Common Data Model (CDM) Specification, Version 3.1

1. Table of Contents

Important Links and References (section added in v3.1):

The PCORnet CDM documentation can be accessed online at: http://www.pcornet.org/resource-center/pcornet-common-data-model/

Note to programmers: The separate "CDM parseable file" is more helpful for direct use in implementation, and contains the complete table specifications. All documentation is available here: http://www.pcornet.org/pcornet-common-data-model/

View useful tools for the CDM, such as the CDM-ERRATA and CDM-GUIDANCE issue trackers, on the PCORnet GitHub CDM Forum: https://github.com/CDMFORUM

A lay guide and glossary of terms for this document and can be accessed online at: http://www.pcornet.org/pcornet-common-data-model/

For more information about PCORnet, please visit http://www.pcornet.org/

2. Overview of the PCORnet Common Data Model (CDM)

2.1. License and Use

The PCORnet Distributed Research Network (DRN) operations center and infrastructure, including the Common Data Model (CDM), is led by the PCORnet Coordinating Center and overseen by PCORnet governance including the Data Committee, Executive Committee, and Council. (Description updated in v3.1.)

The PCORnet CDM was originally based on the Mini-Sentinel Common Data Model v4.0 (MSCDM v4.0; www.mini-sentinel.org) and has been informed by other distributed initiatives such as the HMO Research Network, the Vaccine Safety Datalink, various AHRQ Distributed Research Network projects, and the ONC Standards & Interoperability Framework Query Health Initiative. The PCORnet CDM is positioned within healthcare standard terminologies (including ICD, SNOMED, CPT, HCPCS, and LOINC®) to enable interoperability with and responsiveness to evolving data standards.

2.2. Introduction (updated in v3.1)

What is the CDM?

The PCORnet Common Data Model (CDM) is a ***specification*** *that defines a* ***standard organization*** *and* ***representation*** *of data for the PCORnet Distributed Research Network.*

The PCORnet CDM is a key component of the PCORnet Distributed Research Network (DRN) infrastructure. PCORnet developed the PCORnet DRN design to be a "…functional distributed research network that facilitates multi-site patient–centered research across the Clinical Data Research Networks (CDRNs), Patient-Powered Research Networks (PPRNs), and other interested contributors. The distributed network will enable the conduct of observational research and clinical trials while allowing each participating organization to maintain physical and operational control over its data." [Data Standards, Security, and Network Infrastructure Task Force (DSSNI charter), 2014]

For more details of CDM development, additional references include:

- CDM abstracts presented at scientific conferences: https://github.com/CDMFORUM/CDM-GUIDANCE/wiki/CDM-related-Abstracts
- CDM Forum materials: https://github.com/CDMFORUM/CDM-GUIDANCE/wiki/CDM-Forum-Materials

2.3. History of Releases and Modifications

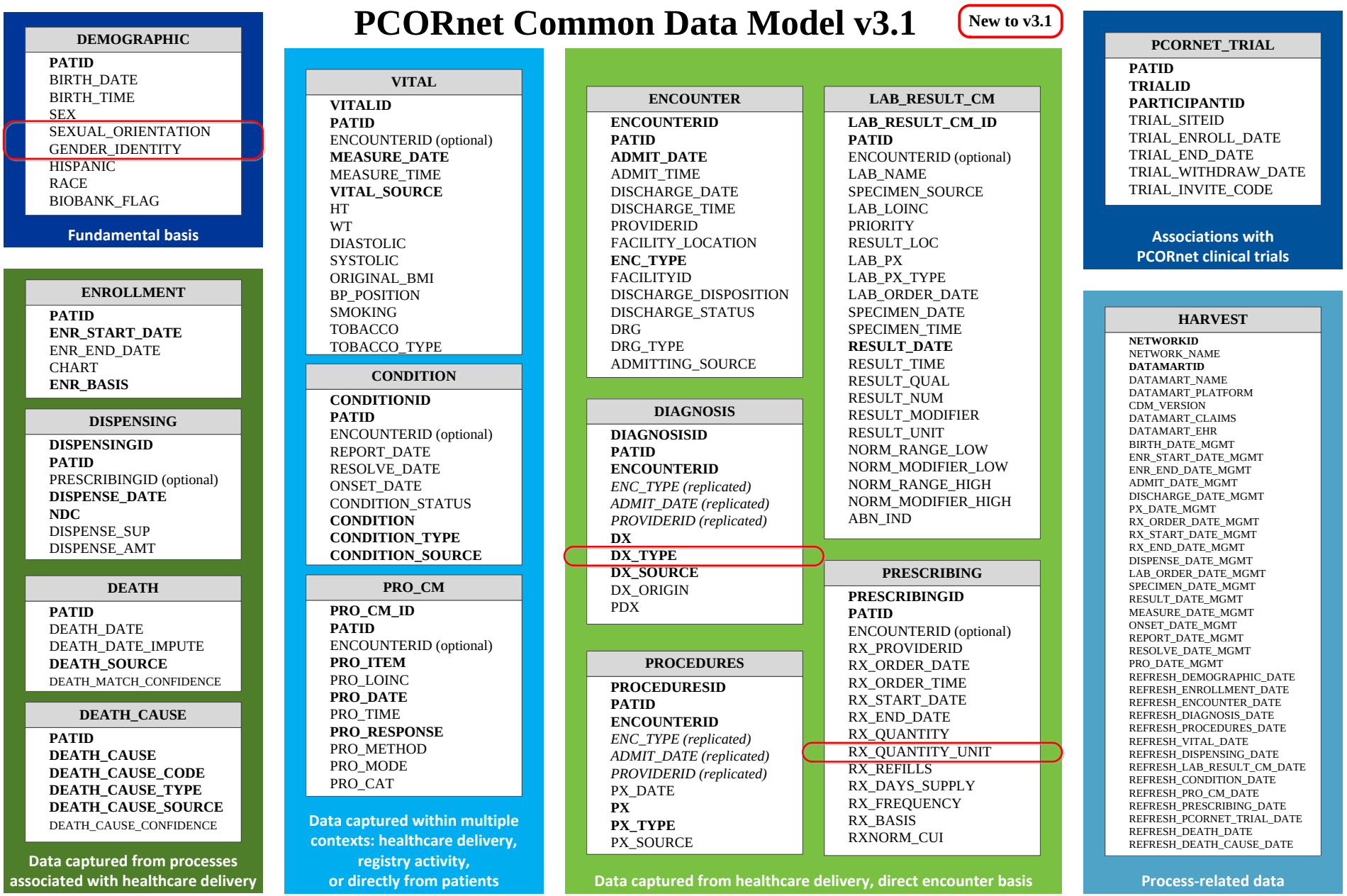
Note on version conventions: Major releases are denoted with whole number incrementation (eg, v1.0, v2.0, v3.0). Minor releases are denoted with decimal incrementation (eg, v1.1, v1.2) and will be used for bug fixes and minor adjustments.

Reference Table: History of Releases

Version	*Date of Release*	*Description of Release*
v1.0	2014-05-30	The DSSNI Task Force thanks the many individuals who provided thoughtful feedback, comments, and suggestions for this first release of the PCORnet CDM. A special thanks to members of the task force who volunteered to serve on the CDM working group.
v2.0	2015-02-27	The v2.0 release includes: • Four new tables (DISPENSING, CONDITION, PRO_CM, LAB_RESULT_CM) • Four new fields in existing tables (VITAL.TOBACCO, VITAL.TOBACCO_TYPE, PROCEDURE.PX_TYPE, PROCEDURE.PX_SOURCE) • Additional guidance and descriptions
v3.0	2015-06-01	The v3.0 release includes: • Five new tables (PRESCRIBING, PCORNET_TRIAL, DEATH, DEATH_CAUSE, and HARVEST) • Ten new fields in existing tables (DISPENSING.DISPENSINGID, DISPENSING.PRESCRIBINGID, VITAL.VITALID, VITAL.SMOKING, CONDITION.CONDITIONID, CONDITION.ONSET_DATE, PRO_CM.PRO_CM_ID, DIAGNOSIS.DIAGNOSISD, PROCEDURES.PROCEDURESID, LAB_RESULT_CM.LAB_RESULT_CM_ID) • Modification to relational integrity specifications • Modification to date formatting practices • New specifications specific to SAS data types • Additional guidance, clarifications, and descriptions
v3.0	2015-07-29	Document updated with licensing information and new PCORnet.org URL. No technical specifications have been modified.

Reference Table: History of Releases		
Version	*Date of Release*	*Description of Release*
v3.1	2016-11-15	Please note: New and modified fields have been indicated in green to assist with visually scanning the document (in addition to the descriptive comments). The v3.1 release includes: • Four new fields (DEMOGRAPHIC.SEXUAL_ORIENTATION, DEMOGRAPHIC.GENDER_IDENTITY, DIAGNOSIS.DX_ORIGIN, PRESCRIBING.RX_QUANTITY_UNIT) • Encounter types value set expanded to include observation stays and institutional professional consults • Collapsed value set of procedure terminologies so that CPT and HCPCS are grouped into single category • Clarified expected number of digits for RDBMS number formatting • Date of death no longer a required field for DEATH table • Enrollment table basis now includes drug coverage

2.4. Overview Diagram



http://www.pcornet.org/pcornet-common-data-model/

Bold font indicates fields that cannot be null due to primary key definitions or record-level constraints.

2.5. Implementation Expectations

CDM Table	*CDRN Expectation**	*PPRN Expectation**
DEMOGRAPHIC	Expected	Expected
ENROLLMENT	Expected	Optional
ENCOUNTER	Expected	Optional
DIAGNOSIS	Expected	Optional
PROCEDURES	Expected	Optional
VITAL	Optional	Optional
DISPENSING	Optional	Optional
LAB_RESULT_CM	Optional	Optional
CONDITION	Optional	Expected
PRO_CM	Optional	Optional
PRESCRIBING	Optional	Optional
PCORNET_TRIAL	Expected for PCORnet trials	Optional
DEATH	Optional	Optional
DEATH_CAUSE	Optional	Optional
HARVEST	Expected	Expected

*Any table may be required for a given PCORnet study or trial.

In v3.1, the VITAL table has been changed to Optional for CDRN Expectation because this domain would not be expected from claims-based datamarts.

Please note that all tables must be present in an instantiation of the CDM even if the table is empty. This is important because some components of the querying platform need to locate a given table, even if zero records are present in that table. This applies even if the table is not expected. (Guidance added in v3.1.)

3. Design of the CDM

3.1. Special Topics for CDM Modeling

Prioritization of Analytic Functionality

> PCORnet CDM Guiding Principle #5 states,
>
> "Documentation will be clear and transparent so that its contents are understandable to all contributors. The CDM will be intuitive and easy for analysts and investigators to use. **Investigators and analysts with prior experience using research data will not need additional skills or knowledge to use the CDM."** [emphasis added]

This guiding principle is expressed in the CDM design through prioritization of **analytic** functionality, and a parsimonious approach based upon analytic utility. At times, this results in decisions that are not based in relational database modeling principles such as normalization. The model is designed to facilitate routine and rapid execution of distributed complex analytics. To meet this design requirement, some fields are duplicated across multiple tables to support faster analytic operations for distributed querying. The PCORnet CDM is based on the FDA Mini-Sentinel CDM. This allows PCORnet to more easily leverage the large array of analytic tools and expertise developed for the MSCDM v4.0, including data characterization approaches and the various tools for complex distributed analytics.

Relational Integrity

Database programmers will notice that fields used as primary/foreign keys, especially PATID and ENCOUNTERID, are specified as text instead of numbers. This approach, informed by prior experience in developing large-scale multi-site distributed networks, makes it easier to implement than requiring new key generation that could impact database management within source systems.

Please note that all tables must be present in an instantiation of the CDM, even if data are not populated in every table.

Date Formatting

Because the PCORnet CDM is intended to support multiple Relational Database Management Systems (RDBMS), date format consistency is an issue, given that most RDBMS's have platform-specific native date representation.

To address this issue, each RDBMS will be expected to implement its own native date data type for dates, which will be supported by the Entity Framework technology stack[1]. The CDM will always separate date fields and time fields for consistency, and employ a naming convention of suffix "_DATE" or "_TIME".

All times should be recorded within the local time zone. A uniform time stamp or GMT offset is not expected.

A SAS date is different from an RDBMS date. A SAS date is a value that represents the number of days between January 1, 1960 and the specified date. SAS can perform calculations on dates ranging from A.D. 1582 to A.D. 19,900. Dates before January 1, 1960, are negative numbers; dates after are positive numbers. (Guidance added in v3.1.)

Number Formatting (guidance added in v3.1)

Numeric data elements in v3.1 have been updated to indicate **both** precision (total number of digits) **and** scale (digits to the right of a decimal point) for RDBMS Number data types. For example, **RDBMS Number(15,8)** specifies a precision of 15 and scale of 8. The maximum precision used in the CDM was chosen to harmonize with SAS numbers.

However, the indication of SAS Numeric(8) is not the same as an RDBMS precision/scale. SAS Numeric(8) indicates a *byte length* of 8; of note, the length is not the same as the number of significant digits contained within this variable in SAS. SAS numeric variables (including integers and dates) are almost always stored in 8-byte floating-point form. These numbers have a precision of about 16 significant digits.

