

| **Field** | **Type** | **Required** | **Description** | **PCORnet Conventions** |
| --- | --- | --- | --- | --- |
| drug\_exposure\_id | integer | Yes | A system-generated unique identifier for each Drug utilization event. |  |
| person\_id | integer | Yes | A foreign key identifier to the person who is subjected to the Drug. The demographic details of that person are stored in the person table. | Derive prescribing.patid from this value. |
| drug\_concept\_id | integer | Yes | A foreign key that refers to a Standard Concept identifier in the Standardized Vocabularies for the Drug concept. |  |
| drug\_exposure\_start\_date | date | Yes | The start date for the current instance of Drug utilization. Valid entries include a start date of a prescription, the date a prescription was filled, or the date on which a Drug administration procedure was recorded. | Target for this value is prescribing.rx\_start\_date. |
| drug\_exposure\_end\_date | date | No | The end date for the current instance of Drug utilization. It is not available from all sources. | Target for this value is prescribing.rx\_end\_date. |
| drug\_type\_concept\_id | integer | Yes | A foreign key to the predefined Concept identifier in the Standardized Vocabularies reflecting the type of Drug Exposure recorded. It indicates how the Drug Exposure was represented in the source data. |  |
| stop\_reason | varchar(20) | No | The reason the Drug was stopped. Reasons include regimen completed, changed, removed, etc. |  |
| refills | integer | No | The number of refills after the initial prescription. The initial prescription is not counted, values start with 0. | Target for this value is prescribing.rx\_refills. |
| quantity | float | No | The quantity of drug as recorded in the original prescription or dispensing record. | Target for this value is prescribing.rx\_quantity. |
| days\_supply | integer | No | The number of days of supply of the medication as recorded in the original prescription or dispensing record. | Target for this value is prescribing.rx\_days\_supply. |
| sig | clob | No | The directions (“signetur”) on the Drug prescription as recorded in the original prescription (and printed on the container) or dispensing record. |  |
| route\_concept\_id | integer | No | A foreign key to a predefined concept in the Standardized Vocabularies reflecting the route of administration. |  |
| effective\_drug\_dose | float | No | Numerical value of Drug dose for this Drug Exposure record. |  |
| dose\_unit\_concept\_ id | integer | No | A foreign key to a predefined concept in the Standardized Vocabularies reflecting the unit the effective\_drug\_dose value is expressed. |  |
| lot\_number | varchar(50) | No | An identifier assigned to a particular quantity or lot of Drug product from the manufacturer. |  |
| provider\_id | integer | No | A foreign key to the provider in the provider table who initiated (prescribed or administered) the Drug Exposure. | Target for this value would be prescribing.rx\_providerid.  However, we have not populated the provider\_id field in any OMOP tables so far. |
| visit\_occurrence\_id | integer | No | A foreign key to the visit in the visit table during which the Drug Exposure was initiated. | Derive prescribing.encounterid from this value. |
| drug\_source\_value | varchar(50) | No | The source code for the Drug as it appears in the source data. This code is mapped to a Standard Drug concept in the Standardized Vocabularies and the original code is, stored here for reference. | Target for this value is prescribing.raw\_rx\_med\_name. |
| drug\_source\_concept\_id | integer | No | A foreign key to a Drug Concept that refers to the code used in the source. |  |
| route\_source\_value | varchar(50) | No | The information about the route of administration as detailed in the source. |  |
| dose\_unit\_source\_value | varchar(50) | No | The information about the dose unit as detailed in the source. |  |

