Specification for ETL
from OMOP CDM v5.1 to PCORnet CDM v3

Revision Date: December 19, 2016

Versions

4/7/2016	Toan	Added the version table
5/5/2016	Don Torok	Added section for Death and Death Cause
8/25/2016	Toan	Added Dispensing
12/19/2016	Don Torok	Reconciled with Conventions Document
		Response to Comments after 12/19/2016 release

Table of Contents

1.0   Introduction

The purpose of this document is to provide a mechanism for PCORnet data partners to communicate information about how they transformed data stored in OMOP Common Data Model (CDM) Version 5.1 format into the PCORnet Common Data Model (CDM) Version 3.0. To describe how the information will be used to help the PCORnet Coordinating Center better understand the transformation process, appropriate uses of the PCORnet data, and the comparability of data sources. This document details the approach used for the Extract, Transform, and Load (ETL) process to transform OMOP CDMv5 data elements to the data elements in the PCORnet CDM Version 3.0.

The document, 2015-07-29-PCORnet-Common-Data-Model-v3dot0-RELEASE.pdf, should be used in conjunction with this document, as the PCORnet Common Data Model has the data types and descriptions of the PCORnet tables.

This document assumes that the conventions outlined in CDRN Conventions for Populating OMOP CDM were followed in populating the OMOP CDMv5 database. It also requires that OMOP Vocabulary 5 or later be used for the ETL.

2.0   Source Data Mapping Approach

This document describes mapping of the target PCORnet Common Data Model (CDM) tables and columns from source OMOP CDM model v5.

The mapping was designed based on OMOP CDM v5 specification, PCORnet CDM v3 specification and Conventions for Populating OMOP CDM v5 for PCORnet v3. The mapping should provide sufficient information in order to design and develop ETL processes.

3.0   Source Data Mapping

This section describes mapping process and ETL conversions for transforming data from an OMOP CDM (source) to a PCORNet CDM (destination).

3.1   Data Mapping

Data mapping expects source and target data to be stored in any conventional relational database system per OMOP CDM v5 and PCORNet CDM v3 specifications respectively.

3.1.1 Table: Demographic

PCORI DEMOGRAPHIC table contains one record per patient. Load Demographic data from OMOP Person table as described below.

Demographic field mapping:

Destination Field	Source Field	Applied Rule	Comment
PATID	Person.person_id		Convert to text
BIRTH_DATE	Use Person.year_of_birth, month_of_birth and day_of_birth to construct date. Substitute month and day (each) as '01' if not available in the source.		Changed to date data type.
BIRTH_TIME	Person.time_of_birth	Use NULL if not available	Convert to text format 'HH:MI' using 24-hour clock and zero-padding for hour and minute
SEX	Person.gender_concept_id Char 2	*See OMOP to PCORnet table below*	*See OMOP Concepts table below*
HISPANIC	Derive from Person.ethnicity_concept_id	*See OMOP to PCORnet table below*	*See OMOP Concepts table below*

SEX — OMOP to PCORnet:

44814664	A = Ambiguous
8532	F = Female
8507	M = Male
44814650	NI = No information
44814653	UN = Unknown
44814649	OT = Other
0	NULL

SEX — OMOP Concepts:

44814664	Ambiguous
8532	Female
8507	Male
44814650	No Information
44814653	Unknown
44814649	Other
0	Field does not exist in the source

HISPANIC — OMOP to PCORnet:

38003563	Y = Yes
38003564	N = No
44814650	NI = No information
44814653	UN = Unknown
44814649	OT = Other

HISPANIC — OMOP Concepts:

38003563	Hispanic or Latino
38003564	Not Hispanic or Latino
44814650	No Information
44814653	Unknown
44814649	Other

		0 NULL	0 Field does not exist in the source
RACE	Derive from Person.Race_Concept_id		The mapping for Race from OMOP to PCORnet is given in a table below.
BIOBANK_FLAG	Observation.value_as_concept_id	If at least one record in Specimen table for the patient exist or in the Observation table observation_concept_id is 4001345 (Biobank flag) with value_as_concept_id = 4188539 (Yes) then set biobank_flag as 'Y' else 'N'	The allowable values are 'Y' or 'N'. The absence of a record indicates that there are no biobank specimens.
RAW_SEX	Person.gender_source_value		
RAW_ HISPANIC	Person.ethnicity_source_value		
RAW_RACE	Person.race_source_value		

OMOP to PCORnet Race Mapping

OMOP		PCORnet
concept	description	Value
38003600	African	03 = Black or African American
38003599	African American	03 = Black or African American
38003573	Alaska Native	01 = American Indian or Alaska Native
38003572	American Indian	01 = American Indian or Alaska Native
8657	American Indian or Alaska Native	01 = American Indian or Alaska Native
38003616	Arab	05 = White
8515	Asian	02 = Asian
38003574	Asian Indian	02 = Asian
38003601	Bahamian	03 = Black or African American
38003575	Bangladeshi	02 = Asian

38003602	Barbadian	03 = Black or African American
38003576	Bhutanese	02 = Asian
38003598	Black	03 = Black or African American
8516	Black or African American	03 = Black or African American
38003577	Burmese	02 = Asian
38003578	Cambodian	02 = Asian
38003579	Chinese	02 = Asian
38003604	Dominica Islander	03 = Black or African American
38003603	Dominican	03 = Black or African American
38003614	European	05 = White
38003581	Filipino	02 = Asian
38003605	Haitian	03 = Black or African American
38003582	Hmong	02 = Asian
38003583	Indonesian	02 = Asian
38003593	Iwo Jiman	02 = Asian
38003606	Jamaican	03 = Black or African American
38003584	Japanese	02 = Asian
38003585	Korean	02 = Asian
38003586	Laotian	02 = Asian
38003597	Madagascar	02 = Asian
38003587	Malaysian	02 = Asian
38003594	Maldivian	02 = Asian
38003612	Melanesian	04 = Native Hawaiian or OtherPacific Islander
38003611	Micronesian	04 = Native Hawaiian or OtherPacific Islander
38003615	Middle Eastern or North African	05 = White
8557	Native Hawaiian or Other Pacific Islander	04 = Native Hawaiian or OtherPacific Islander