In the CDM specification, data elements will never have an RDBMS size that is *larger* than the SAS size. This will ensure that there is no loss of data when outputting a SAS dataset from an RDBMS source.

"RDBMS Number" can be implemented as any appropriate RDBMS number concept, such as DECIMAL or DOUBLE data types. Although some RDBMS's have a specific data type called "NUMBER" (such as Postgres), the CDM does not imply that this specific data type should be implemented.

In CDM version 3.1, to prevent further confusion, the specifications for numeric data types have been updated to make them consistent between RDBMS and SAS. Most importantly, this will prevent any loss of data when converting from RDBMS to SAS (or vice versa).

Original values should not be modified to add additional, artificial decimal precision. Original values that are integers should not be modified to add artificial decimal precision. For example:

- The value **1. 1** would **never** be modified to become **1.10000000**
- The integer value of **1** would **never** be modified to become **1.0** or **1.10000000**

In summary:

[1] https://msdn.microsoft.com/en-us/data/ef.aspx

- RDBMS Number(15,8) is a RDBMS number concept with up to 15 total digits and up to 8 digits to the right of the decimal
- SAS Numeric(8) is a floating point number with 15 total digits stored in 8 bytes; because it is a floating point number, the digits to the right of the decimal are not specified

Missing or Unknown Data Values
The PCORnet CDM will use the HL7 conventions of "Null Flavors" (http://hl7.org/implement/standards/fhir/v3/NullFlavor/) as a basis for representing missing or unknown values. Specifically, for fields where an enumeration is present (i.e., a categorical set of values), we will populate null values as follows:

1. **A data field is not present** in the source system. (populate with null)
2. A data field for an enumeration is present in the source system, **but the source value is null or blank**. (populate with NI=No Information)
3. A data field for an enumeration is present in the source system, but the source value **explicitly denotes an unknown value**. (populate with UN=Unknown)
4. A data field for an enumeration is present in the source system, but the source value **cannot be mapped to the CDM**. (populate with OT=Other)

This guidance is only applicable for categorical text fields, not for numbers or dates.

Source Data Consistency
The CDM does not include data consistency rules or edits, such as upper and lower limits of numeric values. The value recorded in the originating source system should be the value populated in the CDM, even if the value is outside a normally acceptable limit. Inclusion of all originating data, without modification, supports data characterization and better data provenance.

Decisions about inclusion (or censoring) of outlier values will be made as part of each analysis or query, allowing for these decisions to be driven by appropriateness for each individual analysis.

> PCORnet CDM Guiding Principle #7 states,
>
> "**The CDM will reflect variables and values found in the local data**. If some data are coded in a way that is unique to a site, mapping the data to a standardized format will be necessary. **Values in the source data before mapping will also be included in the CDM**. Derived variables should be avoided." [emphasis added]

"Raw" Fields
The data model uses a convention for "raw data fields." These are optional fields for storing the originating source value of a field, prior to mapping into PCORnet CDM value set. It may also be used for source-specific ontologies.

The "RAW" fields are intended to support data provenance and facilitate quality control checking by local implementation, if desired. These fields will have a naming convention of prefix "RAW_". We will not include these fields in the Entity-Relationship (ER) diagram.

Case Sensitivity
All RDBMS implementations of the PCORnet Common Data Model should be case-insensitive.

Avoidance of Padding
Numbers should not be "padded" with extra zeroes. Text fields should not be "padded" with spaces before or after the actual textual values.

Additional Fields
PCORnet sites are welcome to include additional fields in their local CDM implementation that will assist with transformation or clarity.

> As stated in PCORnet CDM Guiding Principle #8,
> "CDRNs and PPRNs may include additional domains and data elements in localized versions of the PCORnet CDM."

Incomplete Date Guidance
In situations where the exact day or month is unknown or not available, it is still necessary to have a valid date for native RDBMS and SAS date data types. In this situation, please use this specific imputation strategy:

- If the day is missing, use the **first day of the month** to create a valid date value with the existing month and year.
- In the uncommon situation where a month is missing, **use January 1** to create a value date value with the existing year.

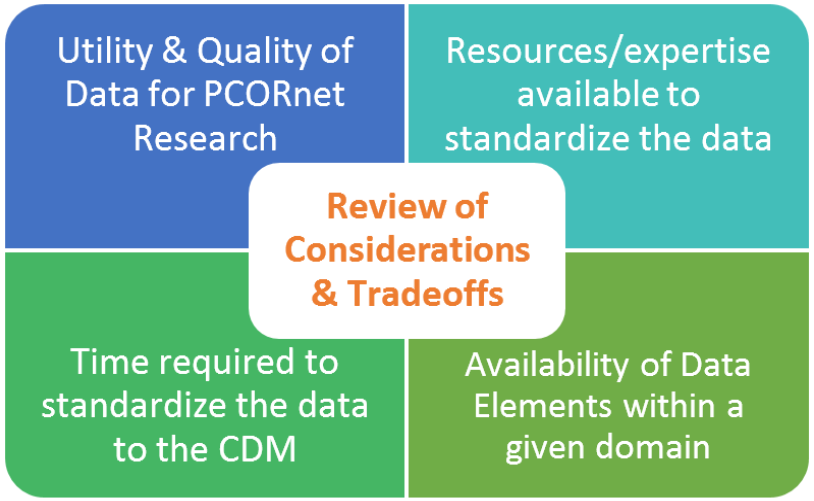
The HARVEST table indicators of DATE_ fields are used to indicate the presence of incomplete dates within the data, and the specific details of imputation would be described in the ETL Annotated Data Dictionary (ADD). The convention of the RAW_ fields can also be deployed to indicate the presence and original value of incomplete dates, if desired.

3.2. Development Notes (updated in v3.1)

PCORnet CDM Guiding Principle #2 states,

> "**It is expected that <u>not all</u> CDRNs and PPRNs will have data needed to populate all parts of the PCORnet CDM.** It is the responsibility of the CDRNs and PPRNs to communicate availability of each data domain and element." [emphasis added]

The PCORnet CDM will be implemented in phases. This will allow incorporation of new data domains and fields throughout the life of the project, building based on PCORnet needs, lessons learned from use, and data availability. The assessment of considerations and tradeoffs is an integral part of decision-making based on pragmatism and analytic value.



Because the PCORnet DRN has independent objectives and priorities, the PCORnet CDM will not reuse an existing data model, but will develop a stand-alone PCORnet CDM based on existing data models, as appropriate.

PCORnet CDM Guiding Principle #6 states,

> "Other common data elements and common data model initiatives exist. PCORnet will draw from the experience of others within and outside of PCORI, leveraging existing successful approached and data model definitions wherever possible."

The model was initially informed by results from the PCORnet DSSNI Preliminary Partner Survey (also known as the "Tech Survey") completed in December 2013 and January 2014. Recommendations from the PCORnet CDM Working Group have been a basis for strategy and decisions. The PCORnet CDM priority data domains and implementation approach are based on PCORI needs, planned future capabilities, and the data sources and expertise of the PCORnet partners.

> As stated in PCORnet CDM Guiding Principle #4,
>
> "The PCORnet CDM will be developed in a **modular, incremental, and extensible fashion.** New types of data will be needed, or newly available, during the life of PCORnet. Data domains and data elements will be added, revised, and deprecated throughout an iterative CDM lifecycle." [emphasis added]

3.3. Comments on Protected Health Information (PHI)

The CDM will contain some of the 18 elements that define PHI under HIPAA, including encounter dates and date of birth. However, these dates will remain under the control of the institutions that already maintain PHI. To maximize analytic flexibility and allow for all types of analyses, complete and exact dates should be included in the CDM. Distributed analytic programs will use the date fields for analysis, but will generate results that contain the minimum necessary information to address the question. The results returned to the requester will typically be aggregated and not include any PHI. Queries that generate results sets with PHI (eg, a person-level analysis under an IRB, with all necessary data agreements in place) will be clearly flagged as such and will only be distributed with the appropriate approvals clearly documented. As with all distributed queries, sites should review all results before release.

> PCORnet Distributed Research Network Guiding Principle #2 states,
>
> "**CDRNs and PPRNs will control how their data are used as allowed by internal governance policies**. Data resources developed for PCORnet will stay within the CDRNs and PPRNs and under their control." [emphasis added]

The necessary "cross-walks" between the arbitrary identifiers included in the CDM and their originating data are not specified in the scope of the CDM, but are expected to be maintained by each data partner.

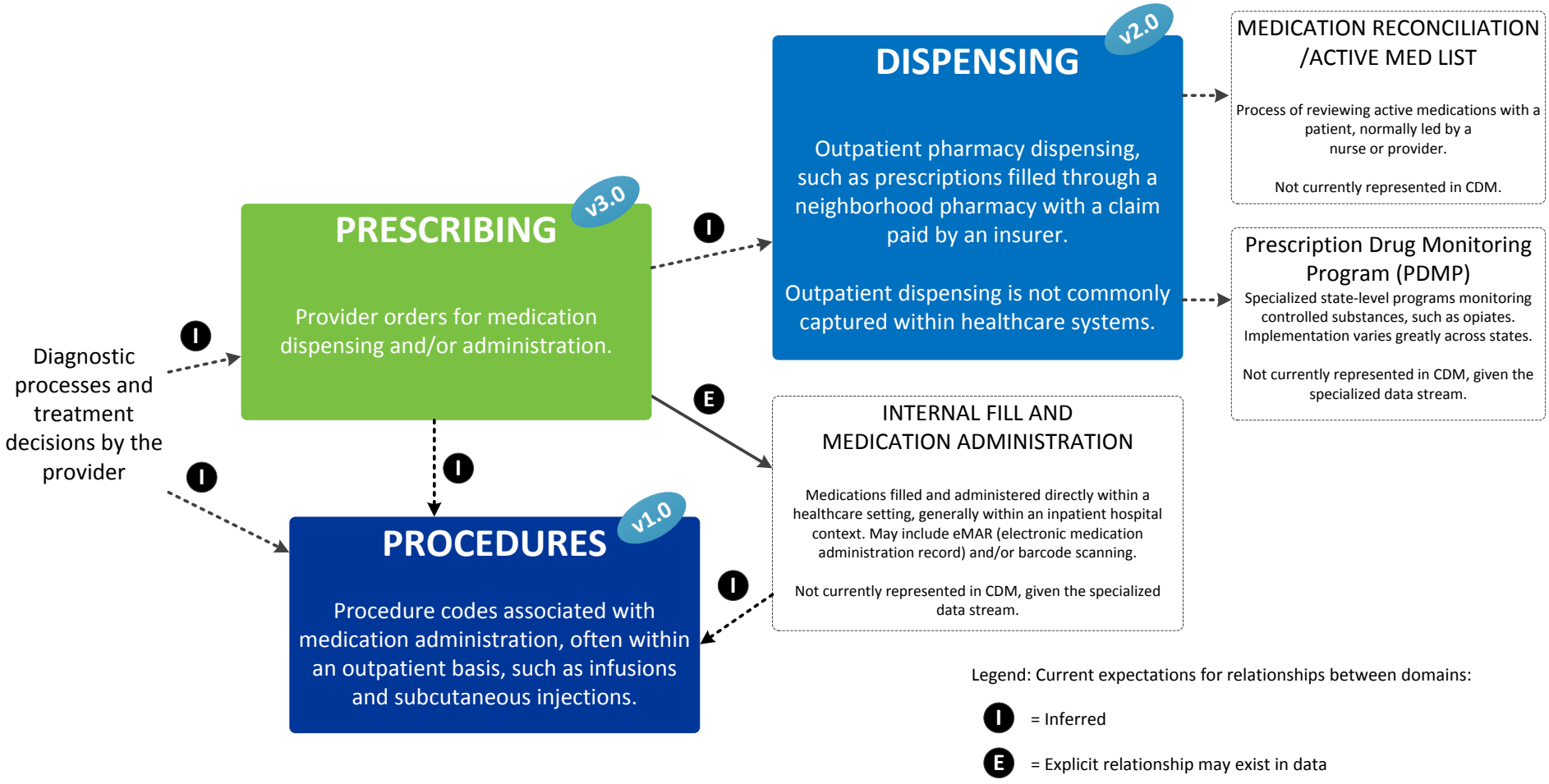
- PATID is a pseudoidentifier with a consistent crosswalk to the true identifier retained by the source site. For analytical data sets requiring patient-level data, only the pseudoidentifier is used to link across all information belonging to a patient.
- The PATID pseudoidentifier should not be a true identifier. It is not appropriate to use Medical Record Identifiers (MRNs) for this purpose because MRN is a direct patient identifier.
- Locally maintained "mapping tables" are tables necessary to implement so that each data partner has the ability to map arbitrary identifiers back to the originating data and patient.
- These mapping tables are not part of the PCORnet DRN.

Mapping tables for implementation of the CDM should include (but are not limited to):

- PATID crosswalk
- PROVIDER crosswalk

3.4. The Continuum of Medication-related Data Domains

This diagram represents our expectations for the **current state** of medication-related data stores in clinical systems, and is meant to assist in the assessment of data availability for PCORnet CDM implementation.



