# Overview

This document describes the ETL specifications/conventions that enable data stored in the [OMOP CDM v4.0] (http://omop.org/CDM) model to be mapped to the PCORnet v1 model. In particular, this document describes the business logic corresponding to the ETL code in the OMOP-PCORNet r3.1.sql file. In this process, we assume that an OMOP CDM v4 instance is already available for transformation to the PCORnet model instance.

This document should be used in conjunction with the omop_pcornet_mappings.csv file (referred to as the *vocabulary mapping document* hereafter). The vocabulary mapping document is implemented into the table "cz_omop_pcornet_concept_map," referred to as the *source-to-concept mapping table* hereafter.

Each (destination) field in the target PCORnet table is described based on the following dimensions:

1. **Source Table**: the OMOP table to be referred
2. **Source Field(s)**: the field(s) in the source table that is used for deriving the value of the destination field; the value could also be a concatenation of values from multiple source fields, e.g. the value for the BIRTH_DATE field in the PCORnet table DEMOGRAPHIC.
3. **Transformation Logic**: the following dimensions describe the logic to map the OMOP field to PCORnet field:
   ◦ **General**: Any general comment on transformation, such as formatting or re-computation
   ◦ **Concept Class for mapping to PCORnet vocabulary**: the concept class for the destination field in the vocabulary mapping document that maps the OMOP concept identifiers to PCORnet values; this dimension is only applicable for the destination fields that are required to be coded in the PCORnet vocabulary. If empty, the source value should be copied to the destination value without any transformation.
   ◦ **OMOP Concept ID for Observations**: The OMOP concept identifier for any destination field that represents an observable concept, e.g. height, weight, etc.
   ◦ **Required Joins**: The SQL JOIN operations required to derive the correct value for the destination field, e.g. total five join operations are required to generate the DEMOGRAPHIC table.

# PCORnet Table: DEMOGRAPHIC

Reading from OMOP tables: Person and Observation

| Destination Field (PCORnet) | Source table (OMOP) | Source Field(s)(OMOP) | *General* | *Concept Class for mapping to PCORnet vocabulary* | *OMOP Concept_ID for observations* | *Required Join* |
|---|---|---|---|---|---|---|
| PATID | person | person_id | The Data Coordinating Center (DCC) is responsible to compile and re-assign PATID into common schema. | | | |
| BIRTH_DATE | person | year_of_birth, month_of_birth, day_of_birth | Concatenate the source fields | | | |
| BIRTH_TIME | | | | | | |
| SEX | person | gender_concept_id | | 'Gender' | | Person and source-to-concept mapping table (1st instance) |
| HISPANIC | person | ethnicity_concept_id | | 'Hispanic' | | Person and source-to-concept mapping table (2nd instance) |
| RACE | person | race_concept_id | | 'Race' | | Person and source-to-concept mapping table (3rd instance) |
| BIOBANK_FLAG | observation | value_as_concept_id | | 'Biobank_flag' 4001345 | | Person and Observation, Observation and source-to-concept mapping table (4th instance) |
| RAW_SEX | person | gender_source_value | | | | |
| RAW_HISPANIC | person | ethnicity_source_value | | | | |
| RAW_RACE | person | race_source_value | | | | |

# PCORnet Table: ENCOUNTER

Reading from: OMOP tables, Visit_Occurrence, Person, Location, Observation, Concept

**Assumptions:** The provider_id field is left null as the sites do not include this information in the ETL. In the future, we may select the most frequent provider_id from the condition_occurrence/procedure_occurrence/observation tables.