**Conventions**

* Valid Concepts for the drug\_concept\_id field belong to the “Drug” domain. Most Concepts in the Drug domain are based on RxNorm, but some may come from other sources. Concepts are members of the Clinical Drug or Pack, Branded Drug or Pack, Drug Component or Ingredient classes.
* Source drug identifiers, including NDC codes, Generic Product Identifiers, etc. are mapped to Standard Drug Concepts in the Standardized Vocabularies (e.g., based on RxNorm). When the Drug Source Value of the code cannot be translated into standard Drug Concept IDs, a Drug exposure entry is stored with only the corresponding source\_concept\_id and drug\_source\_value and a drug\_concept\_id of 0.
* The Drug Concept with the most detailed content of information is preferred during the mapping process. These are indicated in the concept\_class\_id field of the Concept and are recorded in the following order of precedence: “Branded Pack”, “Clinical Pack”, “Branded Drug”, “Clinical Drug”, “Branded Drug Component”, “Clinical Drug Component”, “Branded Drug Form”, “Clinical Drug Form”, and only if no other information is available “Ingredient”. Note: If only the drug class is known, the drug\_concept\_id should contain 0.
* A Drug Type is assigned to each Drug Exposure to track from what source the information was drawn or inferred from. The valid domain\_id for these Concepts is “Drug Type”.
* The content of the refills field determines the current number of refills, not the number of remaining refills. For example, for a drug prescription with 2 refills, the content of this field for the 3 Drug Exposure events are null, 1 and 2.
* The route\_concept\_id refers to a Standard Concepts of the “Route” domain. Note: Route information can also be inferred from the Drug product itself by determining the Drug Form of the Concept, creating some partial overlap of the same type of information. However, the route\_concept\_id could resolve ambiguities of how a certain Drug Form is actually applied. For example, a “Solution” could be used orally or parentherally, and this field will make this determination.
* The Effective Drug Dose and the Dose Unit Concepts are provided in cases when dose information is explicitly provided, as it is typically for pediatric and chemotherapeutic treatments. The domain\_id for the Dose Unit Concept is “Unit”. Note: this information can only be present if the Drug contains a single active ingredient. Combination products which have doses for each ingredient need to be recorded as separate records.
* The lot\_number field contains an identifier assigned from the manufacturer of the Drug product.
* If possible, the visit in which the drug was prescribed or delivered is recorded in the visit\_occurrence\_id field through a reference to the visit table.
* If possible, the prescribing or administering provider (physician or nurse) is recorded in the provider\_id field through a reference to the provider table.

## OUTSTANDING ISSUES

Immediate

1. Add new concept for ‘Occasional tobacco smoker’.
2. Determine relationship\_concept\_id for linking tobacco status and tobacco type.
3. ‘U/L’ unit is missing in UCUM vocabulary?
4. Check with Chris if concept 44814723 has been corrected: ‘Period while enrolled in study’ should be changed to ‘Geography based’.

Parking lot

1. Lab handling
   1. PRIORITY
   2. RESULT\_LOC
   3. LAB\_PX
   4. LAB\_PX\_TYPE
   5. LAB\_ORDER\_DATE– add to OMOP CDM and also add TIME
   6. RESULT\_DATE – add to OMOP CDM
   7. RESULT\_TIME - add to OMOP CDM
   8. NORM\_MODIFIER\_LOW
   9. NORM\_MODIFIER\_HIGH
   10. Abnormal indicator
   11. Lab order – add concept
2. ADD Death Handling to OMOP v5-PCORnet v2.

Will we record death information based on ICD9 or other conditions that indicate death?

Handle source of death data

Will we exclude death records if there are conditions/drugs/procs/observations 60 days after indication of death?

1. When observation period and chart availability determination is clear, address how Chart Availability in the Observation table connects with Observation\_period table.
2. Handling of Providers with multiple NPIs in PCORnet???
3. Discuss with OMOP: make care\_site.place\_of\_service\_concept\_id a required field
4. Discuss with OMOP: make Provider.specialty\_concept\_id a required field
5. Questions for OHDSI: Will there be gender, race, ethnicity source vocabularies?
6. Discuss distributions of records across the CDM tables based on the concept domain and how this affects interoperability with PCORnet.
7. Discuss doubling diagnosis, procedure and other records based on code mappings and how these affect records related to these duplicates such as cost records.