4. Individual Table Specifications

4.1. Table: DEMOGRAPHIC

DEMOGRAPHIC Domain Description:
Demographics record the direct attributes of individual patients.

Relational Integrity:
The DEMOGRAPHIC table contains one record per patient.

Primary Key: PATID

Constraints:
PATID (unique; required, not null)

DEMOGRAPHIC Table Specification

Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier used to link across tables. PATID is a pseudoidentifier with a consistent crosswalk to the true identifier retained by the source data partner. For analytical data sets requiring patient-level data, only the pseudoidentifier is used to link across all information belonging to a patient. The PATID must be unique within each PCORnet data mart. Creating a unique identifier within a network would be beneficial and acceptable. The PATID is not the basis for linkages across data partners.	MSCDM v4.0
BIRTH_DATE	RDBMS Date	SAS Date (Numeric)	.	Date of birth.	MSCDM v4.0

DEMOGRAPHIC Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
BIRTH_TIME	RDBMS Text(5): Format as HH:MI using 24-hour clock and zero-padding for hour and minute	SAS Time (Numeric)	.	Time of birth.	PCORnet Source of time format: ISO 8601
SEX	RDBMS Text(2)	SAS Char(2)	A=Ambiguous F=Female M=Male NI=No information UN=Unknown OT=Other	Sex assigned at birth (definition updated in v3.1).	MSCDM v4.0 with modified field size and value set Source: Administrative Sex (HL7) http://phinvads.cdc.gov/vads/ViewValueSet.action?id=06D34BBC-617F-DD11-B38D-00188B398520
SEXUAL_ORIENTATION	RDBMS Text(2)	SAS Char(2)	AS=Asexual BI=Bisexual GA=Gay LE=Lesbian QU=Queer QS=Questioning ST=Straight SE=Something else MU=Multiple sexual orientations DC=Decline to answer NI=No information UN=Unknown OT=Other	Sexual orientation. (New field added in v3.1.)	Source: Health IT Certification Criteria, 2015 Base Edition, modified with expert advisory within PCORnet https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base

DEMOGRAPHIC Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
GENDER_IDENTITY	RDBMS Text(2)	SAS Char(2)	M=Man F=Woman TM=Transgender male/Trans man/Female-to-male TF=Transgender female/Trans woman/Male-to-female GQ=Genderqueer SE=Something else MU=Multiple gender categories DC=Decline to answer NI=No information UN=Unknown OT=Other	Current gender identity. (New field added in v3.1.)	Source: Health IT Certification Criteria, 2015 Base Edition, modified with expert advisory within PCORnet https://www.federalregister.gov/documents/2015/10/16/2015-25597/2015-edition-health-information-technology-health-it-certification-criteria-2015-edition-base
HISPANIC	RDBMS Text(2)	SAS Char(2)	Y=Yes N=No R=Refuse to answer NI=No information UN=Unknown OT=Other	A person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race.	MSCDM v4.0 with modified field size and value set Compatible with "OMB Hispanic Ethnicity" (Hispanic or Latino, Not Hispanic or Latino)

DEMOGRAPHIC Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RACE	RDBMS Text(2)	SAS Char(2)	01=American Indian or Alaska Native 02=Asian 03=Black or African American 04=Native Hawaiian or Other Pacific Islander 05=White 06=Multiple race 07=Refuse to answer NI=No information UN=Unknown OT=Other	Please use only one race value per patient. **Details of categorical definitions:** American Indian or Alaska Native: A person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment. Asian: A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam. Black or African American: A person having origins in any of the black racial groups of Africa. Native Hawaiian or Other Pacific Islander: A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands. White: A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.	MSCDM v4.0 with modified field size and value set Original value set is based upon U.S. Office of Management and Budget (OMB) standard, and is compatible with the 2010 U.S. Census
BIOBANK_FLAG	RDBMS Text(1)	SAS Char(1)	Y=Yes N=No	Flag to indicate that one or more biobanked specimens are stored and available for research use. Examples of biospecimens could include blood, urine, or tissue (eg, skin cells, organ tissues). If biospecimens are available, locally maintained "mapping tables" would be necessary to map between the DEMOGRAPHIC record and the originating biobanking system(s). If no known biobanked specimens are available, this field should be marked "No".	PCORnet
RAW_SEX	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_ SEXUAL_ORIENTATION	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_ GENDER_IDENTITY	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

DEMOGRAPHIC Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RAW_HISPANIC	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_RACE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

4.2. Table: ENROLLMENT

ENROLLMENT Domain Description:

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

(Domain description updated in v3.1.)

Relational Integrity:
The ENROLLMENT table contains one record per unique combination of PATID, ENR_START_DATE, and ENR_BASIS.

Please note: Each form of coverage (the ENR_BASIS) would have a separate record; for example, if a patient has both medical coverage and drug coverage, these would be 2 separate records, potentially with different enrollment dates for each record (new guidance added in v3.1).

Composite Primary Key: PATID, ENR_START_DATE, ENR_BASIS

Foreign Key:
ENROLLMENT.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)

Constraints:
PATID + ENR_START_DATE + ENR_BASIS (unique)
PATID (required, not null)
ENR_START_DATE (required, not null)
ENR_BASIS (required, not null)

ENROLLMENT Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier used to link across tables.	MSCDM v4.0
ENR_START_DATE	RDBMS Date	SAS Date (Numeric)	.	Date of the beginning of the enrollment period. If the exact date is unknown, use the first day of the month.	MSCDM v4.0 with modified field name
ENR_END_DATE	RDBMS Date	SAS Date (Numeric)	.	Date of the end of the enrollment period. If the exact date is unknown, use the last day of the month.	MSCDM v4.0 with modified field name
CHART	RDBMS Text(1)	SAS Char(1)	Y=Yes N=No	Chart abstraction flag is intended to answer the question, "Are you able to request (or review) charts for this person?" This flag does not address chart availability. Note: This field is most relevant for health insurers that can request charts from affiliated providers. This field allows exclusion of patients from studies that require chart review to validate exposures and/or outcomes. It identifies patients for whom charts are **never** available and for whom the chart can never be requested.	MSCDM v4.0

ENROLLMENT Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
ENR_BASIS	RDBMS Text(1)	SAS Char(1)	I=Medical insurance coverage D=Outpatient prescription drug coverage G=Geography A=Algorithmic E=Encounter-based	ENR_BASIS is a property of the time period defined. A patient can have multiple entries in the table. **Details of categorical definitions:** Medical insurance coverage: The start and stop dates are based upon enrollment where the health plan has any responsibility for covering medical care for the member during this enrollment period (i.e., if you expect to observe medical care provided to this member during the enrollment period). Outpatient prescription drug coverage: The start and stop dates are based on enrollment where the health plan has any responsibility for covering outpatient prescription drugs for the member during this enrollment period (i.e., if you expect to observe outpatient pharmacy dispensings for this member during this enrollment period). (New value set item added in v3.1.) Geography: An assertion of complete data capture between the start and end dates based upon geographic characteristics, such as regional isolation. Algorithmic: An assertion of complete data capture between the start and end dates, based on a locally developed or applied algorithm, often using multiple criteria. Encounter-based: The start and stop dates are populated from the earliest-observed encounter and latest-observed encounter. Field definition and value sets modified in v3.1 to include drug coverage.	PCORnet Based upon the HMORN VDW and Sentinel CDM v6.0

4.3. Table: ENCOUNTER

ENCOUNTER Domain Description:
Encounters are interactions between patients and providers within the context of healthcare delivery.

Relational Integrity:
The ENCOUNTER table contains one record per unique encounter.

Primary Key: ENCOUNTERID

Foreign Key:
ENCOUNTER.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)

Constraints:
ENCOUNTERID (unique; required, not null)
PATID (required, not null)
ADMIT_DATE (required, not null)
ENC_TYPE (required, not null)

ENCOUNTER Table Specification

Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
ENCOUNTERID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary encounter-level identifier. Used to link across tables, including the ENCOUNTER, DIAGNOSIS, and PROCEDURES tables.	MSCDM v4.0
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier used to link across tables.	MSCDM v4.0
ADMIT_DATE	RDBMS Date	SAS Date (Numeric)	.	Encounter or admission date.	MSCDM v4.0 with modified field name

ENCOUNTER Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
ADMIT_TIME	RDBMS Text(5): Format as HH:MI using 24-hour clock and zero-padding for hour and minute	SAS Time (Numeric)	.	Encounter or admission time.	PCORnet Source of time format: ISO 8601
DISCHARGE_DATE	RDBMS Date	SAS Date (Numeric)	.	Discharge date.	MSCDM v4.0 with modified field name
DISCHARGE_TIME	RDBMS Text(5): Format as HH:MI using 24-hour clock and zero-padding for hour and minute	SAS Time (Numeric)	.	Discharge time.	PCORnet Source of time format: ISO 8601
PROVIDERID	RDBMS Text(x)	SAS Char(x)	.	Provider code for the provider who is most responsible for this encounter. For encounters with multiple providers choose one so the encounter can be linked to the diagnosis and procedure tables. As with the PATID, the provider code is a pseudoidentifier with a consistent crosswalk to the real identifier.	MSCDM v4.0
FACILITY_LOCATION	RDBMS Text(3)	SAS Char(3)	.	Geographic location (3 digit zip code).	MSCDM v4.0

ENC_TYPE	RDBMS Text(2)	SAS Char(2)	AV=Ambulatory Visit ED=Emergency Department EI=Emergency Department Admit to Inpatient Hospital Stay (permissible substitution) IP=Inpatient Hospital Stay IS=Non-Acute Institutional Stay OS=Observation Stay IC=Institutional Professional Consult (permissible substitution) OA=Other Ambulatory Visit NI=No information UN=Unknown OT=Other	Encounter type. **Details of categorical definitions:** 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 (ED): Includes ED encounters that become inpatient stays (in which case inpatient stays would be a separate encounter). Excludes urgent care facility visits. ED claims should be pulled before hospitalization claims to ensure that ED with subsequent admission won't be rolled up in the hospital event. Does not include observation stays, where known (guidance added in v3.1). Emergency Department Admit to Inpatient Hospital Stay: Permissible substitution for preferred state of separate ED and IP records. Only for use with data sources where the individual records for ED and IP cannot be distinguished. 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. Does not include observation stays, where known (guidance added in v3.1). Observation Stay: "Hospital outpatient services given to help the doctor decide if the patient needs to be admitted as an inpatient or can be discharged. Observations services may be given in the emergency department or another area of the hospital." Definition from Medicare, CMS Product No. 11435, https://www.medicare.gov/Pubs/pdf/11435.pdf (new value set item added in v3.1). Institutional Professional Consult: Permissible substitution when services provided by a medical professional cannot be combined with the given encounter record, such as a specialist consult in an inpatient setting; this situation can be common with claims data sources (new value set item added in v3.1). 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 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.	MSCDM v4.0 with modified value set

ENCOUNTER Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
FACILITYID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary local facility code that identifies the hospital or clinic. Used for chart abstraction and validation. FACILITYID can be a true identifier, or a pseudoidentifier with a consistent crosswalk to the true identifier retained by the source data partner.	MSCDM v4.0 with modified field name
DISCHARGE_DISPOSITION	RDBMS Text(2)	SAS Char(2)	A=Discharged alive E=Expired NI=No information UN=Unknown OT=Other	Vital status at discharge.	MSCDM v4.0 with modified field size and value set
DISCHARGE_STATUS	RDBMS Text(2)	SAS Char(2)	AF=Adult Foster Home AL=Assisted Living Facility AM=Against Medical Advice AW=Absent without leave EX=Expired HH=Home Health HO=Home / Self Care HS=Hospice IP=Other Acute Inpatient Hospital NH=Nursing Home (Includes ICF) RH=Rehabilitation Facility RS=Residential Facility SH=Still In Hospital SN=Skilled Nursing Facility NI=No information UN=Unknown OT=Other	Discharge status.	MSCDM v4.0 with modified value set

ENCOUNTER Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
DRG	RDBMS Text(3)	SAS Char(3)	.	3-digit Diagnosis Related Group (DRG). The DRG is used for reimbursement for inpatient encounters. It is a Medicare requirement that combines diagnoses into clinical concepts for billing. Frequently used in observational data analyses.	MSCDM v4.0
DRG_TYPE	RDBMS Text(2)	SAS Char(2)	01=CMS-DRG (old system) 02=MS-DRG (current system) NI=No information UN=Unknown OT=Other	DRG code version.	MSCDM v4.0 with modified field size and value set
ADMITTING_SOURCE	RDBMS Text(2)	SAS Char(2)	AF=Adult Foster Home AL=Assisted Living Facility AV=Ambulatory Visit ED=Emergency Department HH=Home Health HO=Home / Self Care HS=Hospice IP=Other Acute Inpatient Hospital NH=Nursing Home (Includes ICF) RH=Rehabilitation Facility RS=Residential Facility SN=Skilled Nursing Facility NI=No information UN=Unknown OT=Other	Admitting source.	MSCDM v4.0 with modified value set

ENCOUNTER Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RAW_SITEID	RDBMS Text(x)	SAS Char(x)	.	Optional field for locally-defined identifier intended for local use; for example, where a network may have multiple sites contributing to a central data repository. This attribute may be sensitive in certain contexts; the intent is for internal network use only, and not to enable site quality comparisons.	PCORnet
RAW_ENC_TYPE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_DISCHARGE_DISPOSITION	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_DISCHARGE_STATUS	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_DRG_TYPE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_ADMITTING_SOURCE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

4.4. Table: DIAGNOSIS

DIAGNOSIS Domain Description:

Diagnosis codes indicate the results of diagnostic processes and medical coding within healthcare delivery.