38003595	Nepalese	02 = Asian
38003588	Okinawan	02 = Asian
38003613	Other Pacific Islander	04 = Native Hawaiian or OtherPacific Islander
38003589	Pakistani	02 = Asian
38003610	Polynesian	04 = Native Hawaiian or OtherPacific Islander
38003596	Singaporean	02 = Asian
38003590	Sri Lankan	02 = Asian
38003580	Taiwanese	02 = Asian
38003591	Thai	02 = Asian
38003607	Tobagoan	03 = Black or African American
38003608	Trinidadian	03 = Black or African American
38003592	Vietnamese	02 = Asian
38003609	West Indian	03 = Black or African American
8527	White	05 = White
44814659	Multiple Race	06 = Multiple Race
44814660	Refuse to answer	07 = Refuse to answer
44814650	No Information	NI = No information
44814653	Unknown	UN = Unknown
44814649	Other	OT = Other
0	Field does not exist in the source	NULL

3.1.2 Table: Enrollment

The ENROLLMENT table has a start/stop structure that contains records for continuous enrollment periods.

"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.

This table is designed to identify periods during which a person is expected to have complete data capture. Members with medical coverage, drug coverage, or both should be included.

A record is expected to represent a unique combination of PATID, ENR_START_DATE.

Currently OMOP CDM is using the earliest and latest encounter dates ('E'), which is in violation of the PCORnet requirement. This is to be discussed with PCORnet.

Enrollment field mapping:

Destination Field	Source Field	Applied Rule	Comment
PATID	Observation_Period.person_id		Convert to text
ENR_START_DATE	Observation_Period.observation_period_start_date		
ENR_END_DATE	Observation_Period.observation_period_end_date		
CHART	Observation.value_as_concept_id where observation_type_concept_id = 4030450 (Patient chart)	Join to Observation table on person_id, observation_start_date and observation_type_concept_id = 4030450 (Patient chart). If the value_as_concept_id = 4188539 (Yes) then 'Y' else 'N'	The absence of an Observation record for a person for an Observation Period will be interpreted as No.
ENR_BASIS	Observation_Period.period_type_concept_id		

OMOP to PCORnet

44814722	I = Insurance
44814723	G = Geography
44814725	A = Algorithmic
44814724	E = Encounter Based

OMOP Concepts

44814722	Insurance
44814723	Geography
44814725	Algorithmic
44814724	Encounter Based

3.1.3 Table: Encounter

 The ENCOUNTER Table contains one record per ENCOUNTERID (which reflects a unique combination of PATID, ADMIT_DATE, PROVIDERID and ENC_TYPE).

The encounter table should include information on interactions between patients and providers. Each diagnosis and procedure recorded during the encounter should have a separate record in the Diagnosis or Procedure Tables.

Multiple visits to the same provider on the same day may be considered one encounter (especially if defined by a reimbursement basis); if so, the ENCOUNTER record should be associated with all diagnoses and procedures that were recorded during those visits.

Note: PCORnet requires that all Procedure and Diagnosis records link back to an Encounter record. This is not a requirement for OMOP CDM. It may be necessary for the ETL to 'create' Encounters for some Procedures and Diagnosis records.

Encounter field mapping:

Destination Field	Source Field	Applied Rule	Comment
PATID	Visit_Occurrence.person_id		Convert to text, Required
ENCOUNTERID	Visit_Occurrence.visit_occurrence_id		Convert to text, Required
ADMIT_DATE	Visit_Occurrence.visit_start_date		This also will hold the encounter date for encounters that are not ER or Inpatient. Required
ADMIT_TIME	Visit_Occurrence.visit_start_time	If available format as 'hh:mm:ss' military time otherwise is should be NULL	
DISCHARGE_DATE	Visit_Occurrence.visit_end_date		
DISCHARGE_TIME	Visit_Occurrence.visit_end_time	If available format as 'hh:mm:ss' otherwise is should be NULL	
PROVIDERID	Visit_Occurrence.provider_id		Convert to text
FACILITY_LOCATION	Location.zip (first 3 digits only)	Join Visit_Occurrence to Care_Site on care_site_id, then to Location on location_id. NULL if it cannot be derived.	Only if zipcode is available. Otherwise NULL 3-digit zip for PCORNet
ENC_TYPE	Visit_Occurrence.visit_concept_id	OMOP to PCORnet table (below)	Required

OMOP to PCORnet

9201	IP = Inpatient Hospital Stay
9202	AV = Ambulatory Visit
9203	ED = Emergency Department
TBD	EI = Emergency Department Inpatient Hospital Stay

OMOP Concepts

9201	Inpatient Visit
9202	Outpatient Visit
9203	Emergency Room Visit
42898160	Long Term Care Visit
TBD	Emergency Room - Inpatient Visit

		OMOP to PCORnet		OMOP Concepts	
		42898160	IS = Non-Acute Institutional Stay	44814710	Non-Acute Institutional Stay
		44814710	IS = Non-Acute Institutional Stay	44814711	Other ambulatory visit
		44814711	OA = Other Ambulatory Visit	44814650	No information
		44814650	NI = No information	44814653	Unknown
		44814653	UN = Unknown	44814649	Other
		44814649	OT = Other	0	Field does not exist in the source
		0	NULL		

Until OMOP adds a concept for EI, Emergency Department to Inpatient, set the encounter type to EI when the admitting source concept id is 8870 (Emergency Room - Hospital) and the visit concept id is 9201 (Inpatient Visit).

FACILITYID	Visit_Occurrence.care_site_id		Convert to text

DISCHARGE_DISPOSITION — Visit_occurrence .discharge_to_concept_id

OMOP uses a single value, discharge_to_concept_id for both the discharge disposition and status. If the patient discharge concept id is to any of the location listed in the discharge status section, then the assumption is the person was discharged alive. Set to NULL for ambulatory or OA visits.

OMOP to PCORnet	
4161979	A = Discharged alive
4216643	E = Expired
44814650	NI = No information
44814653	UN = Unknown
44814649	OT = Other
Any other code	A = Discharged alive
0	NULL

OMOP Concepts	
4161979	Discharged alive
4216643	Patient died
44814650	No Information
44814653	Unknown
44814649	Other
0	Field does not exist in the source

DISCHARGE_STATUS	Visit_occurrence .discharge_to_concept_id	Set to NULL for ambulatory or OA visits.	