| Destination Field (PCORnet) | Source table (OMOP) | Source Field(s)(OMOP) | *General* | *Concept Class for mapping to PCORnet vocabulary* | *OMOP Concept_ID for observations* | *Required Join* |
|---|---|---|---|---|---|---|
| PATID | Visit_occurrence | person_id | | | | |
| ENCOUNTERID | Visit_occurrence | visit_occurrence_id | | | | |
| ADMIT_DATE | Visit_occurrence | visit_start_date | | | | |
| ADMIT_TIME | | | | | | |

| Destination Field (PCORnet) | Source table (OMOP) | Source Field(s)(OMOP) | General | Concept Class for mapping to PCORnet vocabulary | OMOP Concept_ID for observations | Required Join |
|---|---|---|---|---|---|---|
| DISCHARGE_DATE DISCHARGE_TIME PROVIDERID | Visit_occurrence | visit_end_date | | | | |
| FACILITY_LOCATION | Location | Zip | | | | Visit_occurrence and Care_Site, Care_Site and Location |
| ENC_TYPE | Visit_occurrence | Place_of_service_concept_id | | 'Encounter_type' | | Visit_occurrence and source-to-concept mapping table (1st instance) |
| FACILITY_ID | Visit_occurrence | care_site_id | | | | |
| DISCHARGE_DISPOSITION | Observation (1st instance) | value_as_concept_id | | 'Discharge Disposition' | 44813951 | Visit_occurrence and Observation (1st instance), Observation (1st instance) and source-to-concept mapping table (2nd instance) |
| DISCHARGE_STATUS | Observation(2nd instance) | value_as_concept_id | | 'Discharge Status' | 4137274 | Visit_occurrence and Observation (2nd instance), Observation (2nd instance) and source-to-concept mapping table (3rd instance) |
| DRG | Concept | Concept_code | | | 3040646 | Visit_occurrence and Observation (3rd instance), Observation (3rd instance) and Concept |
| DRG_TYPE | | | Corresponds to the field DRG; '01' if concept_class is 'DRG', else '02' | | | |
| ADMITTING_SOURCE | Observation (4th instance) | value_as_concept_id | | 'Admitting source' | 4145666 | Visit_occurrence and Observation (4th instance), Observation (4th instance) and source-to-concept mapping table (4th instance) |
| RAW_ENC_TYPE | Visit_occurrence | place_of_service_concept_id | Corresponds to the field ENC_TYPE | | | |
| RAW_DISCHARGE_DISPOSITION | Observation (1st instance) | value_as_concept_id | Corresponds to the field DISCHARGE_DISPOSITION | | | |
| RAW_DISCHARGE_STATUS | Observation (2nd instance) | value_as_concept_id | Corresponds to the field DISCHARGE_STATUS | | | |
| RAW_DRG_TYPE | | | Corresponds to the field DRG_TYPE | | | |
| RAW_ADMITTING_SOURCE | Observation (4th instance) | value_as_concept_id | Corresponds to the field ADMITTING_SOURCE | | | |

# PCORnet Table: ENROLLMENT

Reading from OMOP tables: Observation_period, and Observation

**Assumptions:** *We assume that the enrollment information is encounter-based ('E') and hence we use the observation_period table to determine the enrollment duration of a patient. For sites that have the insurance information must use the payer_plan_period table instead and set the BASIS field to 'I', i.e. insurance-based.*

| Destination Field (PCORnet) | Source table (OMOP) | Source Field(s)(OMOP) | General | Concept Class for mapping to PCORnet vocabulary | OMOP Concept_ID for observations | Required Join |
|---|---|---|---|---|---|---|
| PATID | Observation_period | person_id | | | | |
| ENR_START_DATE | Observation_period | observation_period_start_date | | | | |
| ENR_END_DATE | Observation_period | observation_period_end_date | | | | |
| CHART | Observation | value_as_concept_id | | 'Chart availability' | 4030450 | Observation_period and Observation, Observation and source-to-concept mapping table |
| BASIS | | | Hardcore to 'E' (encounter-based) | | | |

# PCORnet Table: DIAGNOSIS

Reading from OMOP tables: Condition_occurrence, and PCORnet table/view ENCOUNTER

| Destination Field (PCORnet) | Source table (OMOP) | Source Field(s)(OMOP) | General | Concept Class for mapping to PCORnet vocabulary | OMOP Concept_ID for observations | Required Join |
|---|---|---|---|---|---|---|
| PATID | Condition_occurrence | Person_id | | | | |
| ENCOUNTERID | Condition_occurrence | Visit_occurrence_id | | | | |
| ENC_TYPE | ENCOUNTER | ENC_TYPE | | | | Condition_occurrence and ENCOUNTER |
| ADMIT_DATE | ENCOUNTER | ADMIT_DATE | | | | |
| PROVIDERID | ENCOUNTER | PROVIDERID | | | | |
| DX | Condition_occurrence | condition_concept_id | | | | |
| DX_TYPE | | | Hardcore to 'SM' (SNOMED-CT) | | | |
| DX_SOURCE | | | | | | |
| PDX | | | | | | |
| RAW_DX | | | | | | |
| RAW_DX_TYPE | | | | | | |
| RAW_DX_SOURCE | | | | | | |
| RAW_PDX | | | | | | |