Data in this table are expected to be from healthcare-mediated processes and reimbursement drivers.

Relational Integrity:
The DIAGNOSIS table contains one record per DIAGNOISID.

Primary Key: DIAGNOSISID

Foreign Keys:
DIAGNOSIS.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)
DIAGNOSIS.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (one-to-many relationship)

Constraints:
DIAGNOSISID (unique; required, not null)
PATID (required, not null)
ENCOUNTERID (required, not null)
DX (required, not null)
DX_TYPE (required, not null)
DX_SOURCE (required, not null)

DIAGNOSIS Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
DIAGNOSISID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary identifier for each unique record. Does not need to be persistent across refreshes, and may be created by methods such as sequence or GUID.	PCORnet
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier. Used to link across tables.	MSCDM v4.0
ENCOUNTERID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary encounter-level identifier. Used to link across tables.	MSCDM v4.0
ENC_TYPE	RDBMS Text(2)	SAS Char(2)	AV=Ambulatory Visit ED=Emergency Department EI=Emergency Department Admit to Inpatient Hospital Stay (permissible substitution) IP=Inpatient Hospital Stay IS=Non-Acute Institutional Stay OS=Observation Stay IC=Institutional Professional Consult (permissible substitution) OA=Other Ambulatory Visit NI=No information UN=Unknown OT=Other	Please note: This is a field replicated from the ENCOUNTER table. See the ENCOUNTER table for definitions.	MSCDM v4.0 with modified value set
ADMIT_DATE	RDBMS Date	SAS Date (Numeric)	.	Please note: This is a field replicated from the ENCOUNTER table. See the ENCOUNTER table for definitions.	MSCDM v4.0 with modified field name
PROVIDERID	RDBMS Text(x)	SAS Char(x)	.	Please note: This is a field replicated from the ENCOUNTER table. See the ENCOUNTER table for definitions.	MSCDM v4.0

DIAGNOSIS Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
DX	RDBMS Text(18)	SAS Char(18)	.	Diagnosis code. Some codes will contain leading zeroes, and different levels of decimal precision may also be present. This field is a character field, not numeric, to accommodate these coding conventions. (Description updated in v3.1.)	MSCDM v4.0
DX_TYPE	RDBMS Text(2)	SAS Char(2)	09=ICD-9-CM 10=ICD-10-CM 11=ICD-11-CM SM=SNOMED CT NI=No information UN=Unknown OT=Other	Diagnosis code type. We provide values for ICD and SNOMED code types. Other code types will be added as new terminologies are more widely used. Please note: The "Other" category is meant to identify internal use ontologies and codes.	MSCDM v4.0 with modified field name
DX_SOURCE	RDBMS Text(2)	SAS Char(2)	AD=Admitting DI=Discharge FI=Final IN=Interim NI=No information UN=Unknown OT=Other	Classification of diagnosis source. We include these categories to allow some flexibility in implementation. The context is to capture available diagnoses recorded during a specific encounter.	PCORnet
DX_ORIGIN	RDBMS Text(2)	SAS Char(2)	OD=Order BI=Billing CL=Claim NI=No information UN=Unknown OT=Other	Source of the diagnosis information. Billing pertains to internal healthcare processes and data sources. Claim pertains to data from the bill fulfillment, generally data sources held by insurers and other health plans. New field added in v3.1.	PCORnet
PDX	RDBMS Text(2)	SAS Char(2)	P=Principal S=Secondary X=Unable to Classify NI=No information UN=Unknown OT=Other	Principal discharge diagnosis flag.	MSCDM v4.0 with modified field size and value set

DIAGNOSIS Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RAW_DX	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_DX_TYPE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_DX_SOURCE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_PDX	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

4.5. Table: PROCEDURES

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

(Domain description updated in v3.1.)

Relational Integrity:
The PROCEDURES table contains one record per PROCEDURESID.

Primary Key: PROCEDURESID

Foreign Keys:
PROCEDURES.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)
PROCEDURES.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (one-to-many relationship)

Constraints:
PROCEDURESID (unique; required, not null)
PATID (required, not null)
ENCOUNTERID (required, not null)
PX (required, not null)
PX_TYPE (required, not null)

Note: This table uses the plural form of "procedures" because "procedure" (singular) is often a reserved word in RDBMS's.

PROCEDURES Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PROCEDURESID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary identifier for each unique record. Does not need to be persistent across refreshes, and may be created by methods such as sequence or GUID.	PCORnet
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier. Used to link across tables.	MSCDM v4.0
ENCOUNTERID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary encounter-level identifier. Used to link across tables.	MSCDM v4.0

PROCEDURES Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
ENC_TYPE	RDBMS Text(2)	SAS Char(2)	AV=Ambulatory Visit ED=Emergency Department EI=Emergency Department Admit to Inpatient Hospital Stay (permissible substitution) IP=Inpatient Hospital Stay IS=Non-Acute Institutional Stay OS=Observation Stay IC=Institutional Professional Consult (permissible substitution) OA=Other Ambulatory Visit NI=No information UN=Unknown OT=Other	Please note: This is a field replicated from the ENCOUNTER table. See ENCOUNTER table for definitions.	MSCDM v4.0 with modified field name and value set
ADMIT_DATE	RDBMS Date	SAS Date (Numeric)	.	Please note: This is a field replicated from the ENCOUNTER table. See ENCOUNTER table for definitions.	MSCDM v4.0 with modified field name
PROVIDERID	RDBMS Text(x)	SAS Char(x)	.	Please note: This is a field replicated from the ENCOUNTER table. See ENCOUNTER table for definitions.	MSCDM v4.0
PX_DATE	RDBMS Date	SAS Date (Numeric)	.	Date the procedure was performed.	PCORnet
PX	RDBMS Text(11)	SAS Char(11)	.	Procedure code.	MSCDM v4.0

PROCEDURES Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PX_TYPE	RDBMS Text(2)	SAS Char(2)	09=ICD-9-PCS 10=ICD-10-PCS 11=ICD-11-PCS CH = CPT or HCPCS LC=LOINC ND=NDC RE=Revenue NI=No information UN=Unknown OT=Other	Procedure code type. We include a number of code types for flexibility, but the basic requirement that the code refer to a medical procedure remains. Revenue codes are a standard concept in Medicare billing and can be useful for defining care settings. If those codes are available they can be included. Medications administered by clinicians can be captured in billing data and Electronic Health Records (EHRs) as HCPCS procedure codes. Administration (infusion) of chemotherapy is an example. We are now seeing NDCs captured as part of procedures because payers are demanding it for payment authorization. Inclusion of this code type enables those data partners that capture the NDC along with the procedure to include the data. Please note: The "Other" category is meant to identify internal use ontologies and codes. Value set modified in v3.1 to combine CPT and HCPCS codes into one category; also, ICD-9 value set item typo corrected to ICD-9-PCS.	MSCDM v4.0 with modified field name and value set
PX_SOURCE	RDBMS Text(2)	SAS Char(2)	OD=Order BI=Billing CL=Claim NI=No information UN=Unknown OT=Other	Source of the procedure information. Order and billing pertain to internal healthcare processes and data sources. Claim pertains to data from the bill fulfillment, generally data sources held by insurers and other health plans. Length of data types updated in v3.1.	PCORnet

PROCEDURES Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RAW_PX	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_PX_TYPE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

4.6. Table: VITAL

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

Relational Integrity:
The VITAL table contains one record per VITALID.

Primary Key: VITALID

Foreign Keys:
VITAL.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)
VITAL.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero-to-many relationship)

Constraints:
VITALID (unique; required, not null)
PATID (required, not null)
MEASURE_DATE (required, not null)
VITAL_SOURCE (required, not null)

Figure 1. Example of populated VITAL table.

VITALID	PATID	ENCOUNTERID	MEASURE_DATE	MEASURE_TIME	VITAL_SOURCE	HT	WT	DIASTOLIC	SYSTOLIC	ORIGINAL_BMI	BP_POSITION	TOBACCO	TOBACCO_TYPE
f5a9a07a-f910-11e4-a322-1697f925ec7b	123	98765	1/5/2014	13:51	HC	67							
f5a9a2be-f910-11e4-a322-1697f925ec7b	123	98765	1/5/2014	13:52	HC		150						
f5a9a3fe-f910-11e4-a322-1697f925ec7b	123	98765	1/5/2014	13:55	HC			120	80		01		
f5a9a52a-f910-11e4-a322-1697f925ec7b	123	98765	1/5/2014		HC							01	NI
f5a9a822-f910-11e4-a322-1697f925ec7b	123	98765	1/5/2014	14:02	HC			122	86		01		
f5a9a94e-f910-11e4-a322-1697f925ec7b	123		3/22/2014		PR		145.6						
f5a9aa7a-f910-11e4-a322-1697f925ec7b	123	65432	11/30/2014		HC	67							
f5a9ab9c-f910-11e4-a322-1697f925ec7b	123	65432	11/30/2014		HC		149.3						

The encounter basis is optional.

Measurements on the same date are recorded in different records; however, it is permissible to consolidate into one record if none of the measures were repeated.

In this example, no time was recorded for several of the measures. Although preferable to capture time, we recognize that some source data may not include time precision.

More than one blood pressure reading was collected during this encounter on January 5.

Note: Completely fake data example created de novo, not from existing data.

VITAL Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
VITALID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary identifier for each unique VITAL record. Does not need to be persistent across refreshes, and may be created by methods such as sequence or GUID.	PCORnet
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier. Used to link across tables.	MSCDM v4.0
ENCOUNTERID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary encounter-level identifier. This is an optional relationship because not all vital sign measures will be associated with a healthcare encounter. (New description added in v3.1.)	PCORnet
MEASURE_DATE	RDBMS Date	SAS Date (Numeric)	.	Date of vitals measure.	MSCDM v4.0

VITAL Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
MEASURE_TIME	RDBMS Text(5): Format as HH:MI using 24-hour clock and zero-padding for hour and minute	SAS Time (Numeric)	.	Time of vitals measure.	MSCDM v4.0 with modified data type Source of time format: ISO 8601
VITAL_SOURCE	RDBMS Text(2)	SAS Char(2)	PR=Patient-reported PD=Patient device direct feed HC=Healthcare delivery setting HD=Healthcare device direct feed NI=No information UN=Unknown OT=Other	Please note: The "Patient-reported" category can include reporting by patient's family or guardian.	PCORnet
HT	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Height (in inches) measured by standing. Only populated if measure was taken on this date. If missing, this value should be null. Decimal precision is permissible. Number precision and scale updated in v3.1.	MSCDM v4.0
WT	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Weight (in pounds). Only populated if measure was taken on this date. If missing, this value should be null. Decimal precision is permissible. Number precision and scale updated in v3.1.	MSCDM v4.0
DIASTOLIC	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Diastolic blood pressure (in mmHg). Only populated if measure was taken on this date. If missing, this value should be null. Number precision and scale updated in v3.1.	MSCDM v4.0

VITAL Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
SYSTOLIC	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Systolic blood pressure (in mmHg). Only populated if measure was taken on this date. If missing, this value should be null. Number precision and scale updated in v3.1.	MSCDM v4.0
ORIGINAL_BMI	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	BMI if calculated in the source system. Decimal precision is permissible (new guidance added in v3.1). Important: **Please do not calculate BMI during CDM implementation**. This field should only reflect originating source system calculations, if height and weight are not stored in the source. Number precision and scale updated in v3.1.	PCORnet
BP_POSITION	RDBMS Text(2)	SAS Char(2)	01=Sitting 02=Standing 03=Supine NI=No information UN=Unknown OT=Other	Position for orthostatic blood pressure. This value should be null if blood pressure was not measured.	MSCDM v4.0 with modified field name, field size, and value set

VITAL Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
SMOKING	RDBMS Text(2)	SAS Char(2)	01=Current every day smoker 02=Current some day smoker 03=Former smoker 04=Never smoker 05=Smoker, current status unknown 06=Unknown if ever smoked 07=Heavy tobacco smoker 08=Light tobacco smoker NI=No information UN=Unknown OT=Other	Indicator for any form of tobacco **that is smoked**. Per Meaningful Use guidance, "…smoking status includes any form of tobacco that is smoked, but not all tobacco use." "'Light smoker' is interpreted to mean less than 10 cigarettes per day, or an equivalent (but less concretely defined) quantity of cigar or pipe smoke. 'Heavy smoker' is interpreted to mean greater than 10 cigarettes per day or an equivalent (but less concretely defined) quantity of cigar or pipe smoke." "…we understand that a "current every day smoker" or "current some day smoker" is an individual who has smoked at least 100 cigarettes during his/her lifetime and still regularly smokes every day or periodically, yet consistently; a "former smoker" would be an individual who has smoked at least 100 cigarettes during his/her lifetime but does not currently smoke; and a "never smoker" would be an individual who has not smoked 100 or more cigarettes during his/her lifetime." http://www.healthit.gov/sites/default/files/standards-certification/2014-edition-draft-test-procedures/170-314-a-11-smoking-status-2014-test-procedure-draft-v1.0.pdf [retrieved May 11, 2015]	PCORnet Meaningful Use Core Measures 9 of 13, Stage 1 (2014 definition) http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/9_Record_Smoking_Status.pdf [retrieved January 11, 2015]

VITAL Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
TOBACCO	RDBMS Text(2)	SAS Char(2)	01=Current user 02=Never 03=Quit/former user 04=Passive or environmental exposure 06=Not asked NI=No information UN=Unknown OT=Other	Indicator for any form of tobacco.	MSCDM v4.0 with modified field name, field size, and value set
TOBACCO_TYPE	RDBMS Text(2)	SAS Char(2)	01=Smoked tobacco only 02=Non-smoked tobacco only 03=Use of both smoked and non-smoked tobacco products 04=None 05=Use of smoked tobacco but no information about non-smoked tobacco use NI=No information UN=Unknown OT=Other	Type(s) of tobacco used.	MSCDM v4.0 with modified field size and value set
RAW_DIASTOLIC	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to formatting into the PCORnet CDM.	PCORnet
RAW_SYSTOLIC	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to formatting into the PCORnet CDM.	PCORnet
RAW_BP_POSITION	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_SMOKING	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_TOBACCO	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

VITAL Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RAW_TOBACCO_TYPE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

4.7. Table: DISPENSING

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

(Domain description updated in v3.1.)