Set to NULL for ambulatory or OA visits.

OMOP to PCORnet		OMOP Concepts	
38004205	AF = Adult Foster Home	38004205	Agencies, Foster Care Agency
38004301	AL = Assisted Living Facility	38004301	Nursing & Custodial Care Facilities, Assisted Living Facility
4021968	AM = Against Medical Advice	4021968	Patient self-discharge against medical advice
44814693	AW = Absent without leave		
4216643	EX = Expired	44814693	Absent without leave
38004195	HH = Home Health	4216643	Patient died
8536	HO = Home / Self Care	38004195	Agencies, Home Health
8546	HS = Hospice	8536	Home
38004279	IP = Other Acute Inpatient Hospital	8546	Hospice
8676	NH = Nursing Home (Includes ICF)	38004279	Hospitals, General Acute Care Hospital
8920	RH = Rehabilitation Facility	8676	Nursing Facility
44814680	RS = Residential Facility	8920	Comprehensive Inpatient Rehabilitation Facility:
8717	SH = Still In Hospital	44814680	Residential Facility
8863	SN = Skilled Nursing Facility	8717	Inpatient Hospital
44814650	NI = No information	8863	Skilled Nursing Facility
44814653	UN = Unknown	44814650	No information
44814649	OT = Other	44814653	Unknown
0	NULL		

			44814649	Other
			0	No matching concept

DRG	Cost.DRG_source_value	Join Cost table on visit_occurrence_id Set to NULL for ambulatory or OA visits.	
DRG_TYPE	See Applied Rule	OMOP CDMv5 does not have this information. Use the appropriate value from the vocabulary below: 01 = CMS-DRG (old system) 02 = MS-DRG (current system) NI = No information UN = Unknown OT = Other	02- double check
ADMITTING_SOURCE	Visit_occurrence .admitting_source_concept _id	Not applicable to ambulatory or OA visits.	

OMOP to PCORnet

38004205	AF = Adult Foster Home
38004195	HH = Home Health
38004207	AV = Ambulatory Visit
8920	RH = Rehabilitation Facility
8870	ED = Emergency Department
8536	HO = Home / Self Care
8546	HS = Hospice
38004279	IP = Other Acute Inpatient Hospital
38004301	AL = Assisted Living Facility
8676	NH = Nursing Home (Includes ICF)
44814680	RS = Residential Facility

OMOP Concepts

38004205	Agencies, Foster Care Agency
38004195	Agencies, Home Health
38004207	Ambulatory Health Care Facilities, Clinic/Center, Ambulatory Surgical
8920	Comprehensive Inpatient Rehabilitation Facility
8870	Emergency Room - Hospital
8536	Home
8546	Hospice
38004279	Hospitals, General Acute Care Hospital
38004301	Nursing & Custodial Care Facilities, Assisted Living Facility

		SN = Skilled Nursing Facility		8676	Nursing Facility
	8863			44814680	Residential facility
	44814650	NI = No information		8863	Skilled Nursing Facility
	44814653	UN = Unknown		44814650	No Information
	44814649	OT = Other		44814653	Unknown
	0	NULL		44814649	Other
				0	Field does not exist in the source

RAW_ENC_TYPE	Visit_Occurrence.visit_source_value		
RAW_ DISCHARGE_DISPOSITION	Visit_occurrence .discharge_to_source_value	Not applicable to ambulatory or OA visits.	Note: discharge_to_source_value is used to populate both discharge disposition and discharge status
RAW_ DISCHARGE_STATUS	Visit_occurrence .discharge_to_source_value	Not applicable to ambulatory or OA visits.	
RAW_ DRG_TYPE	NULL	Not applicable to ambulatory or OA visits.	
RAW_ ADMITTING_SOURCE	Visit_occurrence .admitting_source_value	Not applicable to ambulatory or OA visits.	

3.1.4 Table: Diagnosis

DIAGNOSIS should capture unique diagnoses made during an encounters, except those generated from problem lists. If a patient has multiple diagnoses associated with one encounter, then there should be one record in this table for each diagnosis. Exclude records from the OMOP CDM where the Condition Type Concept is EHR problem list entry (38000245). Records where the Condition Type Concept is EHR problem list entry will go into the PCORnet Condition table.

The admit date for the diagnosis is copied from the encounter record which is the admission or appointment date, whereas in the OMOP CDM, the condition occurrence date is when the condition was defined. Therefore, it is possible that there will be more than one of the same diagnoses during a visit in the OMOP CDM. Duplicate records are also possible due to the mapping of one source code to multiple standard codes.

Duplicate diagnosis records should be reduced to a single record for PCORnet based on the following Condition Occurrence attributes: *person_id, visit occurrence id* and *condition source value*. Some of the columns in the duplicated records may have different values, for example, the *condition status source value* might be 'Admitting' for one record and 'Final' for another. The desire is to select the most definitive record. So in dedupping, order the records by the *condition status source values*: Final/Discharge, Admitting, Interim, No Information, Unknown, Other, NULL, and then by the *condition type*: Principal, Secondary, Unable to Classify, No Information, Unknown, Other, NULL. Take the values from the first row.