# PCORnet Table: PROCEDURE

Reading from OMOP tables: Procedure_occurrence, and PCORnet table/view ENCOUNTER

| Destination Field (PCORnet) | Source table (OMOP) | Source Field(s)(OMOP) | General | Concept Class for mapping to PCORnet vocabulary | OMOP Concept_ID for observations | Required Join |
|---|---|---|---|---|---|---|
| PATID | Procedure_occurrence | Person_id | | | | |
| ENCOUNTERID | Procedure_occurrence | Visit_occurrence_id | | | | |
| ENC_TYPE | ENCOUNTER | ENC_TYPE | | | | |

| Destination Field (PCORnet) | Source table (OMOP) | Source Field(s)(OMOP) | General | Concept Class for mapping to PCORnet vocabulary | OMOP Concept_ID for observations | Required Join |
|---|---|---|---|---|---|---|
| ADMIT_DATE | ENCOUNTER | ADMIT_DATE | | | | |
| PROVIDERID | ENCOUNTER | PROVIDERID | | | | |
| PX | Procedure_Occurrence | procedure_concept_id | | | | |
| PX_TYPE | | | | Hardcore to 'SM' (SNOMED-CT) | | |
| RAW_PX | | | | | | |
| RAW_PX_TYPE | | | | | | |

# PCORnet Table: VITAL

Reading from OMOP table: Observation

Notes: The timezone information associated with observation_time is lost during this transformation. The resultant table cannot distinguish between the systolic/diastolic values for two differernt blood pressure position recorded at the same time.

| Destination Field (PCORnet) | Source table (OMOP) | Source Field(s)(OMOP) | General | Concept Class for mapping to PCORnet vocabulary | OMOP Concept_ID for observations | Required Join |
|---|---|---|---|---|---|---|
| PATID | Observation (1st instance) | Person_id | | | | |
| ENCOUNTERID | Observation (1st instance) | Visit_occurrence_id | | | | |
| MEASURE_DATE | Observation (1st instance) | Observation_date | | | | |
| MEASURE_TIME | Observation (1st instance) | Observation_time | | | | |
| VITAL_SOURCE | | | | | | |
| HT | Observation (2nd instance) | Value_as_number | | | 3036277 | Observation (1st instance) and Observation (2nd instance) |
| WT | Observation (3rd instance) | Value_as_number | | | 3025315 | Observation (1st instance) and Observation (3rd instance) |
| ORIGINAL_BMI | Observation (4th instance) | Value_as_number | | | 3038553 | Observation (1st instance) and Observation (4th instance) |
| DIASTOLIC | Observation (5th instance) | Value_as_number | | | 3012888 | Observation (1st instance) and Observation (5th instance) |
| SYSTOLIC | Observation (6th instance) | Value_as_number | | | 3004249 | Observation (1st instance) and Observation (6th instance) |

| Destination Field (PCORnet) | Source table (OMOP) | Source Field(s)(OMOP) | *General* | *Concept Class for mapping to PCORnet vocabulary* | *OMOP Concept_ID for observations* | *Required Join* |
|---|---|---|---|---|---|---|
| BP_POSITION | Observation (7th instance) | Target_concept | | 'BP Position' | | Observation (1st instance) and Observation (7th instance), Observation (7th instance) and source-to-concept-mapping table |
| RAW_VITAL_SOURCE | | | | | | |
| RAW_DIASTOLIC | | | | | | |
| RAW_SYSTOLIC | | | | | | |
| RAW_BP_POSITION | | | | | | |