Relational Integrity:
The DISPENSING table contains one record per DISPENSINGID.

Primary Key: DISPENSINGID

Foreign Keys:
DISPENSING.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)
DISPENSING.PRESCRIBINGID is a foreign key to PRESCRIBING.PRESCRIBINGID (zero-to-many relationship)

Constraints:
DISPENSINGID (unique; required, not null)
PATID (required, not null)
DISPENSE_DATE (required, not null)
NDC (required, not null)

DISPENSING Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
DISPENSINGID	RDBMS Text(x)	SAS Char(x)	·	Arbitrary identifier for each unique record. Does not need to be persistent across refreshes, and may be created by methods such as sequence or GUID.	PCORnet
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier. Used to link across tables.	MSCDM v4.0
PRESCRIBINGID	RDBMS Text(x)	SAS Char(x)	·	This is an optional relationship to the PRESCRIBING table, and may not be generally available. One prescribing order may generate multiple dispensing records.	PCORnet
DISPENSE_DATE	RDBMS Date	SAS Date (Numeric)	.	Dispensing date (as close as possible to date the person received the dispensing).	MSCDM v4.0 with modified field name
NDC	RDBMS Text(11)	SAS Char(11)	.	National Drug Code in the 11-digit, no-dash, HIPAA format. Please expunge any place holders (such as dashes or extra digits).	MSCDM v4.0 with additional guidance
DISPENSE_SUP	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Days supply. Number of days that the medication supports based on the number of doses as reported by the pharmacist. This amount is typically found on the dispensing record. Integer values are expected. Important: Please do not calculate during CDM implementation. This field should only reflect originating source system calculations. Number precision and scale updated in v3.1.	MSCDM v4.0 with modified field name
DISPENSE_AMT	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Number of units (pills, tablets, vials) dispensed. Net amount per NDC per dispensing. This amount is typically found on the dispensing record. Positive values are expected. Important: Please do not calculate during CDM implementation. This field should only reflect originating source system calculations. Number precision and scale updated in v3.1.	MSCDM v4.0 with modified field name

DISPENSING Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RAW_NDC	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	

4.8. Table: LAB_RESULT_CM

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

(Domain description updated in v3.1.)

Relational Integrity:
The LAB_RESULT_CM table contains one record per LAB_RESULT_CM_ID

Primary Key: LAB_RESULT_CM_ID

Foreign Keys:
LAB_RESULT_CM.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)
LAB_RESULT_CM.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero-to-many relationship)

Constraints:
LAB_RESULT_CM_ID (unique; required, not null)
PATID (required, not null)
RESULT_DATE (required, not null)

LAB_RESULT_CM Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
LAB_RESULT_CM_ID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary identifier for each unique LAB_RESULT_CM record. Does not need to be persistent across refreshes, and may be created by methods such as sequence or GUID.	PCORnet
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier. Used to link across tables.	MSCDM v4.0
ENCOUNTERID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary encounter-level identifier used to link across tables. This is an optional field, and should only be populated if the lab was collected as part of a healthcare encounter.	PCORnet (modeled upon VITAL table)
LAB_NAME	RDBMS Text(10)	SAS Char(10)	A1C=Hemoglobin A1c CK=Creatine kinase total CK_MB=Creatine kinase MB CK_MBI=Creatine kinase MB/creatine kinase total CREATININE=Creatinine HGB=Hemoglobin LDL=Low-density lipoprotein INR=International normalized ratio TROP_I=Troponin I cardiac TROP_T_QL=Troponin T cardiac (qualitative) TROP_T_QN=Troponin T cardiac (quantitative) NI=No information UN=Unknown OT=Other	Laboratory result common measure, a categorical identification for the type of test, which is harmonized across all contributing data partners. Please note that it is possible for more than one LOINC® code, CPT code, and/or local code to be associated with one LAB_NAME. Value set modified in v3.1 to add “null value” options.	MSCDM v4.0 with modified field name and subset of categorical values

LAB_RESULT_CM Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
SPECIMEN_SOURCE	RDBMS Text(10)	SAS Char(10)	BLOOD=Blood CSF=Cerebrospinal fluid PLASMA=Plasma PPP=Platelet poor plasma SERUM=Serum SR_PLS=Serum/plasma URINE=Urine NI=No information UN=Unknown OT=Other	Specimen source. All records will have a specimen source; some tests have several possible values for SPECIMEN_SOURCE. Please see the Implementation Guidance for additional details. (Description updated for v3.1.)	MSCDM v4.0 with subset of categorical values
LAB_LOINC	RDBMS Text(10)	SAS Char(10)	.	Logical Observation Identifiers, Names, and Codes (LOINC®) from the Regenstrief Institute. Results with local versions of LOINC codes (e.g., LOINC candidate codes) should be included in the RAW_ table field, but the LOINC variable should be set to missing. Current LOINC codes are from 3-7 characters long but Regenstrief suggests a length of 10 for future growth. The last digit of the LOINC code is a check digit and is always preceded by a hyphen. All parts of the LOINC code, including the hyphen, must be included. Do not pad the LOINC code with leading zeros.	MSCDM v4.0
PRIORITY	RDBMS Text(2)	SAS Char(2)	E=Expedite R=Routine S=Stat NI=No information UN=Unknown OT=Other	Immediacy of test. The intent of this variable is to determine whether the test was obtained as part of routine care or as an emergent/urgent diagnostic test (designated as Stat or Expedite).	MSCDM v4.0 with modified value set and modified field name

LAB_RESULT_CM Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RESULT_LOC	RDBMS Text(2)	SAS Char(2)	L=Lab P=Point of Care NI=No information UN=Unknown OT=Other	Location of the test result. Point of Care locations may include anticoagulation clinic, newborn nursery, finger stick in provider office, or home. The default value is 'L' unless the result is Point of Care. There should not be any missing values.	MSCDM v4.0 with modified value set
LAB_PX	RDBMS Text(11)	SAS Char(11)	.	Optional variable for local and standard procedure codes, used to identify the originating order for the lab test.	MSCDM v4.0 with modified field name
LAB_PX_TYPE	RDBMS Text(2)	SAS Char(2)	09=ICD-9-PCS 10=ICD-10-PCS 11=ICD-11-PCS CH = CPT or HCPCS LC=LOINC ND=NDC RE=Revenue NI=No information UN=Unknown OT=Other	Procedure code type, if applicable. Value set modified in v3.1 to combine CPT and HCPCS codes into one category; also, ICD-9 value set item typo corrected to specify ICD-9-PCS.	MSCDM v4.0 with modified field name and value set
LAB_ORDER_DATE	RDBMS Date	SAS Date (Numeric)	.	Date test was ordered.	MSCDM v4.0 with modified field name
SPECIMEN_DATE	RDBMS Date	SAS Date (Numeric)	.	Date specimen was collected.	MSCDM v4.0 with modified field name
SPECIMEN_TIME	RDBMS Text(5): Format as HH:MI using 24-hour clock and zero-padding for hour and minute	SAS Time (Numeric)	.	Time specimen was collected.	MSCDM v4.0 with modified field name

LAB_RESULT_CM Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RESULT_DATE	RDBMS Date	SAS Date (Numeric)		Result date.	MSCDM v4.0 with modified field name
RESULT_TIME	RDBMS Text(5): Format as HH:MI using 24-hour clock and zero-padding for hour and minute	SAS Time (Numeric)		Result time.	MSCDM v4.0 with modified field name
RESULT_QUAL	RDBMS Text(12)	SAS Char(12)	BORDERLINE POSITIVE NEGATIVE UNDETERMINED NI=No information UN=Unknown OT=Other	Standardized result for qualitative results. This variable should be NI for quantitative results. Please see the Implementation Guidance for additional details and information on acceptable values for each qualitative LAB_NAME (description updated in v3.1).	MSCDM v4.0 with modified field name,value set, and field length
RESULT_NUM	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Standardized/converted result for quantitative results. Please see the Implementation Guidance for additional details (description updated for v3.1). v3.1 modification: SAS data type corrected to Numeric. Number precision and scale updated in v3.1.	MSCDM v4.0 with modified field name

LAB_RESULT_CM Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RESULT_MODIFIER	RDBMS Text(2)	SAS Char(2)	EQ=Equal GE=Greater than or equal to GT=Greater than LE=Less than or equal to LT=Less than TX=Text NI=No information UN=Unknown OT=Other	Modifier for result values.	MSCDM v4.0 with modified field name and value set
RESULT_UNIT	RDBMS Text(11)	SAS Char(11)	.	Converted/standardized units for the result. Please see the Implementation Guidance for additional details (description updated in v3.1).	MSCDM v4.0 with modified field name
NORM_RANGE_LOW	RDBMS Text(10)	SAS Char(10)	.	Lower bound of the normal range assigned by the laboratory. Value should only contain the value of the lower bound. The symbols >, <, >=, <= should be removed. For example, if the normal range for a test is >100 and <300, then "100" should be entered.	MSCDM v4.0
NORM_MODIFIER_LOW	RDBMS Text(2)	SAS Char(2)	EQ=Equal GE=Greater than or equal to GT=Greater than NO=No lower limit NI=No information UN=Unknown OT=Other	Modifier for NORM_RANGE_LOW values. For numeric results one of the following needs to be true: 1) Both MODIFIER_LOW and MODIFIER_HIGH contain EQ (e.g. normal values fall in the range 3-10) 2) MODIFIER_LOW contains GT or GE and MODIFIER_HIGH contains NO (e.g. normal values are >3 with no upper boundary) 3) MODIFIER_HIGH contains LT or LE and MODIFIER_LOW contains NO (e.g. normal values are <=10 with no lower boundary)	MSCDM v4.0 with modified value set and field name

LAB_RESULT_CM Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
NORM_RANGE_HIGH	RDBMS Text(10)	SAS Char(10)	.	Upper bound of the normal range assigned by the laboratory. Value should only contain the value of the upper bound. The symbols >, <, >=, <= should be removed. For example, if the normal range for a test is >100 and <300, then "300" should be entered.	MSCDM v4.0 with modified field length
NORM_MODIFIER_HIGH	RDBMS Text(2)	SAS Char(2)	EQ=Equal LE=Less than or equal to LT=Less than NO=No higher limit NI=No information UN=Unknown OT=Other	Modifier for NORM_RANGE_HIGH values. For numeric results one of the following needs to be true: 1) Both MODIFIER_LOW and MODIFIER_HIGH contain EQ (e.g. normal values fall in the range 3-10) 2) MODIFIER_LOW contains GT or GE and MODIFIER_HIGH contains NO (e.g. normal values are >3 with no upper boundary) 3) MODIFIER_HIGH contains LT or LE and MODIFIER_LOW contains NO (e.g. normal values are <=10 with no lower boundary)	MSCDM v4.0 with modified value set and field name
ABN_IND	RDBMS Text(2)	SAS Char(2)	AB=Abnormal AH=Abnormally high AL=Abnormally low CH=Critically high CL=Critically low CR=Critical IN=Inconclusive NL=Normal NI=No information UN=Unknown OT=Other	Abnormal result indicator. This value comes from the source data; do not apply logic to create it.	MSCDM v4.0 with modified value set
RAW_LAB_NAME	RDBMS Text(x)	SAS Char(x)	.	Local name related to an individual lab test.	PCORnet
RAW_LAB_CODE	RDBMS Text(x)	SAS Char(x)	.	Local code related to an individual lab test.	PCORnet
RAW_PANEL	RDBMS Text(x)	SAS Char(x)	.	Local code related to a battery or panel of lab tests.	PCORnet

LAB_RESULT_CM Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RAW_RESULT	RDBMS Text(x)	SAS Char(x)	.	The original test result value as seen in your source data. Values may include a decimal point, a sign or text (e.g., POSITIVE, NEGATIVE, DETECTED). The symbols >, <, >=, <= should be removed from the value and stored in the Modifier variable instead.	PCORnet
RAW_UNIT	RDBMS Text(x)	SAS Char(x)	.	Original units for the result in your source data.	PCORnet
RAW_ORDER_DEPT	RDBMS Text(x)	SAS Char(x)	.	Local code for ordering provider department.	PCORnet
RAW_FACILITY_CODE	RDBMS Text(x)	SAS Char(x)	.	Local facility code that identifies the hospital or clinic. Taken from facility claims.	PCORnet

4.9. Table: CONDITION

CONDITION Domain Description:
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.