Diagnosis field mapping:

Destination Field	Source Field	Applied Rule	Comment
DIAGNOSISID	Condition_Occurrence .condition_occurence_id		Convert to text Arbitrary id, per PCORNet does not need to be persistent across refreshes. Required
PATID	Condition_Occurrence.person_id		Convert to text. Required
ENCOUNTERID	Condition_Occurrence .visit_occurrence_id		Convert to text. Required
ENC_TYPE	Encounter.enc_type	Join to [target] Encounter table on Encounter.encounterid = Condition_Occurrence.visit_occurrence_id	Copied from ENCOUNTER record. Required
ADMIT_DATE	Encounter.admit_date	Join to Encounter table on Encounter.encounterid = Condition_Occurrence.visit_occurrence_id	Copied from ENCOUNTER record. Required
PROVIDERID	Encounter.provider_id	Join to encounter table on Encounter.encounterid = Condition_Occurrence.visit_occurrence_id	Convert to text Copied from ENCOUNTER record
DX	Condition_Occurrence .condition_source_value		PCORnet expects to see diagnosis codes as they were represented in the source system. Required
DX_TYPE	Derive from Concept .vocabulary_id	Join condition_source_concept_id to Concept.concept_id to get vocabulary_id OMOP to PCORnet Vocabulary Mapping ICD9CM 09 = ICD-9-CM ICD10CM 10 = ICD-10-CM SNOMED SM = SNOMED CT Otherwise	If a local ontology is used, and cannot be mapped to a standard ontology such as ICD-9-CM, DX_TYPE should be populated as "Other". Required.

		use 'OT' ('Other')	
DX_SOURCE	Condition_Occurrence .condition_status_concept_id	**OMOP to PCORnet** | 4203942	AD = Admitting
4230359	FI = Final/Discharge
4033240	IN = Interim
44814650	No Information
44814653	Unknown
44814649	Other		Required **OMOP Concepts** | 4203942	Admitting diagnosis
4230359	Final /discharge diagnosis
4033240	Preliminary diagnosis
44814650	No Information
0	Unknown
44814649	Other	
PDX	Derive from Condition_Occurrence .condition_type_concept_id	If condition_type_concept_id = 44786627 Then 'P' (Principal) Else If condition_type_concept_id = 44786629 Then 'S' (Secondary) Else If respective Visit_Occurrence.visit_concept_id are 9202 (Outpatient Visit) 9203 (Emergency Room Visit) 44814711 (Other ambulatory visit) Then 'X' (Unable to Classify) Else **OMOP to PCORnet** | 44814650	No Information
44814653	Unknown
44814649	Other		Principal discharge diagnosis flag. Relevant only on IP and IS encounters. Primary Condition (44786627) and Secondary Condition(44786629)
RAW_DX	Condition_Occurrence .condition_source_value		Load source values 'as is' - with source-specific suffixes and prefixes.
RAW_ DX_TYPE	Concept.vocabulary_id	Same as dx_type above	
RAW_ DX_SOURCE	Condition status source value	If de-dupping, should match the dx_source	

RAW_PDX	Concept.concept_name	If condition_type_concept_id IN(44786627, 44786629) join to Concept.concept_id Otherwise NULL	Primary Condition (44786627) Secondary Condition (44786629)

3.1.5 Table: Procedure

The PROCEDURE Table contains one record per unique combination of PATID, ENCOUNTERID, PX, and PX_TYPE. Because the date in the procedure table is that of the encounter, not necessarily when the procedure was performed, there may be multiples of the same procedure for the person/encounter/date when selecting from the OMOP procedure table. Duplicate records are also possible due to the mapping of one source code to multiple standard codes. These duplicated procedure records should be reduced to a single record in PCORnet based on procedure_occurrence.visit_occurrence_id and procedure_occurrence. procedure_source_value.

In OMOP CDM Procedure_Occurrence.visit_occurrence_id is optional, however PCORNet CDM specification requires mandatory encounter id for DIAGNOSIS and PROCEDURE. Exclude procedures where the visit_occurrence_id is NULL.

Procedure field mapping:

Destination Field	Source Field	Applied Rule	Comment
PROCEDURE SID			Required
PATID	Procedure_Occurrence.person_id		Convert to text. Requred
ENCOUNTERID	Procedure_Occurrence.visit_occurrence_id		Convert to text. Required
ENC_TYPE	Encounter.enc_type	Join to [target] Encounter table on Procedure_Occurrence.visit_occurrence_id = Encounter.encounterid	Copied from ENCOUNTER record. Required
ADMIT_DATE	Encounter.admit_date	Join to Encounter table on Procedure_Occurrence.visit_occurrence_id = Encounter.encounterid	Copied from ENCOUNTER record. Required
PROVIDERID	Encounter.provider_id	Join to encounter table on Procedure_Occurrence.visit_occurrence_id = Encounter.providerid	Convert to text Copied from ENCOUNTER record.
PX_DATE	procedure_occurrence.procedure_date		
PX	procedure_occurrence.procedure_source _value Otherwise Concept.concept_code	If procedure_source_concept_id is 44814649 (Other) or 0 (No matching concept), use procedure_occurrence. procedure_source_value. Otherwise	PCORnet expects to see all procedure codes as they were represented in the source system. Therefore, use source_concept_id or

		join procedure_ source_concept_id to Concept.concept_id.	source_value to represent PX in the source coding system. Required
PX_TYPE	Derive from Concept.vocabulary_id	Join procedure_source_concept_id to Concept.concept_id to get vocabulary_id **OMOP to PCORnet Vocabulary Mapping** ICD9CM — 09 = ICD-9-CM ICD9Proc — 09 = ICD-9-CM ICD10PCS — 10 = ICD-10-PCS CPT4 — C4 = CPT-4 (i.e., HCPCS Level I) HCPCS — HC = HCPCS (i.e., HCPCS Level II) LOINC — LC = LOINC NDC — ND = NDC Revenue Code — RE = Revenue PCORNet — OT = Other Otherwise Use 'OT' ('Other').	Required **OMOP Vocabulary Codes** ICD9CM ICD9Proc ICD10PCS CPT4 HCPCS LOINC NDC Revenue Code PCORNet
PX_SOURCE	procedure_occurrence.procedure_type_concept_id	If the procedure type is: any of the claims values set to CL EHR Order List, set to OD Hospital Cost Record, set to BI Otherwise set to OT	PCORnet values are OD=Order BI=Billing CL=Claim NI=No information UN=Unknown OT=Other
RAW_PX	Procedure_Occurrence.procedure_source_value		
RAW_PX_TYPE	Concept.concept_id	If source_condition_concept_id is 44814649 (Other) or 0 (No matching concept), use 'OT' ('Other'). Otherwise, join procedure_source_concept_id to Concept.concept_id	

3.1.6 Table: Vital

Measurements to Vital

Multiple measurements per encounter can be populated (for example, 3 blood pressure readings). There will be records where not all the vital statistics are defined. Create a record any time there is at least one of the attributes, weight, blood pressure, height or BMI is defined.
Vital signs data are sourced from OMOP Measurement and Observation tables.
Records corresponding to one visit may be grouped into one Vital record or represented as one Vital record per one vital sign.
Systolic and diastolic blood pressure coming from the same measurement must be by grouped into one record by utilizing Fact_Relationship link between the two records in the Measurement table as follows. Fact_id_1 and fact_id_2 should be equal to the respective measurement_id of diastolic and systolic BP records. Domain_concept_id_1 and domain_concept_id_2 should be equal to 21 ('Measurement'). Relationship_concept_id should be equal to 46233682 ('Diastolic to systolic blood pressure measurement').