Relational Integrity:
The CONDITION table contains one record per CONDITIONID.

Primary Key: CONDITIONID

Foreign Keys:
CONDITION.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)
CONDITION.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero-to-many relationship)

Constraints:
CONDITIONID (unique; required, not null)
PATID (required, not null)
CONDITION (required, not null)
CONDITION_TYPE (required, not null)
CONDITION_SOURCE (required, not null)

CONDITION Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
CONDITIONID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary identifier for each unique record. Does not need to be persistent across refreshes, and may be created by methods such as sequence or GUID.	PCORnet
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier. Used to link across tables.	MSCDM v4.0
ENCOUNTERID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary encounter-level identifier used to link across tables. This is an optional field, and should only be populated if the item was collected as part of a healthcare encounter. If more than one encounter association is present, this field should be populated with the ID of the encounter when the condition was first entered into the system. However, please note that many conditions may be recorded outside of an encounter context.	PCORnet (modeled upon VITAL table)
REPORT_DATE	RDBMS Date	SAS Date (Numeric)	.	Date condition was noted, which may be the date when it was recorded by a provider or nurse, or the date on which the patient reported it. Please note that this date may not correspond to onset date.	PCORnet (informed by ESP model)
RESOLVE_DATE	RDBMS Date	SAS Date (Numeric)	.	Date condition was resolved, if resolution of a transient condition has been achieved. A resolution date is not generally expected for chronic conditions, even if the condition is managed.	PCORnet

CONDITION Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
ONSET_DATE	RDBMS Date	SAS Date (Numeric)	.	The onset date concept here refers to "the date and time when problem (illness, disorder, or symptom) started" (ONC:MU Clinical Data Set, caDSR 4973971). This is a different concept than report date, which is the date on which the medical status was collected. An onset date should generally be considered independently of the observer or provider. However, the judgment of when a condition "started" depends on the disease, the frequency of visits, and many other factors. It is not clear that any facility or physician employs this field in a manner which can be trusted without validation during analysis. (New definition added in v3.1.)	PCORnet
CONDITION_STATUS	RDBMS Text(2)	SAS Char(2)	AC=Active RS=Resolved IN=Inactive NI=No information UN=Unknown OT=Other	Condition status corresponding with REPORT_DATE.	PCORnet (informed by ESP model)
CONDITION	RDBMS Text(18)	SAS Char(18)	.	Condition code. Some codes will contain leading zeroes, and different levels of decimal precision may also be present. This field is a character field, not numeric, to accommodate these coding conventions. Please populate the exact value of this diagnosis code, but remove any source-specific suffixes and prefixes. (Description updated in v3.1.)	PCORnet (modeled upon DIAGNOSIS table)

CONDITION Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
CONDITION_TYPE	RDBMS Text(2)	SAS Char(2)	09=ICD-9-CM/PCS 10=ICD-10-CM/PCS 11=ICD-11-CM/PCS SM=SNOMED CT HP=Human Phenotype Ontology AG=Algorithmic NI=No information UN=Unknown OT=Other	Condition code type. Please note: The "Other" category is meant to identify internal use ontologies and codes. Value set items updated to reference both the Clinical Modification (CM) and the Procedure Coding (PCS) context for ICD-9, ICD-10, and ICD-11 (update made in v3.1).	PCORnet (modeled upon DIAGNOSIS table)
CONDITION_SOURCE	RDBMS Text(2)	SAS Char(2)	PR=Patient-reported medical history HC=Healthcare problem list RG=Registry cohort PC=PCORnet-defined condition algorithm NI=No information UN=Unknown OT=Other	Please note: The "Patient-reported" category can include reporting by a proxy, such as patient's family or guardian.	PCORnet (modeled upon VITAL table)
RAW_CONDITION_STATUS	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_CONDITION	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_CONDITION_TYPE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_CONDITION_SOURCE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

4.10. Table: PRO_CM

PRO_CM Domain Description:

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

Relational Integrity:
The PRO_RESPONSE table contains one record per PRO_CM_ID.

Primary Key: PRO_CM_ID

Foreign Keys:
PRO_CM.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)
PRO_CM.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero-to-many relationship)

Constraints:
PRO_CM_ID (unique; required, not null)
PATID (required, not null)
PRO_ITEM (required, not null)
PRO_DATE (required, not null)
PRO_RESPONSE (required, not null)

Note: This table supports the PCORnet Common Measures established by the PCORnet PRO Task Force. Please see the Common Measures Reference Table for information about these measures.

PRO_CM Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PRO_CM_ID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary identifier for each unique record. Does not need to be persistent across refreshes, and may be created by methods such as sequence or GUID.	PCORnet
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier for the patient for whom the PRO response was captured. Used to link across tables.	MSCDM v4.0
ENCOUNTERID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary encounter-level identifier used to link across tables. This is an optional field, and should only be populated if the item was collected as part of a healthcare encounter.	PCORnet (modeled upon VITAL table)
PRO_ITEM	RDBMS Text(20)	SAS Char(20)	PN_0001=GLOBAL01 PN_0002=GLOBAL02 PN_0003=GLOBAL06 PN_0004=PFA53 PN_0005=EDDEP29 PN_0006=HI7 PN_0007=SLEEP20 PN_0008=SRPPER11_CAPS PN_0009=PAININ9 PN_0010=3793R1 PN_0011=28676R1 PN_0012=EOS_P_011 PN_0013=PEDSGLOBAL2 PN_0014=PEDSGLOBAL5 PN_0015=PEDSGLOBAL6 PN_0016=GLOBAL03 PN_0017=GLOBAL04 PN_0018=EDANX53 PN_0019=SAMHSA PN_0020=CAHPS 4.0 PN_0021=PA070	PCORnet identifier for the specific Common Measure item. Please see the Common Measures Reference Table for more details. Character length increased in v3.1 to accommodate the potential for additional ontologies.	PCORnet

PRO_CM Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PRO_LOINC	RDBMS Text(10)	SAS Char(10)	.	LOINC® code for item context and stem. Please see the Implementation Guidance for known LOINC codes for each common measure (description updated in v3.1). Logical Observation Identifiers, Names, and Codes (LOINC) from the Regenstrief Institute. Results with local versions of LOINC codes (e.g., LOINC candidate codes) should be included in the RAW_ table field, but the PRO_LOINC variable should be set to missing. Current LOINC codes are from 3-7 characters long but Regenstrief suggests a length of 10 for future growth. The last digit of the LOINC code is a check digit and is always preceded by a hyphen. All parts of the LOINC code, including the hyphen, must be included. Do not pad the LOINC code with leading zeros.	PCORnet (modeled on LAB_RESULT_CM table)
PRO_DATE	RDBMS Date	SAS Date (Numeric)	.	The date of the response.	PCORnet
PRO_TIME	RDBMS Text(5): Format as HH:MI using 24-hour clock and zero-padding for hour and minute	SAS Time (Numeric)	.	The time of the response.	PCORnet Source of time format: ISO 8601
PRO_RESPONSE	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	The numeric response recorded for the item. Please see the Common Measures Reference Table, "Value List" column, for the list of numeric valid values for each item. Number precision and scale updated in v3.1.	PCORnet

PRO_CM Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PRO_METHOD	RDBMS Text(2)	SAS Char(2)	PA=Paper EC=Electronic PH=Telephonic IV=Telephonic with interactive voice response (IVR) technology NI=No information UN=Unknown OT=Other	Method of administration. Electronic includes responses captured via a personal or tablet computer, at web kiosks, or via a smartphone.	PCORnet
PRO_MODE	RDBMS Text(2)	SAS Char(2)	SF=Self without assistance SA= Self with assistance PR=Proxy without assistance PA=Proxy with assistance NI=No information UN=Unknown OT=Other	The person who responded on behalf of the patient for whom the response was captured. A proxy report is a measurement based on a report by someone other than the patient reporting as if he or she is the patient, such as a parent responding for a child, or a caregiver responding for an individual unable to report for themselves. Assistance excludes providing interpretation of the patient's response.	PCORnet
PRO_CAT	RDBMS Text(2)	SAS Char(2)	Y=Yes N=No NI=No information UN=Unknown OT=Other	Indicates whether Computer Adaptive Testing (CAT) was used to administer the survey or instrument that the item was part of. May apply to electronic (EC) and telephonic (PH or IV) modes.	PCORnet
RAW_PRO_CODE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating code, such as LOINC candidate codes that have not yet been adopted	PCORnet
RAW_PRO_RESPONSE	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

Reference Table: PRO Common Measures

This table is based upon the Final Report from the PCORnet Patient-reported Outcomes (PRO) Common Measures Working Group (CMWG), October 25, 2014.

Domain	*Item Text*	*Value Set*	*PCORnet Unique Identifier*	*Item Code*	*Item Bank*	*PCORnet PRO TF Recommendation*	*LOINC # for Item Bank / Domain*	*LOINC # for Item Context + Stem (2014)*	*SNOMED SCTID for domain (2010)*
General Health	In general, would you say your health is	5=Excellent 4=Very good 3=Good 2=Fair 1=Poor	PN_0001	GLOBAL 01	PROMIS Global	Core Item (Adult and Pediatric)	Parent panel: 75418-4 Panel: 75421-8	61577-3	406221003
Quality of Life	In general, would you say your quality of life is	5=Excellent 4=Very good 3=Good 2=Fair 1=Poor	PN_0002	GLOBAL 02	PROMIS Global	Core Item (Adult and Pediatric)	Parent panel: 75418-4 Panel: 75421-8	61578-1	406221003
Physical Function (alternate)	To what extent are you able to carry out your everyday physical activities such as walking, climbing stairs, carrying groceries, or moving a chair?	5=Completely 4=Mostly 3=Moderately 2=A little 1=Not at all	PN_0003	GLOBAL 06	PROMIS Global	Core Adult Item	Parent panel: 75418-4 Panel: 75421-8	61582-3	Not available
Physical Function	Are you able to run errands and shop?	5=Without any difficulty 4=With a little difficulty 3=With some difficulty 2=With much difficulty 1=Unable to do	PN_0004	PFA53	PROMIS Physical Function	Core Adult Item	Parent panel: 75418-4 Panel: 75421-8	61635-9	364665006
Depression	In the past 7 days…I felt depressed	1=Never 2=Rarely 3=Sometimes 4=Often 5=Always	PN_0005	EDDEP29	PROMIS Emotional Distress-Depression	Core Adult Item	Parent panel: 75418-4 Panel: 75421-8	61967-6	405049007

Reference Table: PRO Common Measures

This table is based upon the Final Report from the PCORnet Patient-reported Outcomes (PRO) Common Measures Working Group (CMWG), October 25, 2014.

Domain	*Item Text*	*Value Set*	*PCORnet Unique Identifier*	*Item Code*	*Item Bank*	*PCORnet PRO TF Recommendation*	*LOINC # for Item Bank / Domain*	*LOINC # for Item Context + Stem (2014)*	*SNOMED SCTID for domain (2010)*
Fatigue	During the past 7 days…I feel fatigued	1=Not at all 2=A little bit 3=Somewhat 4=Quite a bit 5=Very much	PN_0006	HI7	PROMIS Fatigue	Core Adult Item	Parent panel: 75418-4 Panel: 75421-8	61878-5	440398005
Sleep Disturbance	In the past 7 days…I had a problem with my sleep	1=Not at all 2=A little bit 3=Somewhat 4=Quite a bit 5=Very much	PN_0007	SLEEP20	PROMIS Sleep Disturbance	Core Adult Item	Parent panel: 75418-4 Panel: 75421-8	61998-1	248254009
Social Roles & Activities	I have trouble doing all of my regular leisure activities with others	5=Never 4=Rarely 3=Sometimes 2=Usually 1=Always	PN_0008	SRPPER11_CaPS	PROMIS Social Role Participation	Core Adult Item	Parent panel: 75418-4 Panel: 75421-8	75417-6	405073004
Pain Interference	In the past 7 days…How much did pain interfere with your day to day activities?	1=Not at all 2=A little bit 3=Somewhat 4=Quite a bit 5=Very much	PN_0009	PAININ9	PROMIS Pain Interference	Core Adult Item	Parent panel: 75418-4 Panel: 75421-8	61758-9	405160001
Pain Interference	In the past 7 days…I had trouble sleeping when I had pain	0=Never 1=Almost Never 2=Sometimes 3=Often 4=Almost Always	PN_0010	3793R1	PROMIS Peds - Pain Interference	Core Pediatric Item	Parent panel: 75418-4 Panel: 75420-0	62144-1	405160001
Fatigue	In the past 7 days…I got tired easily	0=Never 1=Almost Never 2=Sometimes 3=Often 4=Almost Always	PN_0011	2876R1	PROMIS Peds - Fatigue	Core Pediatric Item	Parent panel: 75418-4 Panel: 75420-0	62104-5	440398005

Reference Table: PRO Common Measures

This table is based upon the Final Report from the PCORnet Patient-reported Outcomes (PRO) Common Measures Working Group (CMWG), October 25, 2014.

Domain	*Item Text*	*Value Set*	*PCORnet Unique Identifier*	*Item Code*	*Item Bank*	*PCORnet PRO TF Recommendation*	*LOINC # for Item Bank / Domain*	*LOINC # for Item Context + Stem (2014)*	*SNOMED SCTID for domain (2010)*
Stress	In the past 7 days…I felt stressed	1=Never 2=Almost Never 3=Sometimes 4=Often 5=Almost Always	PN_0012	EOS_P_011	PROMIS Peds - Stress	Core Pediatric Item	Parent panel: 75418-4 Panel: 75420-0	62095-5	73595000
Depression	How often do you feel really sad	1=Never 2=Rarely 3=Sometimes 4=Often 5=Always	PN_0013	PEDGLOBAL2	PROMIS Peds - Global	Core Pediatric Item	Parent panel: 75418-4 Panel: 75420-0	75416-8	35489007
Peer Relationships	How often do you have fun with friends	1=Never 2=Rarely 3=Sometimes 4=Often 5=Always	PN_0014	PEDGLOBAL5	PROMIS Peds - Global	Core Pediatric Item	Parent panel: 75418-4 Panel: 75420-0	75415-0	225598002
Family Relationships	How often do your parents listen to your ideas?	1=Never 2=Rarely 3=Sometimes 4=Often 5=Always	PN_0015	PEDGLOBAL6	PROMIS Peds - Global	Core Pediatric Item	Parent panel: 75418-4 Panel: 75420-0	75414-3	225598002
Global Physical Health	In general, how would you rate your physical health?	5=Excellent 4=Very good 3=Good 2=Fair 1=Poor	PN_0016	GLOBAL03	PROMIS Global	Recommended Item (Adult and Pediatric)	Parent panel: 75418-4 Panel: 75419-2	61579-9	406221003
Global Mental Health	In general, how would you rate your mental health including your mood and your ability to think?	5=Excellent 4=Very good 3=Good 2=Fair 1=Poor	PN_0017	GLOBAL04	PROMIS Global	Recommended Item (Adult and Pediatric)	Parent panel: 75418-4 Panel: 75419-2	61580-7	406221003

Reference Table: PRO Common Measures

This table is based upon the Final Report from the PCORnet Patient-reported Outcomes (PRO) Common Measures Working Group (CMWG), October 25, 2014.

Domain	*Item Text*	*Value Set*	*PCORnet Unique Identifier*	*Item Code*	*Item Bank*	*PCORnet PRO TF Recommendation*	*LOINC # for Item Bank / Domain*	*LOINC # for Item Context + Stem (2014)*	*SNOMED SCTID for domain (2010)*
Anxiety	In the past 7 days…I felt uneasy	1=Never 2=Rarely 3=Sometimes 4=Often 5=Always	PN_0018	EDANX53	PROMIS Emotional Distress-Anxiety	Recommended Item (Adult and Pediatric)	Parent panel: 75418-4 Panel: 75419-2	61949-4	405051006
Medication Adherence	In the past 7 days…"People often miss a dose of their medicines from time to time. How many days in the past week did you miss taking one or more of your medications?"	Range between 0-7	PN_0019	SAMHSA	Core Psychosocial & Behavioral	Recommended Item (Adult)	Parent panel: 75418-4 Panel: 75419-2	68513-1	418633004
Experience of Care (Evaluation of Care/"Treatment Satisfaction")	Past 12 months…Using any number from 0 to 10, where 0 is the worst health care possible and 10 is the best health care possible, what number would you use to rate all your health care in the last 12 months?	Range between 0-10	PN_0020	CAHPS 4.0	CAHPS Experience of Care	Recommended Item (Adult and Pediatric)	Parent panel: 75418-4 Panel: 75419-2	75412-7	Not available
Life Satisfaction	I have a good life	1=Strongly disagree 2=Disagree 3=Neither agree nor disagree 4=Agree 5=Strongly agree	PN_0021	PA070	Toolbox Psychological Well-Being – Life Satisfaction	Recommended Item (Adult and Pediatric)	Parent panel: 75418-4 Panel: 75419-2	75413-5	405152002

4.11. Table: PRESCRIBING

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

(Domain description updated in v3.1.)

Relational Integrity:
The PRESCRIBING table contains one record per PRESCRIBINGID.

Primary Key: PRESCRIBINGID

Foreign Keys:
PRESCRIBING.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)
PRESCRIBING.ENCOUNTERID is a foreign key to ENCOUNTER.ENCOUNTERID (zero-to-many relationship)

Constraints:
PRESCRIBINGID (unique; required, not null)
PATID (required, not null)

PRESCRIBING Table Specification

Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PRESCRIBINGID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary identifier for each unique PRESCRIBING record. Does not need to be persistent across refreshes, and may be created by methods such as sequence or GUID.	PCORnet
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier used to link across tables.	MSCDM v4.0
ENCOUNTERID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary encounter-level identifier. This is an optional relationship; the ENCOUNTERID should be present if the prescribing activity is directly associated with an encounter.	MSCDM v4.0

PRESCRIBING Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RX_PROVIDERID	RDBMS Text(x)	SAS Char(x)	.	Provider code for the provider who prescribed the medication. The provider code is a pseudoidentifier with a consistent crosswalk to the real identifier.	PCORnet, based upon ENCOUNTER table
RX_ORDER_DATE	RDBMS Date	SAS Date (Numeric)	.	Order date of the prescription by the provider.	MSCDM v4.0
RX_ORDER_TIME	RDBMS Text(5): Format as HH:MI using 24-hour clock and zero-padding for hour and minute	SAS Time (Numeric)	.	Order time of the prescription by the provider.	PCORnet
RX_START_DATE	RDBMS Date	SAS Date (Numeric)	.	Start date of order. This attribute **may not** be consistent with the date on which the patient actually begin taking the medication.	Based on ESP
RX_END_DATE	RDBMS Date	SAS Date (Numeric)	.	End date of order (if available).	Based on ESP
RX_QUANTITY	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Quantity ordered. Number precision and scale updated in v3.1.	Based on OMOP and ESP

PRESCRIBING Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RX_QUANTITY_UNIT	RDBMS Text(2)	SAS Char(2)	PI=Pill TA=Tablet VI=Vial LI=Liquid SO=Solution SU=Suspension OI=Ointment CR=Cream PO=Powder PA=Patch IN=Inhaler KI=Kit DE=Device NI=No information UN=Unknown OT=Other	The unit associated with the quantity prescribed. New field added in v3.1.	PCORnet, based on NDC attributes
RX_REFILLS	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Number of refills ordered (not including the original prescription). If no refills are ordered, the value should be zero. Number precision and scale updated in v3.1.	Based on OMOP and ESP
RX_DAYS_SUPPLY	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Number of days supply ordered, as specified by the prescription. Number precision and scale updated in v3.1.	Based on OMOP

PRESCRIBING Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RX_FREQUENCY	RDBMS Text(2)	SAS Char(2)	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	Specified frequency of medication.	PCORnet
RX_BASIS	RDBMS Text(2)	SAS Char(2)	01=Order to Dispense 02=Order to administer NI=No information UN=Unknown OT=Other	Basis of the medication order. The PRESCRIBING table can contain orders for many different activities, and this field is intended to connect the provider's prescribing order with how the order was fulfilled (such as outpatient dispensing or administration by a healthcare professional). (Value set items updated and field definition expanded in v3.1.)	PCORnet
RXNORM_CUI	RDBMS Text (8)	SAS Char(8)	.	Where an RxNorm mapping exists for the source medication, this field contains the RxNorm concept identifier (CUI) at the highest possible specificity. v3.1 modification: field types changed to character because the National Library of Medicine specifies this variable as a character type.	PCORnet and NLM UMLS
RAW_RX_MED_NAME	RDBMS Text(x)	SAS Char(x)	.	Optional field for originating, full textual medication name from the source.	PCORnet
RAW_RX_FREQUENCY	RDBMS Text(x)	SAS Char(x)		Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

PRESCRIBING Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RAW_RXNORM_CUI	RDBMS Text(x)	SAS Char(x)		Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_RX_QUANTITY	RDBMS Text(x)	SAS Char(x)		Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet
RAW_RX_NDC	RDBMS Text(x)	SAS Char(x)		Optional field for originating value of field, prior to mapping into the PCORnet CDM value set.	PCORnet

4.12. Table: PCORNET_TRIAL

PCORNET_TRIAL Domain Description:
Patients who are enrolled in PCORnet clinical trials.

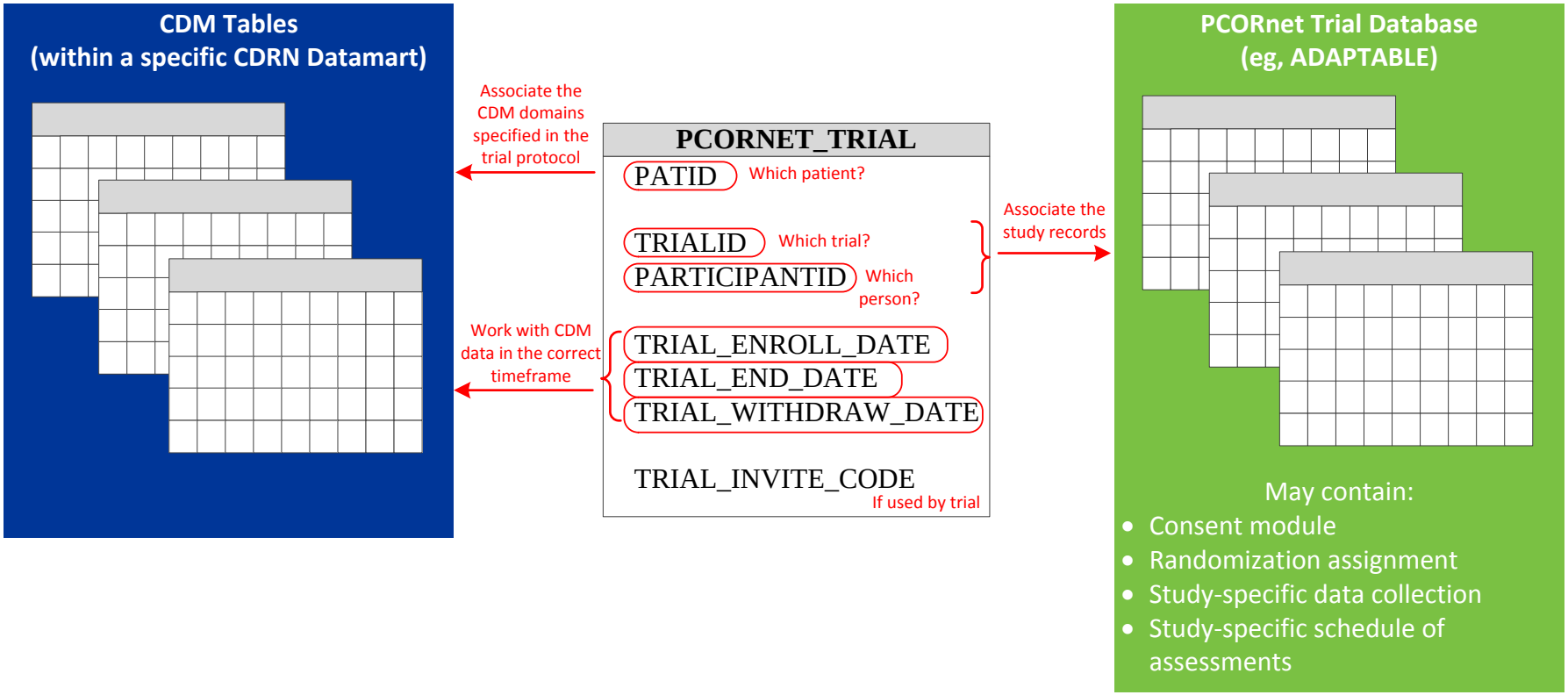
Relational Integrity:
The PCORNET_TRIAL table contains one record per unique combination of PATID, TRIALID, and PARTICIPANTID.