OMOP Measurement to PCORnet VITAL field mapping:

Destination Field	Source Field	Applied Rule	Comment
VITALID			Required
PATID	Measurement.person_id		Convert to text. Required
ENCOUNTERID	Measurement.visit_occurrence_id or NULL		Convert to text Arbitrary encounter-level identifier. This is an optional relationship; the ENCOUNTERID should be present if the vitals were measured as part of healthcare delivery: Measurement.measurement_type_concept_id = 'Observation Recorded from EHR' (38000276).
MEASURE_DATE	Measurement. measurement_date		Required
MEASURE_TIME	Measurement.measurement_time	Text. Format as 'HH:MI' as 24 hours clock with zero-padding for hours and minutes.	
VITAL_SOURCE	Measurement. measurement_type_concept_id	OMOP to PCORnet | Concept Id	PCORnet Value
---	---
45754907	OT = Other
44818701	HC = Healthcare delivery setting			Concept	Description
---	---
45754907	Derived value
44818701	From physical examination
44818702	Lab result
44818703	Pathology finding	

Field	Source	Logic / Concept Mapping	Notes
		<table><tr><td>44818702</td><td>HC = Healthcare delivery setting</td></tr><tr><td>44818703</td><td>HC = Healthcare delivery setting</td></tr><tr><td>44818704</td><td>PR=Patient-reported</td></tr><tr><td>5001</td><td>HC = Healthcare delivery setting</td></tr><tr><td>0</td><td>OT = Other</td></tr></table>	<table><tr><td>44818704</td><td>Patient reported value</td></tr><tr><td>5001</td><td>Test ordered through EHR</td></tr></table> If multiple vital signs are compiled together in one record, Measurement.measurement_Type_Concept_ID must be the same. Required
HT	Measurement.value_as_number	Where Measurement .measurement_concept_id = 3036277 (Body height) Convert to inches.	
WT	Measurement.value_as_number	Where Measurement .measurement_concept_id = 3025315 (Body weight) Convert to pounds	
DIASTOLIC	Measurement.value_as_number	Where Measurement .measurement_concept_id in (3012888, 3034703, 3019962, 3013940)	
SYSTOLIC	Measurement.value_as_number	Where Measurement .measurement_concept_id in (3004249, 3018586, 3035856, 3009395)	
ORIGINAL_BMI	Measurement.value_as_number	Where Measurement .measurement_concept_id = 3038553 (Body mass index)	
BP_POSITION	derived from Measurement.measurement_concept_id	<table><tr><th>Concept</th><th>Description</th><th>PCORnet Value</th></tr><tr><td>3034703</td><td>Diastolic Blood Pressure - Sitting</td><td>'01'</td></tr><tr><td>3019962</td><td>Diastolic Blood Pressure - Standing</td><td>'02'</td></tr><tr><td>3013940</td><td>Diastolic Blood Pressure - Supine</td><td>'03'</td></tr><tr><td>3012888</td><td>Diastolic BP</td><td>'NI'</td></tr><tr><td>3018586</td><td>Systolic Blood Pressure - Sitting</td><td>'01'</td></tr><tr><td>3035856</td><td>Systolic Blood Pressure - Standing</td><td>'02'</td></tr></table>	Position when blood pressure taken is derived from the diastolic and systolic code provided.

| | | 3009395 | Systolic Blood Pressure - Supine | '03' | |
| | | 3004249 | Systolic BP | 'NI' | |

NULL if no blood pressure reading in this record.

	Measurement.measurement_source_value	Where Measurement .measurement_concept_id in (3012888, 3034703, 3019962, 3013940)	OMOP Concepts
			3012888 — BP diastolic
			3034703 — Diastolic blood pressure--sitting
			3019962 — Diastolic blood pressure--standing
RAW_ DIASTOLIC			3013940 — Diastolic blood pressure--supine
	Measurement.measurement_source_value	Measurement .measurement_concept_id in (3004249, 3018586, 3035856, 3009395)	OMOP Concepts
			3004249 — BP systolic
			3018586 — Systolic blood pressure--sitting
			3035856 — Systolic blood pressure--standing
RAW_ SYSTOLIC			3009395 — Systolic blood pressure--supine
RAW _ BP_ POSITION	NULL		Not available

Tobacco to Vital

Tobacco status and tobacco type Observation records are grouped into one VITAL record by *person_id* and *visit_occurrence_id* or *obsevation_date* where the *Observation.observation_concept_id* is set to 4041306 (Tobacco use and exposure). A scheme to convert these records into the PCORNET values for *smoking, tobacco* and *tobacco_type* is explained in the document 'A Solution for Mapping OMOP Observation to PCORnet Vital Smoking'. This will allow the ETL to consolidate the various possible tobacco related responses into a single set of values for PCORnet.

OMOP Observation to PCORnet VITAL field mapping for Tobacco:

Destination Field	Source Field	Applied Rule	Comment
VITALID			
PATID	Observation.person_id		Convert to text
ENCOUNTERID	Observation.visit_occurrence_id or NULL		Convert to text Arbitrary encounter-level identifier. This is an optional relationship; the ENCOUNTERID should be present if the vitals were measured as part of healthcare delivery: Observation.Observation_type_concept_id = 'Observation Recorded from EHR' (38000276).
MEASURE_DATE	Observation. Observation_date		
MEASURE_TIME	Observation.Observation_time	Text. Format as 'HH:MI' as 24 hours clock with zero-padding for hours and minutes.	
VITAL_SOURCE	Observation. Observation_type_concept_id	OMOP to PCORnet 44814721 — PR = Patient-reported 38000276 — HC = Healthcare delivery setting All other codes — OT = Other	Relevant OMOP concepts are: 'Patient reported' (44814721) or 'Observation Recorded from EHR' (38000276). If multiple vital signs are compiled together in one record, Observation.Observation_Type_Concept_ID must be the same.
SMOKING	Observation.value_as_concept_id	Where Observation.observation_concept_id = 4041306 (Tobacco use and exposure) Valid PCORnet values 01 — Current every day smoker 02 — Current some day smoker 03 — Former smoker 04 — Never smoker 05 — Smoker, current status unknown	See 'A Solution for Mapping OMOP Observation to PCORnet Vital Smoking'

		06	Unknown if ever smoked
		07	Heavy tobacco smoker
		08	Light tobacco smoker
		NI	No information
		UN	Unknown
		OT	Other

TOBACCO	Observation.value_as_concept_id	Where Observation.observation_concept_id = 4041306 (Tobacco use and exposure)	See 'A Solution for Mapping OMOP Observation to PCORnet Vital Smoking'

Valid PCORnet values

01	Current user
02	Never
03	Quit/former user
04	Passive or environmental exposure
06	Not asked
NI	No information
UN	Unknown
OT	Other

TOBACCO_TYPE	Observation.value_as_concept_id	Where Observation.observation_concept_id = 4041306 (Tobacco use and exposure)	See 'A Solution for Mapping OMOP Observation to PCORnet Vital Smoking'

Valid PCORnet values

01	01 = Cigarettes only
02	Non-smoked tobacco only
03	Cigarettes and other tobacco
04	None
05	Use of smoked tobacco but no information about

		non-smoked tobacco use	
	NI	No information	
	UN	Unknown	
	OT	Other	

RAW_SMOKING		Set to NULL	
RAW_TOBACCO		Set to NULL	
RAW_TOBACCO_TYPE		Set to NULL	

3.1.7 Table: LAB_RESULT_CM

The LAB_RESULT_CM table contains one record per LAB_RESULT_CM_ID.

Only records with actual lab results should be included in this table. If the results suggest that the test was run (e.g., result is "borderline") include it. But if the test is not resulted for any reason then do not include it.

The source for Lab results in OMOP CDM is the Measurement table, all measurement records where Measurement.measurement_type_concept_id is 44818702 ('Lab result'). In OMOP CDM, lab tests are represented by LOINC codes. The following LAB_RESULT_CM fields are derived from the LOINC codes: LAB_NAME, LAB_LOINC, SPECIMEN_SOURCE, and RESULT_UNIT. Mappings for selected LOINC codes between OMOP concepts and these attributes are presented in the table below. This list will have to be expanded as the list of target LOINC codes grows.

PCORnet Lab Name	OMOP Concept ID	LAB_NAME	LAB_LOINC	SPECIMEN_SOURCE	RESULT_UNIT
Troponin I cardiac	3021337	TROP_I	10839-9	SR_PLS	NG/ML
Creatinine kinase MB/creatinine kinase total	3007150	CK-MBI	12187-1	SR_PLS	PERCENT
Creatinine	3016662	CREATININE	12190-5	OT	MG/DL
Low-density lipoprotein	3028288	LDL	13457-7	SR_PLS	

Creatinine kinase MB	3005785	CK_MB	13969-1	SR_PLS	NG/ML
Low-density lipoprotein	3009966	LDL	18262-6	SR_PLS	MG/DL
Creatinine kinase MB/creatinine kinase total	3016311	CK-MBI	20569-0	SR_PLS	PERCENT
Low-density lipoprotein	3028437	LDL	2089-1	SR_PLS	MG/DL
Creatinine kinase total	3007220	CK	2157-6	SR_PLS	U/L
Creatinine	3016723	CREATININE	2160-0	SR_PLS	MG/DL
Low-density lipoprotein	3001308	LDL	22748-8	SR_PLS	
Creatinine kinase MB	3029790	CK_MB	32673-6	SR_PLS	U/L
Troponin T cardiac (qualitative)	3042837	Trop_T_QL	33204-9	SR_PLS	
Creatinine	3051825	CREATININE	38483-4	BLOOD	MG/DL
Troponin I cardiac	3033745	TROP_I	42757-5	BLOOD	NG/ML
Low-density lipoprotein	3046549	LDL	43727-7	SR_PLS	OT
Hemoglobin A1c	3004410	A1C	4548-4	BLOOD	PERCENT
Low-density lipoprotein	3053190	LDL	47213-4	SR_PLS	
Troponin T cardiac (qualitative)	3048529	Trop_T.QN	48425-3	BLOOD	UG/L

Troponin T cardiac (qualitative)	3052931	Trop_T_QL	48426-1	BLOOD	
Creatinine kinase MB/creatinine kinase total	3048863	CK-MBI	49136-5	SR_PLS	
Low-density lipoprotein	40757565	LDL	54434-6	SR_PLS	OT
Low-density lipoprotein	40758569	LDL	55440-2	SR_PLS	MG/DL
Creatinine kinase MB	3017761	CK_MB	5912-1	SR_PLS	
International normalized ratio	3022217	INR	6301-6	PPP	
Troponin T cardiac (qualitative)	3019572	Trop_T.QN	6597-9	BLOOD	UG/L
Troponin T cardiac (qualitative)	3019800	Trop_T.QN	6598-7	SR_PLS	UG/L
Hemoglobin	3000963	HGB	718-7	BLOOD	G/DL

Lab Result CM field Mapping: (assumes that non-OMOP-standard columns added for Order date and Result date/time)

Destination Field	Source Field	Applied Rule	Comment
LAB_RESULT_CM_ID	Measurement.measurement _id		Convert to text
PATID	Measurement.person_id		Convert to text
ENCOUNTERID	Measurement.visit_occurrence_id or NULL		Convert to text Arbitrary encounter-level identifier. This is an optional relationship; the ENCOUNTERID should be present if the labs were taken as part of healthcare delivery
LAB_NAME	Measurement.measurement_concept_id	Derived from Measurement. measurement_concept_id., see the mapping table above.	