Composite Primary Key: PATID, TRIALID, PARTICIPANTID

Foreign Key:
PCORNET_TRIAL.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many)

Constraints:
PATID + TRIALID + PARTICIPANTID (unique)
PATID (required, not null)
TRIALID (required, not null)
PARTICIPANTID (required, not null)

The PCORNET_TRIAL table serves as a connector and filter for CDM data within the parameters of a given trial protocol:



PCORNET_TRIAL Table Specification					
Field Name	*RDBM Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier used to link across tables.	MSCDM v4.0
TRIALID	RDBMS Text(20)	SAS Char(20)	.	Each TRIALID is assigned by the PCORnet trial's coordinating center.	PCORnet
PARTICIPANTID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier used to uniquely identify a participant in a PCORnet trial. PARTICIPANTID is never repeated or reused for a specific clinical trial, and is generally assigned by trial-specific processes. It may be the same as a randomization ID.	PCORnet
TRIAL_SITEID	RDBMS Text(x)	SAS Char(x)	.	Each TRIAL_SITEID is assigned by the PCORnet trial coordinating center.	PCORnet
TRIAL_ENROLL_DATE	RDBMS Date	SAS Date (Numeric)	.	Date on which the participant enrolled in the trial (generally coincides with trial consent process).	PCORnet
TRIAL_END_DATE	RDBMS Date	SAS Date (Numeric)		Date on which the participant completes participation in the trial.	
TRIAL_WITHDRAW_DATE	RDBMS Date	SAS Date (Numeric)	.	If applicable, date on which the participant withdraws consent from the trial.	PCORnet
TRIAL_INVITE_CODE	RDBMS Text(20)	SAS Char(20)	.	Textual strings used to uniquely identify invitations sent to potential participants, and allows acceptances to be associated back to the originating source. Where used, there should generally be a unique combination of PATID, TRIAL_NAME, and INVITE_CODE within each datamart. For example, this might include "co-enrollment ID strings" for e-mail invites or "verification codes" for letter invites.	PCORnet

4.13. Table: DEATH

DEATH Domain Description:
Reported mortality information for patients.

Relational Integrity:
The DEATH table contains one record per unique combination of PATID and DEATH_SOURCE.

Composite Primary Key: PATID, DEATH_SOURCE

Foreign Key:
DEATH.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)

Constraints (modified in v3.1)
PATID + DEATH_SOURCE (unique)
PATID (required, not null)
DEATH_SOURCE (required, not null)

Note: DEATH_DATE has been removed from the constraints in v3.1 with corresponding changes to relational integrity and the composite primary key.

DEATH Table Specification

Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier used to link across tables.	MSCDM v4.0
DEATH_DATE	RDBMS Date	SAS Date (Numeric)		Date of death.	MSCDM v4.0 with modified field name and data type

DEATH Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
DEATH_DATE_IMPUTE	RDBMS Text(2)	SAS Char(2)	B=Both month and day imputed D=Day imputed M=Month imputed N=Not imputed NI=No information UN=Unknown OT=Other	When date of death is imputed, this field indicates which parts of the date were imputed.	MSCDM v4.0 with modified field name and valueset
DEATH_SOURCE	RDBMS Text(2)	SAS Char(2)	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		MSCDM v4.0 with modified field name and additional guidance
DEATH_MATCH_CONFIDENCE	RDBMS Text(2)	SAS Char(2)	E=Excellent F=Fair P=Poor NI=No information UN=Unknown OT=Other	For situations where a probabilistic patient matching strategy is used, this field indicates the confidence that the patient drawn from external source data represents the actual patient.	MSCDM v4.0 with modified field name and additional guidance

4.14. Table: DEATH_CAUSE

DEATH_CAUSE Domain Description:
The individual causes associated with a reported death.

Relational Integrity:
The DEATH_CAUSE table contains one record per unique combination of PATID, DEATH_CAUSE, DEATH_CAUSE_CODE, DEATH_CAUSE_TYPE, and DEATH_CAUSE_SOURCE.

Composite Primary Key: PATID, DEATH_CAUSE, DEATH_CAUSE_CODE, DEATH_CAUSE_TYPE, DEATH_CAUSE_SOURCE

Foreign Key:
DEATH_CAUSE.PATID is a foreign key to DEMOGRAPHIC.PATID (one-to-many relationship)

Constraints:
PATID + DEATH_CAUSE + DEATH_CAUSE_CODE + DEATH_CAUSE_TYPE + DEATH_CAUSE_SOURCE (unique)
PATID (required, not null)
DEATH_CAUSE (required, not null)
DEATH_CAUSE_CODE (required, not null)
DEATH_CAUSE_TYPE (required, not null)
DEATH_CAUSE_SOURCE (required, not null)

DEATH_CAUSE Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PATID	RDBMS Text(x)	SAS Char(x)	.	Arbitrary person-level identifier used to link across tables.	MSCDM v4.0
DEATH_CAUSE	RDBMS Text(8)	SAS Char(8)	.	Cause of death code. Please include the decimal point in ICD codes (if any).	MSCDM v4.0 with modified field name
DEATH_CAUSE_CODE	RDBMS Text(2)	SAS Char(2)	09=ICD-9 10=ICD-10 NI=No information UN=Unknown OT=Other	Cause of death code type.	MSCDM v4.0 with modified field name
DEATH_CAUSE_TYPE	RDBMS Text(2)	SAS Char(2)	C=Contributory I=Immediate/Primary O=Other U=Underlying NI=No information UN=Unknown OT=Other	Cause of death type. There should be only one underlying cause of death.	MSCDM v4.0 with modified field name
DEATH_CAUSE_SOURCE	RDBMS Text(2)	SAS Char(2)	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	Source of cause of death information.	MSCDM v4.0 with modified field name
DEATH_CAUSE_CONFIDENCE	RDBMS Text(2)	SAS Char(2)	E=Excellent F=Fair P=Poor NI=No information UN=Unknown OT=Other	Confidence in the accuracy of the cause of death based on source, match, number of reporting sources, discrepancies, etc.	MSCDM v4.0 with modified field name

4.15. Table: HARVEST

HARVEST Domain Description:
Attributes associated with the specific PCORnet datamart implementation, including data refreshes.

Relational Integrity:
The HARVEST table contains one record per unique combination of NETWORKID and DATAMARTID.

Composite Primary Key: NETWORKID, DATAMARTID

Constraints:
NETWORKID + DATAMARTID (unique)
NETWORKID (required, not null)
DATAMARTID (required, not null)

Imputation and Obfuscation definitions:

- "No imputation or obfuscation": For any and every date value that is present, no methods of imputation and/or obfuscation have been applied. This does not imply that every record has a date value.
- "Imputation for incomplete dates": Some or all date values were imputed from incomplete dates, but no obfuscation was performed.
- "Date obfuscation": Some or all date values were obfuscated, but no imputation of incomplete dates was performed. Obfuscation can also be called "shifting" or "masking."
- "Both imputation and obfuscation": Some or all date values were imputed, and some or all date values were obfuscated (does not necessarily need to be on the same record).

Imputation refers to the practice of adding day or month precision for incomplete dates (ie, where a specific day or specific month is not present).

Obfuscation, also known as date shifting, is a technique not recommended within PCORnet. However, where this practice exists, this table allows the situation to be recognized for analytic consideration.

HARVEST Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
NETWORKID	RDBMS Text(10)	SAS Char(10)	.	This identifier is assigned by the PCORnet Distributed Research Network Operations Center (DRN OC) (description updated in v3.1)	PCORnet
NETWORK_NAME	RDBMS Text(20)	SAS Char(20)	.	Descriptive name of the network. This identifier is assigned by the PCORnet Distributed Research Network Operations Center (DRN OC) (description updated in v3.1)	PCORnet
DATAMARTID	RDBMS Text(10)	SAS Char(10)	.	This identifier is assigned by the PCORnet Distributed Research Network Operations Center (DRN OC) (description updated in v3.1)	PCORnet
DATAMART_NAME	RDBMS Text(20)	SAS Char(20)	.	Descriptive name of the datamart. This identifier is assigned by the PCORnet Distributed Research Network Operations Center (DRN OC) (description updated in v3.1)	PCORnet
DATAMART_PLATFORM	RDBMS Text(2)	SAS Char(2)	01=SQL Server 02=Oracle 03=PostgreSQL 04=MySQL 05=SAS NI=No information UN=Unknown OT=Other		
CDM_VERSION	RDBMS Number(supports 15 total digits, 8 after decimal)	SAS Numeric(length 8)	.	Version currently implemented within this datamart (for example, 1.0, 2.0, 3.0). Number precision and scale updated in v3.1.	PCORnet

HARVEST Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
DATAMART_CLAIMS	RDBMS Text(2)	SAS Char(2)	01=Not present 02=Present NI=No information UN=Unknown OT=Other	Datamart includes claims data source(s).	PCORnet
DATAMART_EHR	RDBMS Text(2)	SAS Char(2)	01=Not present 02=Present NI=No information UN=Unknown OT=Other	Datamart includes EHR data source(s).	PCORnet
BIRTH_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the BIRTH_DATE field on the DEMOGRAPHIC table. Please see notes for additional definitions.	PCORnet
ENR_START_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the ENR_START_DATE field on the ENROLLMENT table. Please see notes for additional definitions.	PCORnet

HARVEST Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
ENR_END_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the ENR_END_DATE field on the ENROLLMENT table. Please see notes for additional definitions.	PCORnet
ADMIT_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the ADMIT_DATE field on the ENCOUNTER table. Please see notes for additional definitions.	PCORnet
DISCHARGE_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the DISCHARGE_DATE field on the ENCOUNTER table. Please see notes for additional definitions.	PCORnet

HARVEST Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PX_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the PX_DATE field on the PROCEDURES table. Please see notes for additional definitions.	PCORnet
RX_ORDER_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the RX_ORDER_DATE field on the PRESCRIBING table. Please see notes for additional definitions.	PCORnet
RX_START_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the RX_START_DATE field on the PRESCRIBING table. Please see notes for additional definitions.	PCORnet

HARVEST Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
RX_END_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the RX_END_DATE field on the PRESCRIBING table. Please see notes for additional definitions.	PCORnet
DISPENSE_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the DISPENSE_DATE field on the DISPENSING table. Please see notes for additional definitions.	PCORnet
LAB_ORDER_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the LAB_ORDER_DATE field on the LAB_RESULT_CM table. Please see notes for additional definitions.	PCORnet

HARVEST Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
SPECIMEN_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the SPECIMEN_DATE field on the LAB_RESULT_CM table. Please see notes for additional definitions.	PCORnet
RESULT_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the RESULT_DATE field on the LAB_RESULT_CM table. Please see notes for additional definitions.	PCORnet
MEASURE_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the MEASURE_DATE field on the VITAL table. Please see notes for additional definitions.	PCORnet

HARVEST Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
ONSET_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the ONSET_DATE field on the CONDITION table. Please see notes for additional definitions.	PCORnet
REPORT_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the REPORT_DATE field on the CONDITION table. Please see notes for additional definitions.	PCORnet
RESOLVE_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the RESOLVE_DATE field on the CONDITION table. Please see notes for additional definitions.	PCORnet

HARVEST Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
PRO_DATE_MGMT	RDBMS Text(2)	SAS Char(2)	01=No imputation or obfuscation 02=Imputation for incomplete dates 03=Date obfuscation 04=Both imputation and obfuscation NI=No information UN=Unknown OT=Other	Data management strategy currently present in the PRO_DATE field on the PRO_CM table. Please see notes for additional definitions.	PCORnet
REFRESH_DEMOGRAPHIC_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the DEMOGRAPHIC table. This date should be null if the table does not have records.	PCORnet
REFRESH_ENROLLMENT_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the ENROLLMENT table. This date should be null if the table does not have records.	PCORnet
REFRESH_ENCOUNTER_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the ENCOUNTER table. This date should be null if the table does not have records.	PCORnet
REFRESH_DIAGNOSIS_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the DIAGNOSIS table. This date should be null if the table does not have records.	PCORnet
REFRESH_PROCEDURES_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the PROCEDURES table. This date should be null if the table does not have records.	PCORnet
REFRESH_VITAL_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the VITAL table. This date should be null if the table does not have records.	PCORnet

HARVEST Table Specification					
Field Name	*RDBMS Data Type*	*SAS Data Type*	*Predefined Value Sets and Descriptive Text for Categorical Fields*	*Definition / Comments*	*Data Element Provenance*
REFRESH_DISPENSING_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the DISPENSING table. This date should be null if the table does not have records.	PCORnet
REFRESH_LAB_RESULT_CM_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the LAB_RESULT_CM table. This date should be null if the table does not have records.	PCORnet
REFRESH_CONDITION_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the CONDITION table. This date should be null if the table does not have records.	PCORnet
REFRESH_PRO_CM_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the PRO_CM table. This date should be null if the table does not have records.	PCORnet
REFRESH_PRESCRIBING_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the PRESCRIBING table. This date should be null if the table does not have records.	PCORnet
REFRESH_PCORNET_TRIAL_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the PCORNET_TRIAL table. This date should be null if the table does not have records.	PCORnet
REFRESH_DEATH_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the DEATH table. This date should be null if the table does not have records.	PCORnet
REFRESH_DEATH_CAUSE_DATE	RDBMS Date	SAS Date (Numeric)	.	Most recent date on which the present data were loaded into the DEATH_CAUSE table. This date should be null if the table does not have records.	PCORnet