SPECIMEN_SOURCE	Measurement.measurement_concept_id	Derived from Measurement. measurement_concept_id., see the mapping table above.	
LAB_LOINC	Measurement.measurement_concept_id	Derived from Measurement. measurement_concept_id., see the mapping table above.	
PRIORITY	measurement.priority		Optional column added to OMOP Measurement table
RESULT_LOC	NULL		Not populated
LAB_PX	NULL		Not populated
LAB_PX_TYPE	NULL		Not populated
LAB_ORDER_DATE	measurement.lab_order_date		Optional column added to OMOP Measurement table
SPECIMEN_DATE	Measurement.measurement_date		
SPECIMEN_TIME	Measurement.measurement_time	Text. Format as 'HH:MI' as 24 hours clock with zero-padding for hours and minutes.	
RESULT_DATE	Measurement.result_date		Optional column added to OMOP Measurement table
RESULT_TIME	Measurement.result time	Text. Format as 'HH:MI' as 24 hours clock with zero-padding for hours and minutes.	Optional column added to OMOP Measurement table
RESULT_QUAL	Measurement.value_as_concept_id		Not populated for the PCORnet required labs but may be populated for other labs.
RESULT_NUM	Measurement.value_as_number		
RESULT_MODIFIER	Concept.concept_code	Derived by linking Measurement.measurement_source_concept_id to Concept.concept_ id.	

Derived by linking Measurement.measurement_source_concept_id to Concept.concept_ id.

OMOP to PCORnet	
4171756	LT = Less than
4171754	LE = Less than or equal to
4172703	EQ = Equal
4172704	GT = Greater than

OMOP Concepts for Operators	
Concept ID	Concept Name
4171756	<
4171754	<=
4172703	=

		4171755	GE = Greater than or equal to

4172704	>
4171755	>=

Field	Source	Derivation	Notes
RESULT_UNIT	Measurement.measurement_concept_id	Derived from Measurement.measurement_concept_id., see the mapping table above.	
NORM_RANGE_LOW	measurement.range_low		
NORM_MODIFIER_LOW	NULL		Not populated
NORM_RANGE_HIGH	measurement.range_high		
NORM_MODIFIER_HIGH	NULL		Not populated
ABN_IND	measurement.value_as_concept_id	If the value as concept id is populated with values that indicate low, high or abnormal the map these to PCORnet values

Concept Id	PCORnet Value
4135493	AB=Abnormal
4267416	AL=Abnormally low
4328749	AH=Abnormally high
4069590	NL=Normal

If the value is NULL in OMOP set to NULL IN PCORnet
Otherwise set to OT = Other |

Concept	Descr
4135493	Abnormal
4267416	Low
4328749	High
4069590	Normal
RAW_ LAB_NAME	Measurement.measurement_source_value		
RAW_ LAB_CODE	Concept.concept_code	Derived by linking Measurement.measurement_source_concept_id to Concept.concept_id	
RAW_ PANEL	NULL		Not populated
RAW_ RESULT	Measurement.value_source_value		
RAW_ UNIT	Measurement.unit_source_value		
RAW_ ORDER_DEPT	NULL		
RAW_ FACILITY_CODE	NULL		

3.1.8 Table: CONDITION

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.

CONDITION should capture all records from the OMOP CDM where the Condition Type Concept is EHR problem list entry (38000245) or Patient Self-Reported Condition (45905770).

Exclude CONDITION row is the *condition_source_value* is a test description rather than a coding vocabulary value.

Condition field mapping:

Destination Field	Source Field	Applied Rule	Comment
CONDITIONID			
PATID	Condition_Occurrence.person_id		Convert to text
ENCOUNTERID	Condition_Occurrence.visit_occurrence_id		Convert to text
REPORT_DATE	Encounter.admit_date	Join to [target] Encounter table on Condition_Occurrence.visit_occurrence_id = Encounter.encounterid	Copied from ENCOUNTER record
RESOLVE_DATE	condition_occurrence.condition_end_date		
ONSET_DATE	condition_occurrence.condition_start_date		
CONDITION_STATUS	Derive from condition_end_date	If condition_end_date is null then 'AC' – Active Else 'RS' - Resolved	Condition Status = Inactive will not be populated.
CONDITION	Condition_Occurrence.condition_source_value	Exclude rows if the condition source value is a text description rather than a code.	
CONDITION_TYPE	Derive from Concept.vocabulary_id	Join source_condition_concept_id to Concept.concept_id to get vocabulary_id OMOP to PCORnet Vocabulary Mapping ICD9CM — 09 = ICD-9-CM ICD10CM — 10 = ICD-10-CM	OMOP Vocabularies ICD9CM ICD10CM SNOMED PCORNet

		SNOMED / SM = SNOMED CT PCORNet / OT = Other Otherwise use 'OT' ('Other')	
CONDITION_SOURCE	Derive from Condition_Occurrence.condition_type_concept_id	OMOP to PCORnet 38000245 / HC=Healthcare problem list 44819221 / PC=PCORnet-defined condition algorithm NI=No information 44814649 / Other	Concept / Desciption 38000245 / EHR problem list entry 44819221 / Patient-reported
RAW_CONDITION_STATUS	NULL		
RAW_CONDITION	Condition_Occurrence.condition_source_value		
RAW_CONDITION_TYPE	Observation.vocabulary_id	If condition_ source_concept_id is 44814649 ('Other'), use 'OT' ('Other'). Otherwise, join condition_source_concept_id to Concept.concept_id	
RAW_CONDITION_SOURCE	NULL		

3.1.9 Table: PRESCRIBING

Provider orders for medication dispensing and/or administration.

PRESCRIBING should capture all uniquely recorded in-patient medication dispensing and administration. The PRESCRIBING table in the PCORnet CDM is populated with all records in the DRUG EXPOSURE table with Drug_type_concept_id =
- 38000180 (Inpatient administration),
- 38000179 (Physician administered drug (identified as procedure)),

- 43542358 (Physician administered drug (identified from EHR observation)),
- 43542357 (Physician administered drug (identified from referral record)),
- 38000177 (Prescription written)

If a medication cannot be mapped to RxNorm, it should still be present

PRESCRIBING field mapping:

Destination Field	Source Field	Applied Rule	Comment
PRESCRIBINGID			
PATID	Drug_exposure.person_id		Convert to text
ENCOUNTERID	Drug_exposure.visit_occurrence_id		Convert to text
RX_PROVIDERID	Drug_exposure.provider_id		Convert to text
RX_ORDER_DATE	Drug_exposure.order_date		Order date added to Drug Exposure
RX_ORDER_TIME		Set to NULL	Not available in OMOP
RX_START_DATE	Drug_exposure.drug_exposure_start_date		
RX_END_DATE	Drug_exposure.drug_exposure_end_date		
RX_QUANTITY	Drug_exposure.quantity		
RX_REFILLS	Drug_exposure.refills		
RX_DAYS_SUPPLY	Drug_exposure.days_supply		
RX_FREQUENCY	drug_exposure.frequency	Map the frequency to the PCORnet values: 01=Every day 02=Two times a day (BID) 03=Three times a day (TID) 04=Four times a day (QID) 05=Every morning 06=Every afternoon 07=Before meals 08=After meals 09=As needed (PRN) NI=No information UN=Unknown OT=Other	Frequency added to drug exposure

RX_BASIC	Derive from drug_type_concept_id	If drug_type_concept_id is '38000177' Then 01 (Dispensing) If drug_type_concept_id is IN ('38000180',' 38000179',' 43542358',' 43542357') Then 02 (Administration) Else Other	
RXNORM_CUI	Concept.concept_code	Join to Concept table on drug_concept_id = concept_id	vocabulary_id = 'RxNorm'
RAW_RX_MED_NAME	Concept.concept_name	Join to Concept table on drug_concept_id = concept_id	vocabulary_id = 'RxNorm'
RAW_RX_FREQUENCY	Drug_exposure.effective_dose		
RAW_RXNORM_CUI	Concept.concept_code	Join to Concept table on drug_source_concept_id = concept_id	Not limited to RXNORM

3.1.10 Table: DISPENSING

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

DISPENSING should capture all uniquely recorded <u>outpatient</u> medication dispensing usually from claims data. The DISPENSING table in the PCORnet CDM is populated with all records in the DRUG EXPOSURE table with Drug_type_concept_id =
- 38000175 (Prescription dispensed in pharmacy),
- 38000176 (Prescription dispensed through mail order)

DISPENSING field mapping:

Destination Field	Source Field	Applied Rule	Comment
DISPENSINGID			
PATID	Drug_exposure.person_id		Convert to text
PRESCRIBINGID	NULL		How to link Precsribing ID to Dispensing ID? Per PCORNet: This is an optional relationship to the PRESCRIBING table, and may not be generally

			available. One prescribing order may generate multiple dispensing records.
DISPENSE_DATE	Drug_exposure.drug_exposure_start_date		
NDC	JOIN Drug_exposure WITH Vocabulary ON Drug_exposure.drug_source_concept_id = Vocabulary.Concept_ID If vocabulary.vocabulary_id = 'NDC' Then Vocabulary.concept_code Else Map RxNORM code in drug_concept_id to NDC	If one rxNORM code maps to multiple NDC codes, pick the one with lowest concept code.	
DISPENSE_SUP	Drug_exposure.day_supply		
DISPENSE_AMT	Drug_exposure.quantity		
RAW_NDC	Drug_exposure.drug_source_value		

3.1.11 Table: Death

The DEATH table contains one record per unique combination of PATID, DEATH_DATE, and DEATH_SOURCE.

Death field mapping:

Destination Field	Source Field	Applied Rule	Comment
PATID	Death.person_id		Convert to text
DEATH_DATE	Death.death_date		
DEATH_DATE_IMPUTE	'N'= Not Imputed		When date of death is imputed, this field indicates which parts of the date were imputed. Current assumption is that dates are not imputed.
DEATH_SOURCE	'L' = Other, Locally defined	"Other, locally defined" may be used to indicate presence of deaths reported from EHR systems, such as in-patient hospital deaths or dead on arrival.	Possible values: L=Other, locally defined N=National Death Index D=Social Security S=State Death files T=Tumor data NI=No information UN=Unknown OT=Other

DEATH_MATCH_ CONFIDENCE	'E' Excellent		Possible values: E=Excellent F=Fair P=Poor NI=No information UN=Unknown OT=Other

3.1.12 Table: Death Cause

The DEATH_CAUSE table contains one record per unique combination of PATID, DEATH_CAUSE, DEATH_CAUSE_CODE, DEATH_CAUSE_TYPE, and DEATH_CAUSE_SOURCE.
Only create a record when the cause_of_death_concept_id is defined in the OMOP death table.

Death Cause field mapping:

Destination Field	Source Field	Applied Rule	Comment
PATID	Death.person_id		Convert to text
DEATH_CAUSE	Death.cause_of_death_source_value		
DEATH_CAUSE_ CODE	Death.cause_of_death_concept_id	Join to vocabulary.concept and derive from vocabulary_id. Use the following: 2 – '09 34 – '10' 0 – 'UN' Else 'OT'	Cause of death code type. Possible values: 09=ICD-9 10=ICD-10 NI=No information UN=Unknown OT=Other
DEATH_CAUSE_ TYPE	'NI'= No Information		Possible values: C=Contributory I=Immediate/Primary O=Other U=Underlying NI=No information UN=Unknown OT=Other
DEATH_CAUSE_ SOURCE	'L' = Other, Locally defined	"Other, locally defined" may be used to indicate presence of deaths reported from EHR systems, such as in-patient hospital deaths or dead on arrival.	Possible values: C=Contributory I=Immediate/Primary O=Other U=Underlying NI=No information UN=Unknown OT=Other

DEATH_CAUSE_ CONFIDENCE	F=Fair		Possible values: E=Excellent F=Fair P=Poor NI=No information UN=Unknown OT=Other

4.0 Outstanding Issues

Immediate
1. Check with OHDSI if concept 44814723 has been corrected: 'Period while enrolled in study' should be changed to 'Geography based'.

Parking lot
1. TBD: Do we transfer to PCORI 'invalid' DX's? For example '250.x' Indicates diabetes but it is not a valid ICD9 code. Those will be stored in OMOP with source_concept_id=0. If yes, will we have DX_TYPE of 'Other'? Need PCORnet feedback

---

[dmt1]Change to agree with updated conventions document
[dmt2]Change to agree with current conventions